Complex Genetics in Pancreatitis

Insights Gained From a New Candidate Locus Panel

Mitchell A. Ellison, MS,† Daniel M. Spagnolo, PhD,* Celeste Shelton, PhD, CGC,*
Ekatarina Orlova, MS, CGC,*‡ Jessica LaRusch, PhD,*
David C. Whitcomb, MD, PhD,§||¶ and Mark Haupt, MD**

Objectives: Chronic pancreatitis is the end stage of a pathologic inflammatory syndrome with multiple etiological factors, including genetic. We hypothesized that some pancreatitis etiology originates in pancreatic acinar or duct cells and requires both injury and compensatory mechanism failure.

Methods: One hundred pancreatitis patients were assessed using a DNA sequencing panel for pancreatitis. Cooccurrence of variants within and between genes was measured. Gene coexpression was confirmed via published single-cell RNA sequencing.

Results: One hundred and twenty-one variants were identified in 2 or more patients, 15 of which were enriched compared with reference populations. Single cell RNA-sequencing data verified coexpression of *GGT1*, *CFTR*, and *PRSS1* in duct cells, *PRSS1, CPA1, CEL, CTRC*, and *SPINK1* in acinar cells, and *UBR1* in both. Multiple-risk variants with injury/stress effects (*CEL, CFTR, CPA1, PRSS1*) and impaired cell protection (*CTRC, GGT1, SPINK1, UBR1*) cooccur within duct cells, acinar cells, or both.

Conclusions: Pancreatitis is a complex disorder with genetic interactions across genes and cell types. These findings suggest a new, non-Mendelian genetic risk/etiology paradigm where a combination of nonpathogenic genetic risk variants in groups of susceptibility genes and injury/dysfunction response genes contribute to acquired pancreatic disease.

Key Words: pancreatitis, genomics, precision medicine, acinar cell stress, ductal cell stress

Abbreviations: 1KG - 1000 genomes, referring to the 1000 Genomes Project, ACMG - American College of Medical Genetics, AC - allele count, AF - allele frequency, AMR - American haplotype frequency (European ancestry), AN - allele number, AP - acute pancreatitis, DM - diabetes mellitus, CF - cystic fibrosis, CP - chronic pancreatitis, EPI - exocrine pancreatic insufficiency, eQTL - expression quantitative train loci, EUR - European haplotype frequency, FUS - factor of uncertain significance, JBS - Johanson-Blizzard syndrome, LD - linkage disequilibrium, RAP - recurrent acute pancreatitis, scRNA-seq - single-cell RNA sequencing, TF - transcription factor, VUS - variant of uncertain significance

(*Pancreas* 2020;49: 983–998)

Chronic pancreatitis (CP) is the end stage of a progressive, destructive inflammatory disorder of the pancreas. Multiple etiologies follow an unpredictable clinical course with variable complications that include exocrine pancreatic insufficiency, diabetes mellitus, chronic pain syndromes, and risk of pancreatic cancer.[1] Until recently, CP was considered a syndrome of unknown etiology, with inconsistent rates of progression, variable complications, and poor response to therapies.[2] Patients with CP experience among the worst measurements of quality of life of any chronic disease and are often stigmatized by older theories of CP etiology that attribute the disease to severe alcoholism, even if a patient drinks little or no alcohol.[3–5] Optimal management requires identifying the underlying disorders and contributing factors early in the disease course and initiating effective management strategies to prevent progression to irreversible and high-morbidity end stages.[6,7] The intrinsic complexity of the underlying conditions and paucity of simple, accurate diagnostic tools remain major challenges for the field.[8,9]

The etiopathogenesis of CP is now known to be associated with dozens of pathogenic and pathologic factors. Only a small subset of CP patients have a history of very heavy, prolonged alcohol use or a strong family history of pancreatitis or cystic fibrosis (CF) that are suggestive of a Mendelian disorder.[10–12] Instead, most patients have unique combinations of risk factors, such as variable alcohol and smoking use; metabolic factors, such as dyslipidemia or hypercalcemia; and genetic risk variants from over a dozen genes. Furthermore, environmental and genetic factors are associated with modification of disease course and severity.[13–16] These data indicate that CP is a complex syndrome for which the etiology is typically comprised of multiple interacting factors rather than a single cause.

Linkage studies of families with CP identified pathogenic and risk variants in *PRSS1*, *CFTR*, *CEL*, and *UBR1*.[17–25] However, these Mendelian syndromes proved to be rare causes of CP. In contrast, candidate gene and genome-wide association studies (GWAS) identified multiple genetic risk variants, including variants in noncoding regions in *SPINK1*, *CTRC*, *GGT1*, *PRSS1/2* locus, and *CLDN2*.[26,27] Although these variants are common in affected individuals, they are not disease-causing in isolation. This pattern of rare Mendelian genotypes and common risk variants from GWASs is typical of common complex disorders and cannot be explained by clinicopathologic models or Mendelian genetics.[8,28]

Chronic pancreatitis is an acquired syndrome progressing from "at risk" to "end-stage CP" over time, with the pathogenic process often heralded by a sentinel acute pancreatitis (AP) event

From the *Ariel Precision Medicine; †Department of Biological Sciences, ‡Department of Human Genetics, §Department of Cell Biology and Molecular Physiology; ||Department of Human Genetics, and ¶Department of Medicine, University of Pittsburgh Medical Center, Pittsburgh, PA.

Received for publication February 27, 2020; accepted June 5, 2020.

Address correspondence to: Mark Haupt, MD, Suite 270, 5750 Centre Ave, Pittsburgh, PA 15206 (e-mail: mhaupt@arielmedicine.com).

No external funding was used for this study. No funding was received from the National Institutes of Health (NIH) Welcome Trust or the Howard Hughes Medical Institute (HHMI).

M.A.E. and E.O. are consultants for Ariel Precision Medicine; D.M.S., C.S., J.L., M.H. are employees of Ariel Precision Medicine. D.C.W. is cofounder of Ariel Precision Medicine and has consulting agreements with Ariel Precision Medicine.

No external writing support was used in the development of this manuscript.

All authors contributed to study concept and design; acquisition of data; analysis and interpretation of data; drafting of the manuscript; critical revision of the manuscript for important intellectual content. M.A.E. and D.M.S. provided statistical analysis. Study supervision, administrative support, and technical support were provided by M.H.

Supplemental digital contents are available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's Web site (www.pancreasjournal.com).

DOI: 10.1097/MPA.0000000000001612

that activates the immune system and initiates pathology.[1] An AP episode in humans is typically followed by resolution and healing, but about a third of cases develop recurrent AP (RAP), and about a third of RAP patients progress to CP.[29,30] Analyses of patients with AP, RAP, and CP often reveal progressively stronger exposures to environmental stressors (eg, alcohol, smoking, and/or combinations of genetic risk variants).[10,15,16,27,31–35] Furthermore, the function of the risk genes and their regulatory elements appears to be mechanistically linked to processes within the pancreatic acinar cells or duct cells that are related to protection from intrapancreatic trypsin activity, dysfunction of protein quality control (PQC) systems, diminished fluid and bicarbonate secretion, and oxidative stress.[36–39] Thus, the development of AP, and progression to RAP and CP may represent an acquired imbalance where injury or stress overcomes the threshold of protective compensatory mechanisms—especially in patients with genetic risk in 1 or more systems. Under this model, a loss of balance occurs between injury and compensation by either increased stress or decreased compensation, resulting in loss of homeostasis. Eventually, this process leads to clinical signs and symptoms of a pathogenic disorder that may lead to the CP syndrome after the damage becomes irreversible.

Here, we set out to assess the plausibility of a model where stress coupled with failure of compensatory mechanisms leads to injury and thus more severe disease. Under this model, we expect to observe an increased prevalence of pancreatitis patients exhibiting both functional pancreatic (duct and/or acinar) cell genetic variants and stress response genetic variants, compared with a random distribution in accordance with allele frequency (AF). To test this hypothesis, we evaluated the genotypes of 100 consecutive patients with pancreatitis referred to Ariel Precision Medicine (Ariel) (Pittsburgh, Pa) for clinical genetic testing to evaluate the etiology of pancreatitis. Putative molecular partners in pathogenic processes were identified by analyzing intrapatient cooccurrences of variants between 8 established pancreatitis-associated genes, including injury/stress-related genes (*CEL, CFTR, CPA1*, and *PRSSI*) and injury response genes (*CTRC, GGT1, SPINK1, UBR1)*. Cooccurrence of variants in individual patients observed within and between genes was explored in patients diagnosed with AP, RAP, and CP, revealing enriched associations between genetic variants across mechanistic/functional groups, which may be sufficient to cause pancreatitis. Cell-specific gene expression within the pancreas was investigated using publicly available single-cell RNA sequencing (scRNA-Seq) data to confirm feasibility of these findings. This approach provides new insights into pancreatic disease mechanisms and novel methods to evaluate complex data sets and to estimate risk and define disease mechanisms within the context of precision medicine.

MATERIALS AND METHODS

Participants

Ariel is a commercial genetics laboratory/health information technology company. Institutional review board approval was obtained (Western Institutional Review Board; Olympia, Wash) to retrospectively evaluate 100 consecutive patients who underwent clinical genetic testing for pancreatitis using the Ariel *PancreasDx* DNA sequencing panel.

Participant Phenotypic Data

Phenotypic data collected for clinical testing purposes and made available for this analysis include demographic data, ancestry, and relevant patient and family medical history. Before this analysis, all samples were deidentified.

Genetic Sequencing and Variant Interpretation

Sequencing of DNA extracted from saliva samples was performed using the *PancreasDx* next generation DNA sequencing panel. *PancreasDx* evaluates known pancreatitis susceptibility and disease-modifying genes. For the purposes of this study, 8 genetic loci associated with pancreatitis were included in this analysis (*CFTR*, *CPA1*, *CEL*, *CTRC, GGT1*, *PRSS1*, *SPINK1*, and *UBR1*). The sequencing panel was designed to capture all exons and a short distance into introns along with the promoter regions, and deep intronic variants associated with disease. Trans-acting sites associated with tag-single nucleotide polymorphisms (SNPs) were also sequenced (eg, the *PRSS1-PRSS2* haplotype defined by rs6666 and rs6667 in the *PRSS1* coding region).

Genetic variants were classified according the American College of Medical Genetics (ACMG) guidelines for sequence variant interpretation.[40] Common variants considered benign according to ACMG criteria due to AF alone were furthered reviewed as low-to-moderate effect variants according to published case-control studies. Variants significant in 1 or more independent, well-designed case-control study with concordance were classified as "risk variants" or "likely risk variants." The remainder were classified as a "factor of uncertain significance" or "benign."

Association of Variants With Pancreatitis

All variants were organized by chromosome and nucleotide position using version hg38 of the human genome. Alternate allele counts (alleles differing from reference) were totaled at each position and compared with the total allele number of 200 autosomal chromosomes (allele pairs from 100 patients) to compute the 100 patient AF. For each variant, the AF in the 100 patients was compared with the general population AF from publicly available genetic databases including 1000 genomes (1KG), gnomAD, and ExAC. Fisher exact test was used to identify variants present in the pancreatitis cohort at rates higher than expected by chance. *P* values were adjusted using the Bonferroni correction to control for multiple comparisons. Any variants with adjusted *P* values of 0.05 or less were considered significantly enriched.

Cooccurrence Analysis

To identify how often 2 variants cooccur in the same patient and identify statistically associated variant pairs across the 100-patient cohort, we utilized a variant analysis method referred to hereafter as the cooccurrence analysis. All variants identified in at least 2 pancreatitis patients were included in the cooccurrence analysis. Variant *occurrence* data were compiled into a data matrix with unique variants comprising each row and consecutive patients comprising each column. Variants present in a given patient (either heterozygous or homozygous) were assigned a value of 1. Absent variants (nucleotide matching reference genome) were assigned 0. We assumed no missing data based on the deep sequencing of the *PancreasDx* panel for clinical purposes (ie, read depth of >100 at all loci). The sum of the elements of each row in the matrix denotes the number of patients presenting with the variant in question (allele count/2). A variant *cooccurrence* matrix was computed, with rows and columns denoting all 121 observed genetic variants, and cooccurrence counts representing the number of patients containing both of the variants invoked by their respective row and column. Variant *cooccurrence* counts were computed via matrix product of occurrence matrix with its transpose, and the percentage of patients that each variant pair cooccurs in was also calculated. Contingency tables were generated from these counts, and a 1-tailed Fisher exact test was used to test for statistically significant associations between variant pairs.

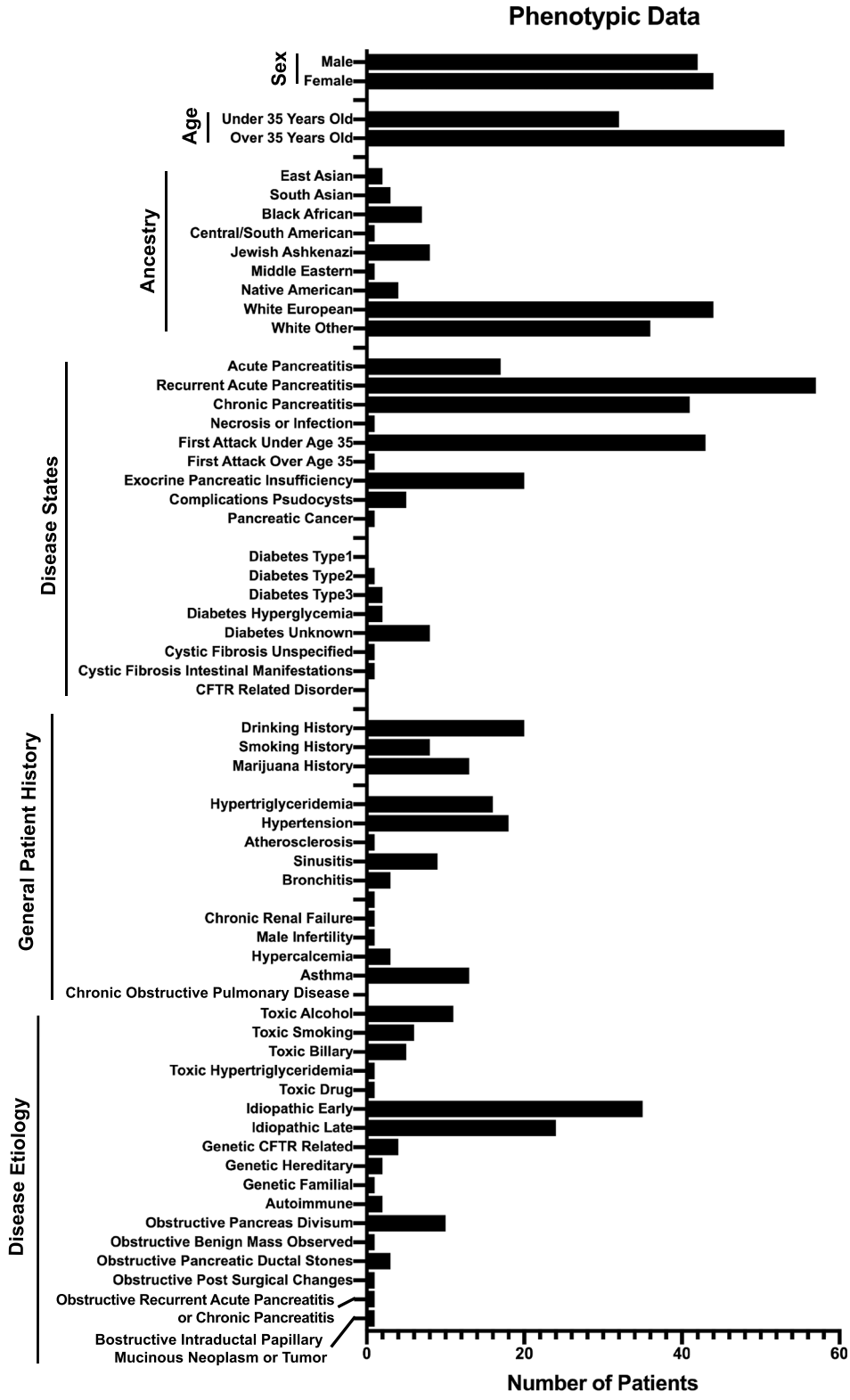

FIGURE 1. Overview of the 100 pancreatitis patient cohort and analysis methods. Phenotypic data for 100 patients described in this study. Bar graphs represent the number of patients falling into each category listed on the *y*-axis.

An effective number of tests was determined using the Li and Ji method, and the Bonferroni correction was done to account for multiple comparisons to yield a nominal *P* value cutoff of 1.6e−4.[41]

Linkage Disequilibrium Statistics

All variants were tested to determine if they were in linkage disequilibrium (LD) by calculating pairwise r^2 values. VCFtools

(Sourceforge, La Jolla, Calif) was used to calculate r^2 values using the -geno-r2 and -ld-window-bp options.[42] All LD calculations were performed using a window of 100,000 bp. Variant pairs with r^2 values greater than 0.8 were considered to be in LD. Analysis of LD between previously published variants (see Discussion section) was conducted using Haploreg4.1.[43]

scRNA-Seq Studies

Gene expression in pancreatic cell types was determined by interrogating publicly available human pancreatic scRNA-Seq data (GSE81547). The chosen data set contains 2544 cells from 8 healthy donors of various ages. The "Seurat" R package (version 3.1.1) was used to log-normalize gene counts.[44,45] Cell type assignments (eg, acinar, etc.) were provided in the Gene Expression Omnibus (GEO) data set, as determined by the authors using flow cell cytometry.[46] Heatmaps of gene expression for the 8 genes in 4 major cell types (acinar, beta, ductal, and mesenchymal) were generated to visualize cell-specific expression data and infer coexpressing genes.

RESULTS

Study Design

The study was designed as a case series with a primary goal of providing a descriptive analysis of patients undergoing clinical genetic testing. To summarize the genetic data, we report common and rare genetic variants and assess combinations of genetic factors according to cell type (eg, duct cell or acinar cell). Using these findings, we make inferences about disease mechanism (eg, CFTR-related disorders, trypsin-related disorders, protein quality defects).

Patient Characteristics

One hundred consecutive patients with AP, RAP, or CP with uncertain etiology underwent clinical DNA sequencing. Phenotypic data reported by patients and their physicians (Fig. 1) indicated that patients in this cohort were evenly distributed between male and female. Over half of the patients were older than 35 years (n = 68), but there were also a large number of younger patients (<35 years old, n = 32). Diverse ancestries were represented in this population, although the majority of patients were of European ancestry (80%).

Patients presented with varying overlapping pancreatitis phenotypes (AP = 17%, RAP = 57%, CP = 41%), with the majority of patients diagnosed with RAP and/or CP. Numerous etiologies were represented in this population. However, the majority of patients were categorized as idiopathic by their ordering physician (59%). Several patients had been diagnosed with diabetes (8%), CF (2%), or cancer (1%) in addition to pancreatitis. Self-reported drug and alcohol use in this population was minimal (alcohol, 20%; tobacco, 8%; marijuana, 13%). Complete population characteristics are presented in Figure 1.

Genetic Variants Identified in the 100-Patient Cohort

For the purpose of these analyses, we limited the genetic variants to those that were identified in at least 2 subjects under the rationale that it would enrich the number of uncommon pathogenic variants and diminish uncommon benign variants, at the cost of eliminating some rare, high-risk variants. We chose not to impose statistical requirements for enrichment of the variants in the primary analysis (eg, 2× expected rate) because the minor AF of low-to-moderate effect risk alleles varies greatly and can exceed 50% (eg, the risk T allele at rs10273639 near *PRSS1/2* has an AF of 0.576).[27]

Of the variants identified in this study, 121 variants were present in 2 or more patients. Genetic variants were detected in all genes tested. Of the 121 variants identified, there were 12 pathogenic variants, 3 variants of uncertain significance (VUS), 5 risk variants, 4 factors of uncertain significance (FUS), and 74 likely benign/benign were identified (Fig. 2). As expected, the 121 variants are largely composed of intronic and protein coding variants because of the use of a targeted a sequencing panel focusing on exons with limited sequence extension into the introns. Most variants (106 [88%] of 121) were no more common than expected by chance in this cohort based on the AF in general populations. However, 15 (12%) of 121 variants were significantly more common in the pancreatitis cohort (Table 1 contains variants with Bonferroni corrected Fisher exact test $P < 0.05$ in at least 1 of the reference populations shown; see Supplementary Fig. 1, http://links.lww.com/MPA/A803, this shows these variants as red data points).

Variants in Multiple Pancreatitis Risk Genes Cooccur Indicating Complex Disease Genetics

Calculating percent cooccurrence for all 121 variants across all patients revealed numerous variants cooccurring at greater than 50% frequency across this population (Fig. 3A, see Supplementary Fig. 2, http://links.lww.com/MPA/A803, which diagrams the steps of the cooccurrence analysis). High intragene percent cooccurrence was observed for all genes in the *PancreasDx* panel, indicating putative haplotypes or common occurrence of common haplotypes. The frequency of intergene percent cooccurrence between variant pairs in *CFTR*, *CPA1*, *PRSS1*, *UBR1*, *GGT1, SPINK1*, and *CTRC* were used to assess possible gene by gene interactions in pancreatitis risk.

To test if variants in this 100-patient cohort were cooccurring more frequently than could be expected by chance, we performed a Fisher exact test on contingency tables generated from pairwise variant combinations. This analysis (Fig. 3B and see Supplementary Fig. 3A, http://links.lww.com/MPA/A803, which shows the heatmap of the cooccurrence analysis) revealed numerous statistically significant intragene cooccurrences in *CFTR*, *CPA1*, *GGT1*, and *CTRC*. Many of these overlapped with the high (>50%) percentage of cooccurrence variants (compare Figs. 3A to B). Significant variant cooccurrence pairs were filtered for variants in known LD ($r^2 \geq 0.8$) (view Supplementary Table 1, http://links.lww.com/MPA/A803, a table containing the variants, view Supplementary Fig. 3B, http://links.lww.com/MPA/A803 which shows the heatmap of LD variants). The remaining significant cooccurring variant pairs, not explained by LD ($r^2 < 0.8$), are provided in Table 2.

Although several intergene variant cooccurrences were biologically compelling and overrepresented in the study cohort, we were unable to identify any statistically significant intergene cooccurrences after correcting our Fisher exact test results for multiple comparisons with the current sample size (Fig. 3B). In summary, these data support the enrichment of uncharacterized haplotypes in patients with pancreatitis and suggest the possibility that numerous genetic interactions exist between variants (Fig. 3C).

Cell-Specific Expression of Risk Genes

To provide a context for the assessment of the observed variant cooccurrences, we analyzed previously published scRNA-seq data collected on pancreatic cells (Fig. 4). These data revealed that observed variant cooccurrences in several genes (variants within and across genes seen together in patients) were consistent with gene coexpression in acinar and/or ductal cells (genes found to

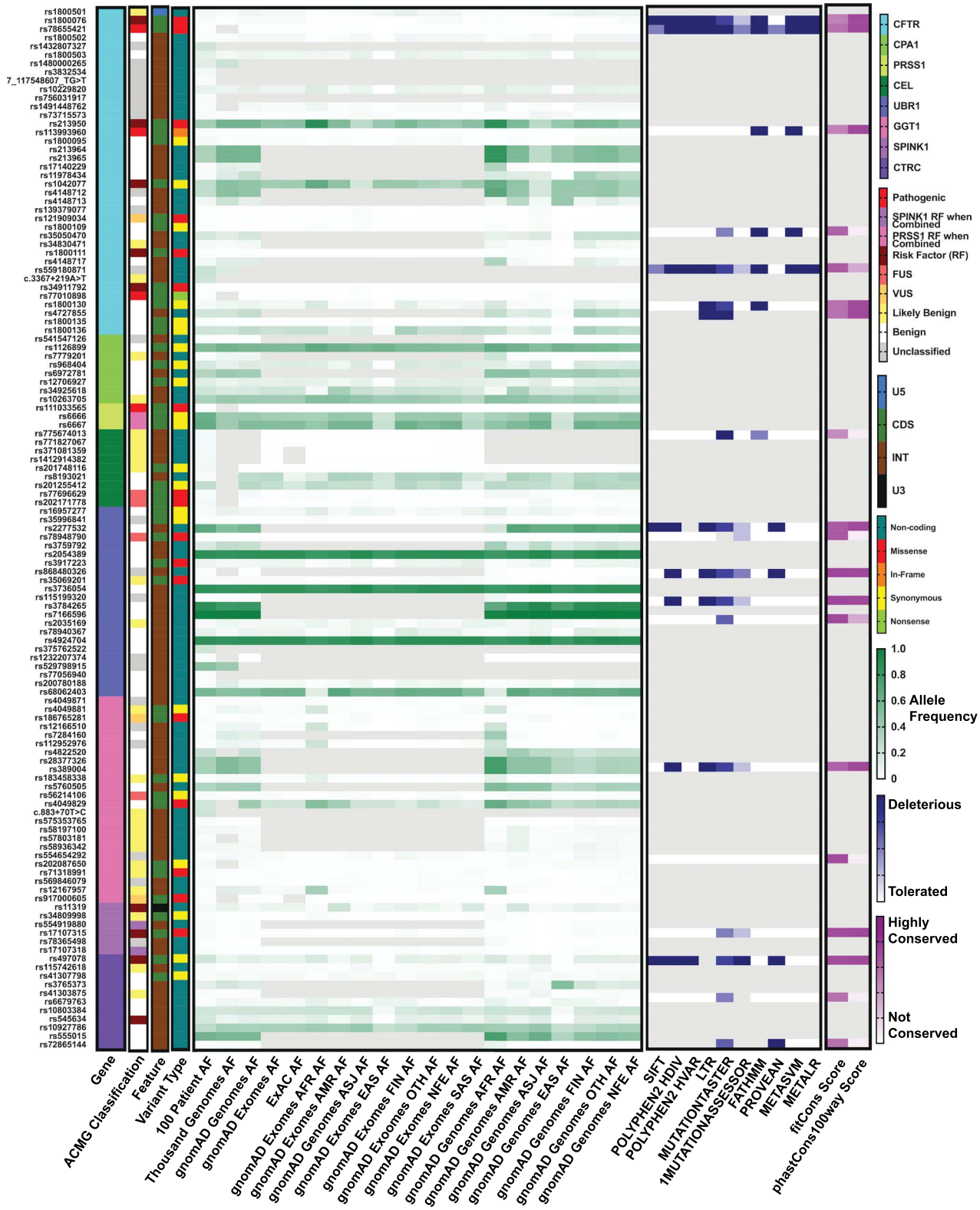

FIGURE 2. Variant type and AF heatmaps. Colored bars on the left-hand side represent genes. From top to bottom: *CFTR*, *CPA1*, *PRSS1*, *CEL*, *UBR1*, *GGT1*, *SPINK1*, and *CTRC*. Categorical data for RNA feature and mutation type are shown to the right of the gene panel followed by the 100 patient allele frequencies and the allele frequencies for various control populations (gnomAD, 1000 Genomes and ExAC) and subpopulations (gnomAD). Results from *in-silico* variant effect prediction algorithms and conservation scores taken from dbNSFP are shown to the right of the AF information.

TABLE 1. Variants Found to Be Significantly Enriched (Bonferroni Adjusted Fisher Exact Test $P < 0.05$) in the 100-Patient Cohort When Compared With at Least 1 of Reference Population Data Set (Exac, Gnomad Genomes, Gnomad Exomes, 1000 Genomes)

Chr	Position	rsID	Gene	Classification	Feature	Amino Acid Change	100-Patient AF	ExAC AF	gnomAD Genomes AF	gnomAD Exome AF	1000 Genomes AF	Adj *P*, ExAC	Adj *P*, gnomAD Genomes	Adj *P*, gnomAD Exomes	Adj *P*, 1000 Genomes
1	15440540	rs497078	*CTRC*	Risk variant	CDS	G60G	1.85E−01	9.30E−02	1.16E−01	9.13E−02	8.29E−02	5.03E−03	3.50E−01	3.44E−03	7.02E−04
1	15443607	rs545634	*CTRC*	Risk variant	INT		1.80E−01	9.00E−02	1.08E−01	8.93E−02	7.29E−02	6.51E−03	1.96E−01	5.12E−03	1.09E−04
7	142751938	rs111033565	*PRSS1*	Pathogenic	CDS	R122H	1.00E−02	4.12E−05	0.00E+00	1.22E−05		6.80E−03	5.05E−03	7.93E−04	
7	142752462	rs6666	*PRSS1*	Risk variant haplotype in homozygotes	CDS	D162D	6.10E−01	3.69E−01	4.12E−01	3.85E−01	3.97E−01	5.58E−10	1.74E−06	1.19E−08	2.35E−07
7	142753014	rs6667	*PRSS1*	Risk variant haplotype in homozygotes	CDS	N246N	6.45E−01	5.04E−01	5.21E−01	5.09E−01	3.94E−01	4.65E−03	3.42E−02	8.64E−03	2.21E−10
9	133064597	rs775674013	*CEL*	Likely benign	INT		3.00E−02	9.51E−04		6.67E−04		7.48E−06		8.90E−07	
9	133064599	rs771827067	*CEL*	Likely benign	INT		3.00E−02	2.48E−05		0.00E+00		1.90E−13		3.33E−17	
9	133064603	rs371081359	*CEL*	Likely benign	INT				8.18E−06				9.38E−16		
9	133068623	rs1412914382	*CEL*	Likely benign	INT				4.08E−06				1.60E−25		
9	133068808	rs201748116	*CEL*	Likely benign	CDS	F347F	4.00E−02	2.27E−03	1.40E−04	1.48E−03		3.58E−06	2.76E−13	1.32E−07	
9	133071003	rs202171778	*CEL*	FUS	CDS	D504H	3.00E−02	2.48E−03	7.42E−03	2.66E−03		1.63E−03	5.20E−01	2.36E−03	
15	42952501	rs2277532	*UBR1*	Benign	INT				5.60E−01		5.10E−01		3.15E−01		2.32E−03
15	43056481	rs78940367	*UBR1*	Benign	INT		9.50E−02	7.10E−02	8.02E−02	7.52E−02	3.33E−02	1.49E+01	3.09E+01	2.11E+01	9.53E−03
22	24627452	rs202087650	*GGT1*	Likely benign	CDS	S347S	3.50E−02	7.83E−04	1.42E−02	1.26E−03		5.73E−08	3.11E+00	1.20E−06	
22	24627670	rs569846079	*GGT1*	Likely benign	INT		2.50E−02	2.57E−03	7.58E−03	9.33E−03	3.79E−03	2.40E−02	2.37E+00	4.94E+00	2.25E−01

AF is a measure of the relative frequency of an allele on a genetic locus in a population. Usually, it is expressed as a proportion or a percentage, here we use proportions.

Chr indicates chromosome; Adj, Bonferroni adjusted Fisher exact test $P < 0.05$; rsID, reference SNP cluster ID; CDS, coding sequence; INT, intronic sequence.

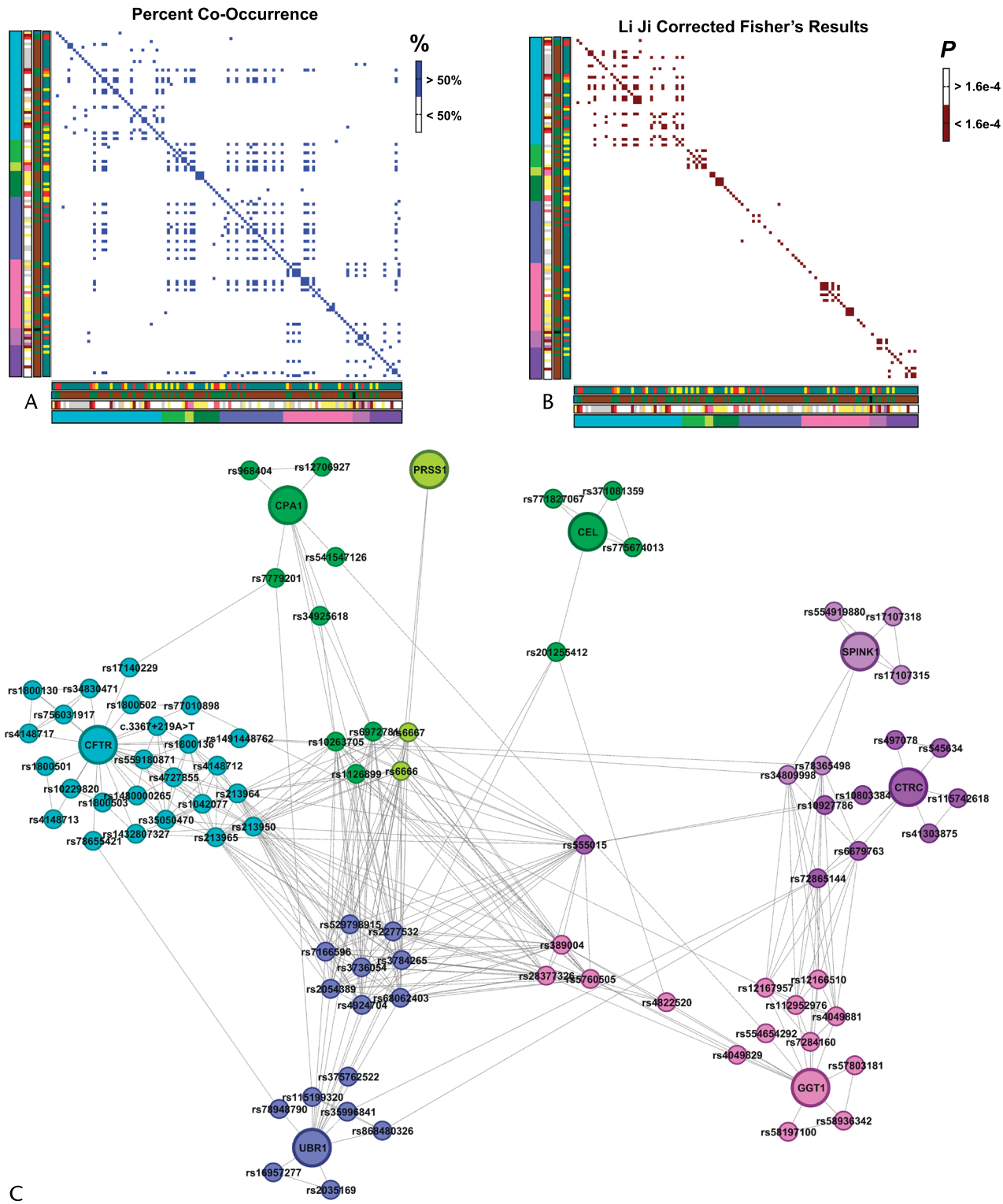

FIGURE 3. Cooccurrence analysis reveals putative intra and intergene genetic interactions. A, Heatmap graphing percent cooccurrence between *CFTR*, *CPA1*, *PRSS1*, *CEL*, *UBR1*, *GGT1*, *SPINK1*, and *CTRC* variants in 100 patients. Only variants with an allele count of 2 were considered in this analysis. Colors of the bars on *x* and *y* axes correspond to gene node colors in C. B, Fisher exact test results for variant cooccurrences shown in A corrected for multiple comparison using the Li Ji method, which resulted in a nominal *P* value cutoff of 1.6e–4.[41] C, Network graph with large nodes representing genes and small nodes representing variants. The graph was generated using only variants that cooccurred in 50% or greater patients. Edge weights were set so that the highest weight was given to interaction between variants and the gene they belong to, followed by intragene cooccurrences and the lowest weight was assigned to intergene cooccurrences. Genes are colored as in Figure 2 and variants belonging to a given gene are colored to match.

TABLE 2. Statistically Significant Intragene Variant Cooccurrences Identified in the 100-Patient Cohort With r^2 <0.80

Variant 1			Variant 2			Statistics		
Gene	ACMG	rsID	Gene	ACMG	rsID	Fisher, *P*	Significant After Li Ji	r^2
CFTR	Likely benign	rs1800501	*CFTR*	Benign	rs4148713	9.88E−06	Yes	NA*
CFTR	Benign	rs1800503	*CFTR*	Risk variant	rs213950	3.45E−06	Yes	0.21
CFTR	Risk variant	rs213950	*CFTR*	Risk variant	rs1042077	2.83E−15	Yes	0.56
CFTR	Risk variant	rs213950	*CFTR*	Benign	rs35050470	1.21E−06	Yes	0.17
CFTR	Risk variant	rs213950	*CFTR*	Benign	rs4727855	2.46E−07	Yes	NA
CFTR	Risk variant	rs213950	*CFTR*	Benign	rs1800136	1.87E−08	Yes	NA
CFTR	Benign	rs213964	*CFTR*	Risk variant	rs1042077	1.26E−11	Yes	0.60
CFTR	Benign	rs213964	*CFTR*	Benign	rs35050470	2.26E−05	Yes	0.09
CFTR	Benign	rs213964	*CFTR*	Benign	rs4727855	2.07E−09	Yes	0.15
CFTR	Benign	rs213964	*CFTR*	Benign	rs1800136	2.97E−07	Yes	0.15
CFTR	Benign	rs213965	*CFTR*	Risk variant	rs1042077	1.26E−11	Yes	0.60
CFTR	Benign	rs213965	*CFTR*	Benign	rs35050470	2.26E−05	Yes	0.09
CFTR	Benign	rs213965	*CFTR*	Benign	rs4727855	2.07E−09	Yes	0.15
CFTR	Benign	rs213965	*CFTR*	Benign	rs1800136	2.97E−07	Yes	0.15
CFTR	Risk variant	rs1042077	*CFTR*	Benign	rs4148713	3.02E−05	Yes	NA
CFTR	Risk variant	rs1042077	*CFTR*	Benign	rs35050470	1.84E−07	Yes	0.28
CFTR	Risk variant	rs1042077	*CFTR*	Benign	rs4727855	1.55E−08	Yes	0.23
CFTR	Risk variant	rs1042077	*CFTR*	Benign	rs1800136	2.38E−10	Yes	0.31
CFTR	Benign	rs35050470	*CFTR*	Likely benign	c.3367+219A>T	1.74E−13	Yes	NA
CFTR	Likely benign	c.3367+219A>T	*CFTR*	Benign	rs4727855	1.05E−10	Yes	NA
CFTR	Likely benign	c.3367+219A>T	*CFTR*	Benign	rs1800136	1.03E−09	Yes	NA
CPA1	Benign	rs1126899	*CPA1*	Benign	rs6972781	5.91E−09	Yes	0.32
CPA1	Benign	rs1126899	*CPA1*	Benign	rs34925618	5.29E−05	Yes	0.10
CPA1	Benign	rs1126899	*CPA1*	Likely benign	rs10263705	1.79E−09	Yes	0.36
CPA1	Benign	rs968404	*CPA1*	Benign	rs12706927	1.00E−11	Yes	0.48
CPA1	Benign	rs6972781	*CPA1*	Benign	rs34925618	3.51E−11	Yes	0.36
CPA1	Benign	rs6972781	*CPA1*	Likely benign	rs10263705	2.05E−23	Yes	0.68
CPA1	Benign	rs34925618	*CPA1*	Likely benign	rs10263705	6.41E−09	Yes	0.20
CEL	Likely benign	rs775674013	*CEL*	Likely benign	rs771827067	8.39E−10	Yes	NA
CEL	Likely benign	rs775674013	*CEL*	Likely benign	rs371081359	8.39E−10	Yes	NA
CEL	Likely benign	rs771827067	*CEL*	Likely benign	rs371081359	8.39E−10	Yes	NA
UBR1	Benign	rs16957277	*UBR1*	Likely benign	rs2035169	6.32E−06	Yes	NA
UBR1	Benign	rs3759792	*UBR1*	Benign	rs3917223	1.76E−09	Yes	0.46
GGT1	Benign	rs4822520	*GGT1*	Benign	rs28377326	3.17E−13	Yes	0.25
GGT1	Benign	rs4822520	*GGT1*	Benign	rs389004	1.09E−14	Yes	0.25
GGT1	Benign	rs4822520	*GGT1*	Benign	rs5760505	1.19E−10	Yes	0.25
GGT1	Benign	rs28377326	*GGT1*	Benign	rs5760505	9.88E−24	Yes	0.77
GGT1	Benign	rs389004	*GGT1*	Benign	rs5760505	1.20E−22	Yes	0.77
GGT1	Likely benign	rs58197100	*GGT1*	Likely benign	rs57803181	3.21E−05	Yes	NA
GGT1	Likely benign	rs58197100	*GGT1*	Likely benign	rs58936342	4.84E−11	Yes	NA
GGT1	Likely benign	rs57803181	*GGT1*	Likely benign	rs58936342	1.79E−05	Yes	NA
SPINK1	Risk variant	rs554919880	*SPINK1*	Risk variant	rs17107315	1.33E−08	Yes	
	Risk variant							NA
SPINK1	Risk variant	rs554919880	*SPINK1*	Risk variant	rs17107318	1.28E−06	Yes	NA
SPINK1	Risk variant	rs17107315	*SPINK1*	Risk variant	rs17107318	1.28E−06	Yes	NA
CTRC	Risk variant	rs497078	*CTRC*	Benign	rs555015	0.000153205	Yes	0.16
CTRC	Likely benign	rs115742618	*CTRC*	Likely benign	rs41303875	2.55E−07	Yes	NA
CTRC	Benign	rs10803384	*CTRC*	Benign	rs10927786	5.50E−08	Yes	0.25
CTRC	Benign	rs10803384	*CTRC*	Benign	rs555015	7.59E−05	Yes	0.06
CTRC	Benign	rs10927786	*CTRC*	Benign	rs555015	3.45E−06	Yes	0.32

*NA is reported for variant pairs where VCFtools[44] returned "nan" indicating that it could not make a reliable estimate of r^2.

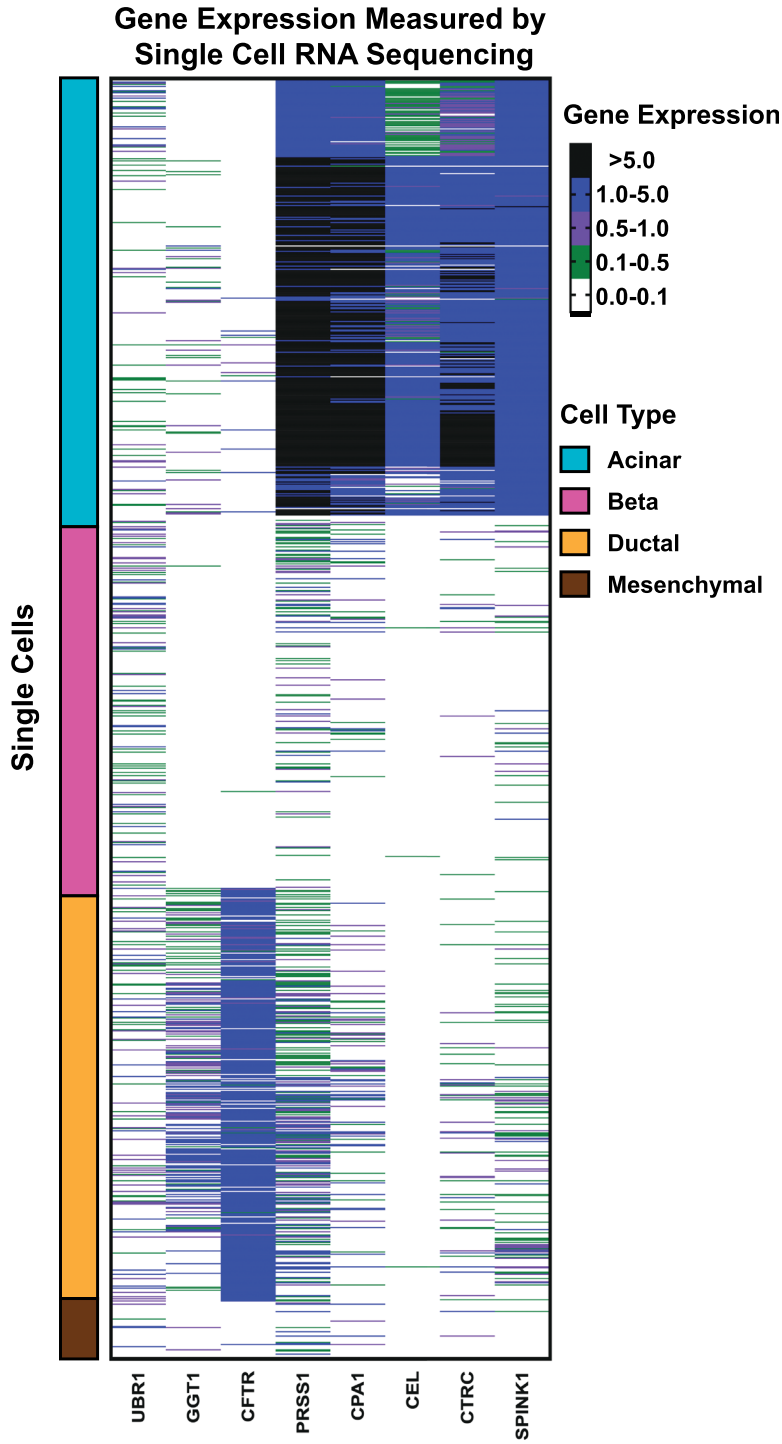

FIGURE 4. Coexpression of pancreatitis-related genes in acinar and ductal cells. Heatmap of log-normalized gene expression levels for *UBR1*, *GGT1*, *CFTR*, *PRSS1*, *CPA1*, *CEL*, *CTRC*, and *SPINK1*.

be synthesized in the same locality). As expected, *PRSS1*, *CPA1*, *CEL*, *CTRC*, and *SPINK1* were generally well expressed in acinar cells and *CFTR* in duct cells compared with the other cell types. Transcripts of *UBR1* and *GGT1* were not well detected in pancreatic acinar cells. In ductal cells, *GGT1* expression was notably higher, and *UBR1* expression was more varied.

Cooccurrence Patterns Change With Pancreatitis Severity

We used percent cooccurrence to test the hypothesis that variants in pancreatitis risk-associated genes more frequently cooccur in patients diagnosed with ongoing disease (CP and RAP) versus

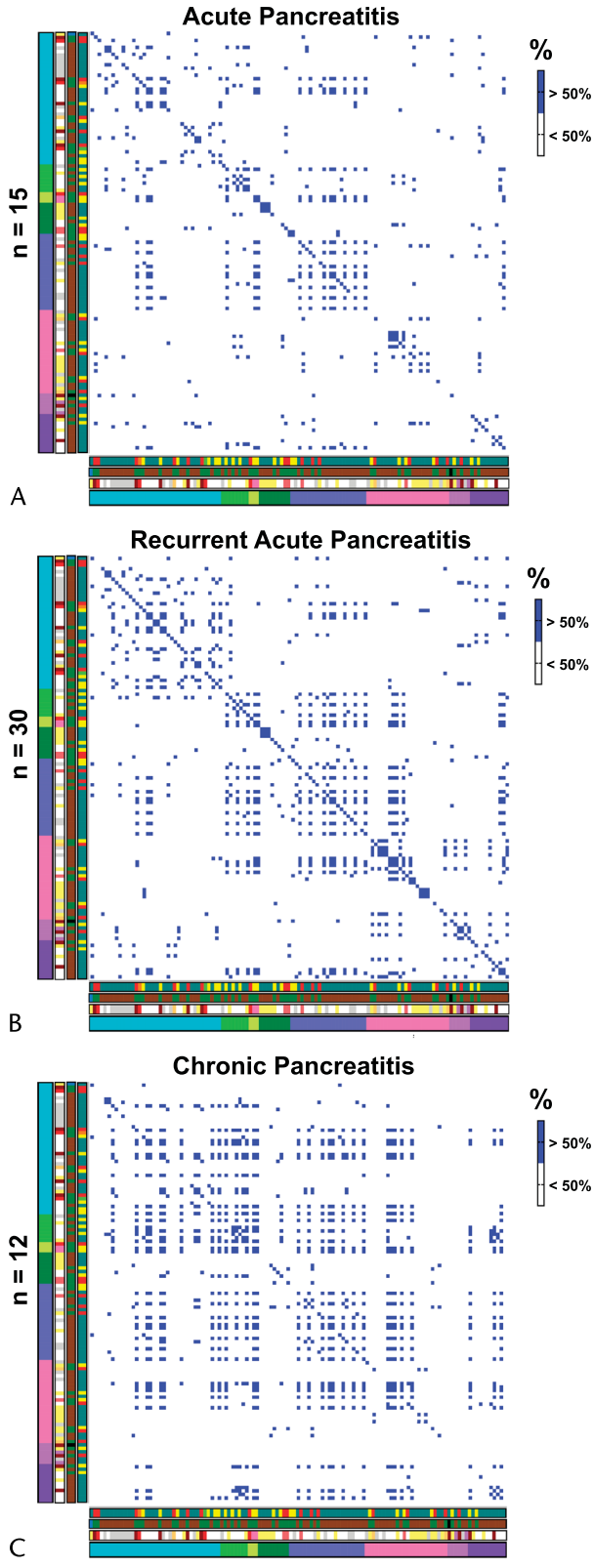

FIGURE 5. Cooccurrence of variants by pancreatitis type. Heatmaps of variants cooccurring greater than 50% of the time in patients diagnosed with (A) acute pancreatitis cooccurrences, (B) recurrent acute pancreatitis, and (C) CP. Genes are colored as in Figure 2 and variants belonging to a given gene are colored to match.

patients with a single episode of pancreatitis (AP). To test this hypothesis, we subset the 100-patient data by diagnosis, excluding patients where the physician had not reported a diagnosis or where multiple pancreatitis diagnoses were reported. The population was divided into the following groups: diagnosis not reported/multiple diagnoses reported (n = 43), AP (n = 15), RAP (n = 30), and CP (n = 12). Cooccurrence analysis was performed on the AP, RAP, and CP subpopulations (Fig. 4).

In patients diagnosed with AP, we observed intragene cooccurrences between variants within nearly all genes tested suggesting ancestry or possibly AP-specific haplotypes (Figs. 5A and 6A). Many high percentage intergene cooccurrences were observed in AP patients (Figs. 5A and B), but far more were seen in patients diagnosed with RAP (Figs. 5B and 6B) and CP (Figs. 5C and 6C). For example, we observe a high degree of cooccurrence between variants in acinar and duct cell genes (*CFTR*, *CPA1*, and *PRSS1*) and stress response genes (*GGT1* and *UBR1*) in CP which is less prominent in patients with AP and RAP. However, because of the lack of statistical power, we were not able to identify many statistically significant interactions between patients once they were split into smaller groups (see Supplementary Fig. 4, http://links.lww.com/MPA/A803 which shows the cooccurrence analysis in specific patient groups).

Cooccurrences between variants were calculated in these disease subgroups (Table 3). Although we observed risk and pathogenic variant intergene cooccurrences in the AP population in multiple genes (*CFTR*, *UBR1*, *GGT1*, *SPINK1*, and *PRSS1*), we were only able to detect intergene cooccurrences between risk variants (nonpathogenic by ACMG criteria) in the CP and RAP subgroups. Very few intergene cooccurrences between risk variants (in *CFTR, PRSS1*, and *CEL*) were identified in the RAP population, indicating that a large number of the cooccurrences we observe in Figure 5B are between benign or likely benign variants using ACMG criteria. Intergene cooccurrences found in the CP group are largely between risk variants present in *PRSS1*, *CFTR*, and *CTRC*. In all groups except for RAP, we observe the cooccurrence of variants in at least 1 acinar- or duct-associated gene with 1 or more genes involved in stress response (*UBR1*, *GGT1*, *SPINK1*, or *CTRC*). This finding suggests that combinations of acinar and ductal cell risk variants and variants in genes encoding proteins involved in stress responses may contribute to the patient's disease phenotype.

Cell-Specific Expression of Risk Genes Supports Their Contribution to Disease Subgroups

In all disease contexts assessed using the cooccurrence analysis, we observed variants cooccurring between genes that are expressed in both acinar and ductal cells (AP: *CFTR*, *UBR1*, *GGT1*, *SPINK1*, and *PRSS1*; RAP: *CFTR, PRSS1*; and *CEL*; CP: *PRSS1*, *CFTR*, and *CTRC*), suggesting that altered function of both acinar and ductal cell types may be necessary to produce pancreatitis in patients without highly penetrant pathogenic variants associated with Mendelian disease (eg, cystic fibrosis, *PRSS1*-hereditary pancreatitis).

DISCUSSION

This study is the first to demonstrate the mechanisms of pancreatitis risk variants in patients with AP, RAP, and CP by cooccurrence of risk variants in susceptibility gene loci (*CFTR*, *CEL*, and *PRSS1*) with variants in *GGT1* and *UBR1* protective/adaptive gene loci. Specifically, we demonstrate cooccurrence of *CFTR* risk variants with *GGT1* risk variants as well as serine protease risk variants and *CFTR* variants with *UBR1* risk variants. Furthermore, this is the first study to demonstrate a role of noncoding *UBR1* variants in the pathogenesis of pancreatic

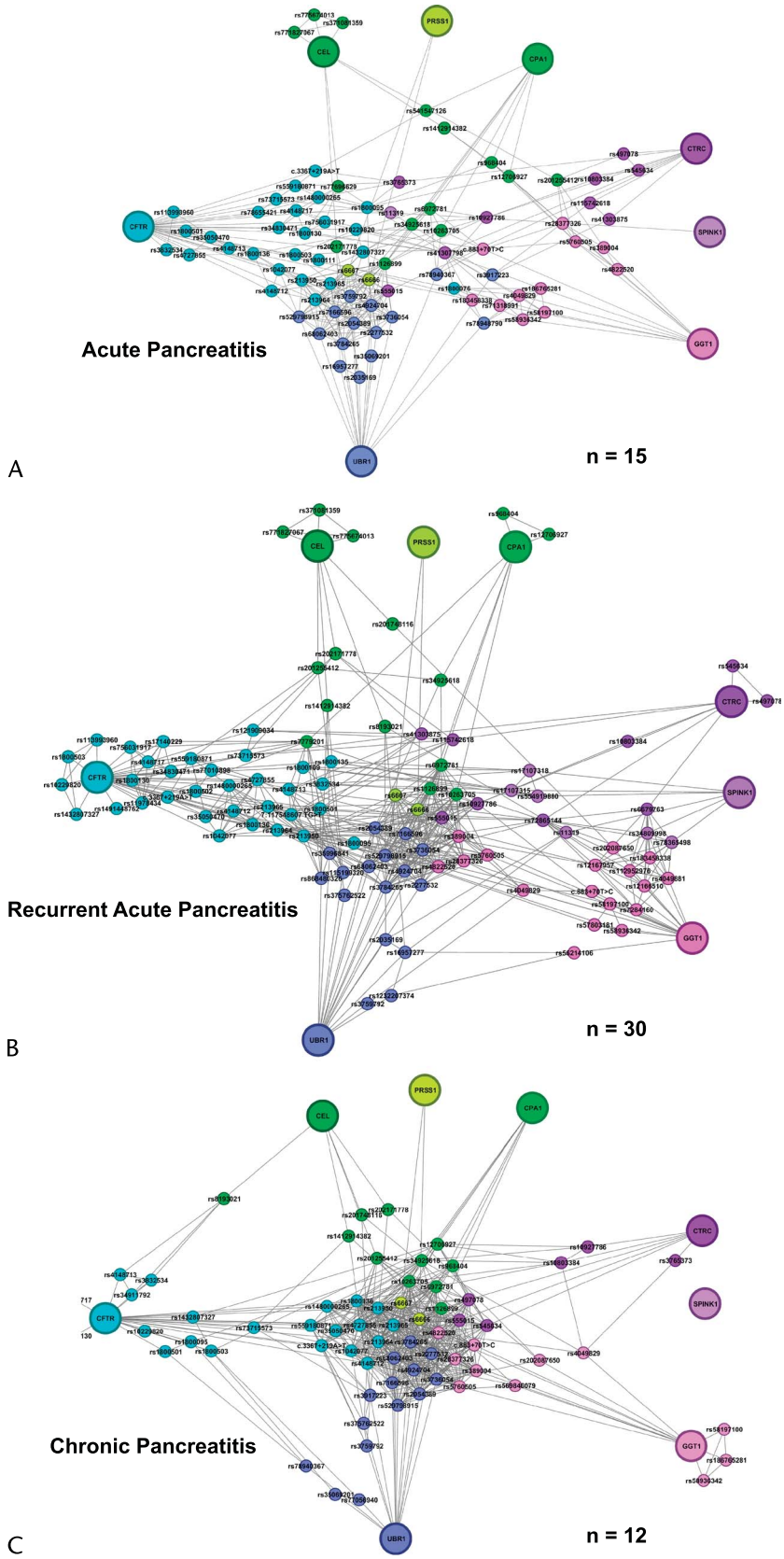

FIGURE 6. Cooccurrence patterns differ between pancreatitis types. Network graphs of variants cooccurring greater than 50% of the time in patients diagnosed with (A) acute pancreatitis cooccurrences, (B) RAP, and (C) CP. Edge weights and node colors as in Figure 3.

TABLE 3. Variants Cooccurring at Over 50% in Each Diagnosis Group (AP, RAP, and CP) Presented in Figure 4

Gene	Classification	rsID	Subgroup AF	gnomAD Exomes AF	Gene	Classification	rsID	Subgroup AF	gnomAD Exomes AF	Percent Cooccurrence	No. Patients With Cooccurrence in Subgroup
AP											
CFTR	Risk variant	rs1800076	0.0333333	0.0153305	*UBR1*	FUS	rs78948790	0.0333333	0.00544111	100	1
CFTR	Risk variant	rs1800076	0.0333333	0.0153305	*GGT1*	VUS	rs186765281	0.0333333	0.0111152	100	1
CFTR	Pathogenic	rs78655421	0.0333333	0.00147328	*SPINK1*	Risk variant	rs11319	0.0666667	0.0833449	50	1
CFTR	Risk variant	rs213950	0.4	0.474443	*PRSS1*	Risk variant haplotype in homozygotes	rs6666	0.533333	0.384531	62	8
CFTR	Risk variant	rs213950	0.4	0.474443	*PRSS1*	Risk variant haplotype in homozygotes	rs6667	0.6	0.50897	64	9
GGT1	VUS	rs186765281	0.0333333	0.0111152	*UBR1*	FUS	rs78948790	0.0333333	0.00544111	100	1
RAP											
CFTR	Risk variant	rs213950	0.383333	0.474443	*PRSS1*	Risk variant haplotype in homozygotes	rs6667	0.716667	0.50897	50	13
CFTR	VUS	rs121909034	0.0166667	0.00107255	*CEL*	FUS	rs202171778	0.0166667	0.0026641	100	1
PRSS1	Risk variant haplotype in homozygotes	rs6667	0.716667	0.50897	*CFTR*	Risk variant	rs213950	0.383333	0.474443	50	13
CP											
CFTR	Risk variant	rs213950	0.333333	0.474443	*PRSS1*	Risk variant haplotype in homozygotes	rs6666	0.541667	0.384531	70	7
CFTR	Risk variant	rs213950	0.333333	0.474443	*PRSS1*	Risk variant haplotype in homozygotes	rs6667	0.541667	0.50897	70	7
CFTR	Risk variant	rs213950	0.333333	0.474443	*CTRC*	Risk variant	rs497078	0.291667	0.0912638	50	4
CFTR	Risk variant	rs213950	0.333333	0.474443	*CTRC*	Risk variant	rs545634	0.291667	0.0893445	50	4
CFTR	Risk variant	rs1042077	0.25	0.382582	*PRSS1*	Risk variant haplotype in homozygotes	rs6666	0.541667	0.384531	50	5
CFTR	Risk variant	rs1042077	0.25	0.382582	*PRSS1*	Risk variant haplotype in homozygotes	rs6667	0.541667	0.50897	50	5
CTRC	Risk variant	rs497078	0.291667	0.0912638	*PRSS1*	Risk variant haplotype in homozygotes	rs6666	0.541667	0.384531	60	6
CTRC	Risk variant	rs497078	0.291667	0.0912638	*PRSS1*	Risk variant haplotype in homozygotes	rs6667	0.541667	0.50897	60	6
CTRC	Risk variant	rs545634	0.291667	0.0893445	*PRSS1*	Risk variant haplotype in homozygotes	rs6666	0.541667	0.384531	60	6
CTRC	Risk variant	rs545634	0.291667	0.0893445	*PRSS1*	Risk variant haplotype in homozygotes	rs6667	0.541667	0.50897	60	6

disease and in RAP/CP beyond Johanson-Blizzard syndrome (JBS).[25,47–49] Additionally, we reveal cooccurrences between variants in *PRSS1* and *CFTR* with variants in *SPINK1* and *CTRC* protective/adaptive genes. These findings suggest a new, non-Mendelian genetic risk/etiology paradigm where a combination of nonpathogenic genetic risk variants in a combination of susceptibility genes and injury/dysfunction response genes contribute to acquired pancreatic disease. Because previous studies have described *SPINK1* and *CTRC* as genetic risk variants for pancreatitis, our discussion will focus on the protective/adaptive genes *GGT1* and *UBR1*.[26,31,50–52]

GGT1-associated Pathophysiology

γ-Glutamyltranspeptidase (GGT), coded for by *GGT1*, is an extracellular enzyme that is primarily expressed in ducts and tubule epithelia (biliary, gall bladder, kidney, epididymis, prostate) and the small intestine where it salvages glutamate, cysteine (Cys), and glycine from glutathione in the ductal fluids.[53,54] It is the only enzyme known to be capable of hydrolyzing the γ-carboxyl-amine bond of reduced glutathione (GSH), thereby supplying Cys to cells which is rate-limiting for intracellular GSH synthesis.[55] Intracellular GSH is present in all mammalian tissues where it counters oxidative stress, regulates redox signaling, plays a vital role in detoxification of xenobiotics, and regulates cell proliferation, apoptosis, immune function, and fibrogenesis.[55,56]

We confirmed that GGT1 transcripts are expressed in the pancreas, and specifically in the pancreatic duct cells using scRNA-Seq data. The GGT1 protein is known to be present in pancreatic exocrine tissue, especially under stress and tumors that express high levels of GGT1 are resistant to chemotherapy and apoptosis, demonstrating its important role in oxidative stress and cell metabolism.[57,58] Knockout mice demonstrate that GGT1 is essential to maintaining Cys levels in the body by intestinal absorption and reclamation in the proximal kidney tubules. The GGT1-deficient mice have stunted growth with mortality within about 12 weeks due to Cys deficiency, and that they have marked reduction of GSH levels in the liver and pancreas, which can be rescued by addition of N-acetylcysteine to their drinking water as an alternate source of Cys.[54]

The role of *GGT1* as a risk factor for RAP/CP was first reported by Brand et al[59] using a candidate gene approach in patients (n = 496) and control subjects (n = 465) from the initial North American Pancreatitis II cohort as an extension of a pool-based GWAS implicating *GGT1* in pancreatic adenocarcinoma.[59,60] The risk-associated variants in this study were noncoding tag SNPs; rs4820599, rs8135987, and rs2017869 (pancreatic cancer). The rs4820599 and rs2017869 variants are part of larger haplotypes, and these variants are also in linkage with each other (r^2 = 0.77). Another variant, rs5751901, is associated with AP risk, and this variant is also linked to rs2017869 risk haplotype (r^2 = 0.89).[61,62] The recent study of *GGT1* variants in AP by Ściskalska et al[61] found that *GGT1* rs5751901 is associated with lower GGT1 serum activity in smokers with the CC genotype and that there is an increased risk of acute AP in smokers with the TC and CC genotypes. This variant was previously linked to serum GGT1 levels in unstimulated lymphocytes of fasting adults with each minor allele of rs5751901 associated with a 21% increase in GGT1 protein levels; however, the authors cautioned that lymphocytes are not the most relevant tissue, and the unstimulated effects may not translate to stimulated cells.[63] The third SNP, reported by Diergaardeand et al[60] and Brand et al,[59] is rs8135987, which is in linkage with rs5760505 (r^2 = 0.82) from this study and is part of another risk haplotype with multiple expression quantitative trait loci (eQTL) including altered *GGT1* expression in the pancreas.[62,64] By identifying additional *GGT1* risk variants we corroborate the findings of Brand et al[59] and Ściskalska et al[61] and supporting the role of *GGT1* in pancreatitis pathophysiology.

Other *GGT1* variants identified in this study are rare, but possibly functional because they are part of eQTLs that alter regulatory elements. The variant rs186765281, which is not part of a risk haplotype, results in both a missense variant, p.I93V, and alters protein binding site for USF1, a transcription factor (TF) that regulates several gene networks, including the stress and immune responses.[65] This variant was found to cooccur with *CFTR* p.R75Q (rs1800076, Table 3), a risk factor for pancreatitis through altered bicarbonate conductance.[34,66] The *GGT1* rs4822520 variant is an intronic SNP that is not part of a major haplotype; however, it alters the nucleotide binding motif of RFX5, a TF associated with immune response and insulin secretion resulting in altered *GGT1* expression in the pancreas.[64,67] The 2-variant intronic haplotype in our study, comprised of rs28377326 and rs389004, alter TF binding sites of several factors including MYC, a family of TFs involved in metabolism, cellular division, differentiation, and cell death. These variants result in altered *GGT1* expression and were identified as an eQTL in multiple tissues including the pancreas.[62] Two very rare variants that were enriched in our study (Table 1), rs202087650 (p.S347S found in an altered HNF4 binding motif) and rs569846079 (intronic), may also be risk factors, but enrichment by chance could not be excluded. Together, these data suggest that *GGT1* SNPs identified in this study alter known TF binding motifs in regions of the genome marked by histone modifications indicative of active chromatin resulting in GGT1 eQTLs.

Our findings here of strong cooccurrence of 2 or more nonlinked *GGT1* risk variants or *GGT1* risk variants with *CFTR* risk variants, combined with our finding of coexpression of *CFTR* and *GGT1* in pancreatic duct cells, provide biologically plausible mechanism for ongoing pathologic signaling from the duct cell to drive RAP/CP. Thus, the *GGT1* variants are not known to be "pathogenic," as in Mendelian genetic disease models, but may be significant risk factors in the context of ongoing injury or stress in the pancreas, especially in the duct cells.

UBR1-Associated Pathophysiology

A key factor in the PQC system is UBR1, which serves as a cytoplasmic and nuclear "garbage collector" of mislocalized and misfolded proteins. The UBR1 protein is an E3 ubiquitin ligase that identifies misfolded/unfolded proteins or proteins misdirected to the cytoplasm by recognition of an N-terminal signal and then links them to ubiquitin via E2 and E1 ligases where they are either directed to the proteasome to be hydrolyzed to amino acids or, if they represent larger aggregates of insoluble proteins, directly to the autophagosome for hydrolysis by lysosomal enzymes.[43] The importance of UBR1 is demonstrated by JBS, a rare, congenital, multisystem autosomal recessive disorder characterized by exocrine pancreatic insufficiency, typical facial features, dental anomalies, hypothyroidism, sensorineural hearing loss, scalp defects, urogenital and anorectal anomalies, short stature, and cognitive impairment of variable degree.[25,47,48] Exocrine insufficiency is also seen in Ubr1(−/−) mice, with impaired stimulus-secretion coupling and increased susceptibility to pancreatic injury.[25] These data indicate that the pancreas is especially sensitive to the loss of UBR1 and that the mechanism being disrupted involves the PQC mechanism linked to secretory proteins, such as digestive enzymes and, possibly, mitochondrial dysfunction—both known to be important in the pathophysiology of pancreatitis.[68,69]

The current study is the first to implicate *UBR1* variants in pancreatic disease outside of JBS. The complete sequencing of

the coding regions, promotor, and into the introns conducted here revealed very few exonic variants; however, many rare and common intronic variants were detected. Although it is not possible to demonstrate the independent deleterious effect of each of these variants, the following evidence suggest the possibility that combinations of these variants could be deleterious. First, the variants are within the *UBR1* locus which codes for an important protein in pancreatic physiology. Second, they were identified in a series of pancreatitis patients with DNA sent for sequencing because of their clinical symptoms and lack of a known etiology, increasing the pretest probability that variants appearing more commonly in this cohort than expected based on minor AF are risk alleles. Third, many of the variants directly, or indirectly, as part of a larger haplotype, disrupts either an important TF binding motif (eg, STAT, rs35996841 and rs3784265; TATA, rs115199320; NfkB, rs68062403), exists in a region of open chromatin (DNase hypersensitivity site), or promoter or enhancer chromatin (marked by histone modifications H3K4me3 and H3K27ac, respectively), and/or has established eQTL characteristics (rs3736054, rs3784265, rs4924704, rs68062403, rs2277532, rs2035169, rs16957277), suggesting deleteriousness. Fourth, the pattern of cooccurrence is consistent with predicted deleteriousness. Clusters of variants in Figures 3C and 4B, D, F differ in minor AF, with variants (nodes) participating in intergene interactions (edges center of the networks) often being more common and variants that do not participate in intergene cooccurrences (near *UBR1* node) often being more severe. It is possible that the cooccurrence reflects linkage within a haplotype, but only rs3784265 and rs4924704 appear to be in LD. Thus, the cooccurrence of these risk variants more likely reflects "recessive" genotypes (ie, both alleles affected as a compound heterozygous risk). The pattern is also instructive in that the cooccurrence of the *UBR1* variants are found with susceptibility variants in *PRSS1*, *CTRC*, *CEL*, and *CFTR*, but less frequently with other protective variants; *SPINK1*, *CTRC*, and *GGT1* and this pattern would not be expected if the associations were random. Finally, the *PRSS1/2* locus risk haplotype (ie, tagged by rs6667) is associated with higher risk from higher trypsinogen levels, and these cooccur with *UBR1* variants predicted to have diminished function, suggesting effects in both the trypsin-dependent pathway and misfolded protein pathway.[27,36,37]

The model of *UBR1* genetic risk variants being deleterious only in the context of other susceptibility genetic variants and recessive *UBR1* genotypes provides insights into studies of *UBR1* sequencing as a common cause of RAP/CP. Masamune et al[49] performed exon sequences of the *UBR1* gene from 389 Japanese patients with CP (188 idiopathic, 172 alcoholic, 20 hereditary, 9 familial). Fifteen nonsynonymous variants including 3 novel ones (c.4514 T > C [p.I1505T], c.4828C > G [p.H1610D], and c.4856A > T [p.D1619V]) and 2 synonymous variants were identified in the exonic regions. They also noted that 7 CP patients had variants in other genes, including *PRSS1*, *SPINK1*, *CTRC*, and *CFTR*. They did not consider noncoding variants and evaluated the data using a Mendelian model. They concluded that no *UBR1* variant caused pancreatitis, but they did not consider dysregulation of *UBR1* as part of a complex disorder as done here.

Implications in Precision Medicine

The strength of our approach includes a focus on a limited number of genetic loci that are known to harbor risk alleles for RAP/CP, the observation of a subset of variants at frequencies that were greater than expected by chance, and the biological plausibility of combined genetic variants in injury/stress response genes impairing protection/adaptation. Additional evidence comes from stepwise enrichment of cooccurring variants in progressively more homogenous etiological and higher-risk patients between AP, RAP, and CP (Figs. 6A, B, and C). The ability to detect this relationship may also be driven by patient preselection by referring physicians who sought advanced genomics and analysis in patients who had already been thoroughly evaluated for pancreatitis etiologies and remained idiopathic.

The discipline of precision medicine for acquired complex common disorders focuses on determining patient-specific gene and environment-associated disorders causing early symptoms of disease, and directing patients to the right treatment to mitigate the problem or prevent the progression to end-stage disease.[8] International consensus efforts in clinical-translational pancreatology provide the groundwork for applying these principles to patients with RAP and CP using updated disease definitions, standardization of risk/etiology lists, and progressive disease models.[1,6,70–73] Furthermore, many important genetic variants beget disease susceptibility by altering regulatory elements, suggesting that pathologic risk may be heightened by failure of programmed stress responses orchestrated through canonical TFs.[27,31,59,74] The work presented here demonstrates the complicated genetics underlying pancreatitis with examples of how multiple genetic changes in key genes in the acinar and ductal cell can be combined with changes to key stress response genes within patients resulting in idiopathic pancreatitis.

REFERENCES

1. Whitcomb DC, Frulloni L, Garg P, et al. Chronic pancreatitis: an international draft consensus proposal for a new mechanistic definition. *Pancreatology*. 2016;16:218–224.
2. Steer ML, Waxman I, Freedman S. Chronic pancreatitis. *N Engl J Med*. 1995;332:1482–1490.
3. Amann ST, Yadav D, Barmada MM, et al. Physical and mental quality of life in chronic pancreatitis: a case-control study from the North American Pancreatitis Study 2 cohort. *Pancreas*. 2013;42:293–300.
4. Machicado JD, Amann ST, Anderson MA, et al. Quality of life in chronic pancreatitis is determined by constant pain, disability/unemployment, current smoking, and associated co-morbidities. *Am J Gastroenterol*. 2017; 112:633–642.
5. Mullady DK, Yadav D, Amann ST, et al. Type of pain, pain-associated complications, quality of life, disability and resource utilisation in chronic pancreatitis: a prospective cohort study. *Gut*. 2011;60:77–84.
6. Whitcomb DC, Shimosegawa T, Chari ST, et al. International consensus statements on early chronic pancreatitis. Recommendations from the working group for the international consensus guidelines for chronic pancreatitis in collaboration with the International Association of Pancreatology, American Pancreatic Association, Japan Pancreas Society, PancreasFest Working Group and European Pancreatic Club. *Pancreatology*. 2018;18:516–527.
7. Lowe ME, Goodman MT, Coté GA, et al. Accelerating the drug delivery pipeline for acute and chronic pancreatitis: summary of the Working Group on Drug Development and Trials in Recurrent Acute Pancreatitis at the National Institute of Diabetes and Digestive and Kidney Diseases workshop. *Pancreas*. 2018;47:1193–1199.
8. Whitcomb DC. Primer on precision medicine for complex chronic disorders. *Clin Transl Gastroenterol*. 2019;10:e00067.
9. Whitcomb DC. Barriers and research priorities for implementing precision medicine. *Pancreas*. 2019;48:1246–1249.
10. Yadav D, Hawes RH, Brand RE, et al. Alcohol consumption, cigarette smoking, and the risk of recurrent acute and chronic pancreatitis. *Arch Intern Med*. 2009;169:1035–1045.
11. Conwell DL, Banks PA, Sandhu BS, et al. Validation of demographics, etiology, and risk factors for chronic pancreatitis in the USA: a report of the

North American Pancreas Study (NAPS) group. *Dig Dis Sci*. 2017; 62:2133–2140.

12. Yadav D, Lowenfels AB. The epidemiology of pancreatitis and pancreatic cancer. *Gastroenterology*. 2013;144:1252–1261.
13. Zator Z, Whitcomb DC. Insights into the genetic risk factors for the development of pancreatic disease. *Therap Adv Gastroenterol*. 2017; 10:323–336.
14. Sofia VM, Surace C, Terlizzi V, et al. Trans-heterozygosity for mutations enhances the risk of recurrent/chronic pancreatitis in patients with cystic fibrosis. *Mol Med*. 2018;24:38.
15. Zou WB, Tang XY, Zhou DZ, et al. SPINK1, PRSS1, CTRC, and CFTR genotypes influence disease onset and clinical outcomes in chronic pancreatitis. *Clin Transl Gastroenterol*. 2018;9:204.
16. Rosendahl J, Landt O, Bernadova J, et al. CFTR, SPINK1, CTRC and PRSS1 variants in chronic pancreatitis: is the role of mutated CFTR overestimated? *Gut*. 2013;62:582–592.
17. Whitcomb DC, Gorry MC, Preston RA, et al. Hereditary pancreatitis is caused by a mutation in the cationic trypsinogen gene. *Nat Genet*. 1996; 14:141–145.
18. Gorry MC, Gabbaizedeh D, Furey W, et al. Mutations in the cationic trypsinogen gene are associated with recurrent acute and chronic pancreatitis. *Gastroenterology*. 1997;113:1063–1068.
19. Kerem B, Rommens JM, Buchanan JA, et al. Identification of the cystic fibrosis gene: genetic analysis. *Science*. 1989;245:1073–1080.
20. Riordan JR, Rommens JM, Kerem B, et al. Identification of the cystic fibrosis gene: cloning and characterization of complementary DNA. *Science*. 1989;245:1066–1073.
21. Sharer N, Schwarz M, Malone G, et al. Mutations of the cystic fibrosis gene in patients with chronic pancreatitis. *N Engl J Med*. 1998;339:645–652.
22. Cohn JA, Friedman KJ, Noone PG, et al. Relation between mutations of the cystic fibrosis gene and idiopathic pancreatitis. *N Engl J Med*. 1998; 339:653–658.
23. Fjeld K, Weiss FU, Lasher D, et al. A recombined allele of the lipase gene CEL and its pseudogene CELP confers susceptibility to chronic pancreatitis. *Nat Genet*. 2015;47:518–522.
24. Xiao X, Jones G, Sevilla WA, et al. A carboxyl ester lipase (CEL) mutant causes chronic pancreatitis by forming intracellular aggregates that activate apoptosis. *J Biol Chem*. 2016;291:23224–23236.
25. Zenker M, Mayerle J, Lerch MM, et al. Deficiency of UBR1, a ubiquitin ligase of the N-end rule pathway, causes pancreatic dysfunction, malformations and mental retardation (Johanson-Blizzard syndrome). *Nat Genet*. 2005;37:1345–1350.
26. Witt H, Luck W, Hennies HC, et al. Mutations in the gene encoding the serine protease inhibitor, Kazal type 1 are associated with chronic pancreatitis. *Nat Genet*. 2000;25:213–216.
27. Whitcomb DC, LaRusch J, Krasinskas AM, et al. Common genetic variants in the CLDN2 and PRSS1-PRSS2 loci alter risk for alcohol-related and sporadic pancreatitis. *Nat Genet*. 2012;44:1349–1354.
28. Manolio TA, Collins FS, Cox NJ, et al. Finding the missing heritability of complex diseases. *Nature*. 2009;461:747–753.
29. Sankaran SJ, Xiao AY, Wu LM, et al. Frequency of progression from acute to chronic pancreatitis and risk factors: a meta-analysis. *Gastroenterology*. 2015;149:1490–1500.e1.
30. Petrov MS, Yadav D. Global epidemiology and holistic prevention of pancreatitis. *Nat Rev Gastroenterol Hepatol*. 2019;16:175–184.
31. LaRusch J, Lozano-Leon A, Stello K, et al. The common chymotrypsinogen C (CTRC) variant G60G (C.180T) increases risk of chronic pancreatitis but not recurrent acute pancreatitis in a North American population. *Clin Transl Gastroenterol*. 2015;6:e68.
32. Takeyama Y. Long-term prognosis of acute pancreatitis in Japan. *Clin Gastroenterol Hepatol*. 2009;7(11 Suppl):S15–S17.
33. Ahmed Ali U, Issa Y, Hagenaars JC, et al. Risk of recurrent pancreatitis and progression to chronic pancreatitis after a first episode of acute pancreatitis. *Clin Gastroenterol Hepatol*. 2016;14:738–746.
34. LaRusch J, Jung J, General IJ, et al. Mechanisms of CFTR functional variants that impair regulated bicarbonate permeation and increase risk for pancreatitis but not for cystic fibrosis. *PLoS Genet*. 2014;10:e1004376.
35. Masson E, Chen JM, Audrézet MP, et al. A conservative assessment of the major genetic causes of idiopathic chronic pancreatitis: data from a comprehensive analysis of PRSS1, SPINK1, CTRC and CFTR genes in 253 young French patients. *PLoS One*. 2013;8:e73522.
36. Hegyi E, Sahin-Tóth M. Genetic risk in chronic pancreatitis: the trypsin-dependent pathway. *Dig Dis Sci*. 2017;62:1692–1701.
37. Sahin-Tóth M. Genetic risk in chronic pancreatitis: the misfolding-dependent pathway. *Curr Opin Gastroenterol*. 2017; 33:390–395.
38. Larusch J, Whitcomb DC. Genetics of pancreatitis with a focus on the pancreatic ducts. *Minerva Gastroenterol Dietol*. 2012;58:299–308.
39. Shalimar, Midha S, Hasan A, et al. Long-term pain relief with optimized medical treatment including antioxidants and step-up interventional therapy in patients with chronic pancreatitis. *J Gastroenterol Hepatol*. 2017; 32:270–277.
40. Richards S, Aziz N, Bale S, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med*. 2015;17:405–424.
41. Li J, Ji L. Adjusting multiple testing in multilocus analyses using the eigenvalues of a correlation matrix. *Heredity (Edinb)*. 2005;95:221–227.
42. Danecek P, Auton A, Abecasis G, et al. The variant call format and VCFtools. *Bioinformatics*. 2011;27:2156–2158.
43. Chen RH, Chen YH, Huang TY. Ubiquitin-mediated regulation of autophagy. *J Biomed Sci*. 2019;26:80.
44. Butler A, Hoffman P, Smibert P, et al. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat Biotechnol*. 2018;36:411–420.
45. Stuart T, Butler A, Hoffman P, et al. Comprehensive integration of single-cell data. *Cell*. 2019;177:1888–1902.e21.
46. Enge M, Arda HE, Mignardi M, et al. Single-cell analysis of human pancreas reveals transcriptional signatures of aging and somatic mutation patterns. *Cell*. 2017;171:321–330.e14.
47. Sukalo M, Fiedler A, Guzmán C, et al. Mutations in the human UBR1 gene and the associated phenotypic spectrum. *Hum Mutat*. 2014;35:521–531.
48. Sukalo M, Schäflein E, Schanze I, et al. Expanding the mutational spectrum in Johanson-Blizzard syndrome: identification of whole exon deletions and duplications in the UBR1 gene by multiplex ligation-dependent probe amplification analysis. *Mol Genet Genomic Med*. 2017;5:774–780.
49. Masamune A, Nakano E, Niihori T, et al. Variants in the UBR1 gene are not associated with chronic pancreatitis in Japan. *Pancreatology*. 2016; 16:814–818.
50. Pfützer RH, Barmada MM, Brunskill AP, et al. SPINK1/PSTI polymorphisms act as disease modifiers in familial and idiopathic chronic pancreatitis. *Gastroenterology*. 2000;119:615–623.
51. Rosendahl J, Witt H, Szmola R, et al. Chymotrypsin C (CTRC) variants that diminish activity or secretion are associated with chronic pancreatitis. *Nat Genet*. 2008;40:78–82.
52. Masson E, Chen JM, Scotet V, et al. Association of rare chymotrypsinogen C (CTRC) gene variations in patients with idiopathic chronic pancreatitis. *Hum Genet*. 2008;123:83–91.
53. Deneke SM, Fanburg BL. Regulation of cellular glutathione. *Am J Physiol*. 1989;257:L163–L173.

54. Hanigan MH. Gamma-glutamyl transpeptidase: redox regulation and drug resistance. *Adv Cancer Res*. 2014;122:103–141.
55. Lu SC. Glutathione synthesis. *Biochim Biophys Acta*. 2013; 1830:3143–3153.
56. Ikeda Y, Taniguchi N. Gene expression of gamma-glutamyltranspeptidase. *Methods Enzymol*. 2005;401:408–425.
57. Nakashima Y, Miyagi-Shiohira C, Kobayashi N, et al. A proteome analysis of pig pancreatic islets and exocrine tissue by liquid chromatography with tandem mass spectrometry. *Islets*. 2017;9:159–176.
58. Ballatori N, Krance SM, Notenboom S, et al. Glutathione dysregulation and the etiology and progression of human diseases. *Biol Chem*. 2009; 390:191–214.
59. Brand H, Diergaarde B, O'Connell MR, et al. Variation in the γ-glutamyltransferase 1 gene and risk of chronic pancreatitis. *Pancreas*. 2013;42:836–840.
60. Diergaarde B, Brand R, Lamb J, et al. Pooling-based genome-wide association study implicates gamma-glutamyltransferase 1 (GGT1) gene in pancreatic carcinogenesis. *Pancreatology*. 2010;10:194–200.
61. Ściskalska M, Ołdakowska M, Marek G, et al. Increased risk of acute pancreatitis occurrence in smokers with rs5751901 polymorphisms in GGT1 gene. *Int J Med Sci*. 2020;17:242–254.
62. Ward LD, Kellis M. HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. *Nucleic Acids Res*. 2012;40:D930–D934.
63. Melzer D, Perry JR, Hernandez D, et al. A genome-wide association study identifies protein quantitative trait loci (pQTLs). *PLoS Genet*. 2008; 4:e1000072.
64. GTEx Consortium. Human genomics. The genotype-tissue expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science*. 2015;348:648–660.
65. Corre S, Galibert MD. Upstream stimulating factors: highly versatile stress-responsive transcription factors. *Pigment Cell Res*. 2005; 18:337–348.
66. Kim Y, Jun I, Shin DH, et al. Regulation of CFTR bicarbonate channel activity by WNK1: implications for pancreatitis and CFTR-related disorders. *Cell Mol Gastroenterol Hepatol*. 2020;9:79–103.
67. Sugiaman-Trapman D, Vitezic M, Jouhilahti EM, et al. Characterization of the human RFX transcription factor family by regulatory and target gene analysis. *BMC Genomics*. 2018;19:181.
68. Gukovskaya AS, Gukovsky I, Algul H, et al. Autophagy, inflammation, and immune dysfunction in the pathogenesis of pancreatitis. *Gastroenterology*. 2017;153:1212–1226.
69. Biczo G, Vegh ET, Shalbueva N, et al. Mitochondrial dysfunction, through impaired autophagy, leads to endoplasmic reticulum stress, deregulated lipid metabolism, and pancreatitis in animal models. *Gastroenterology*. 2018;154:689–703.
70. Guda NM, Muddana V, Whitcomb DC, et al. Recurrent acute pancreatitis: international state-of-the-science conference with recommendations. *Pancreas*. 2018;47:653–666.
71. Etemad B, Whitcomb DC. Chronic pancreatitis: diagnosis, classification, and new genetic developments. *Gastroenterology*. 2001;120:682–707.
72. Pham A, Forsmark C. Chronic pancreatitis: review and update of etiology, risk factors, and management. *F1000Res*. 2018;7:F1000.
73. Whitcomb DC, North American Pancreatitis Study Group. Pancreatitis: TIGAR-O version 2 risk/etiology checklist with topic reviews, updates, and use primers. *Clin Transl Gastroenterol*. 2019;10:e00027.
74. Boulling A, Masson E, Zou WB, et al. Identification of a functional enhancer variant within the chronic pancreatitis-associated SPINK1 c.101A>G (p.Asn34Ser)-containing haplotype. *Hum Mutat*. 2017; 38:1014–1024.