Universiteit Leiden
The Netherlands

Whole-exome sequencing identifies SLC52A1 and ZNF106 variants as novel genetic risk factors for (early) multiple-organ failure in acute pancreatitis

Berg, F.F. van den; Issa, Y.; Vreijling, J.P.; Lerch, M.M.; Weiss, F.U.; Besselink, M.G.; ... ; Santvoort, H.C. van

Citation
Berg, F. F. van den, Issa, Y., Vreijling, J. P., Lerch, M. M., Weiss, F. U., Besselink, M. G., ... Santvoort, H. C. van. (2021). Whole-exome sequencing identifies SLC52A1 and ZNF106 variants as novel genetic risk factors for (early) multiple-organ failure in acute pancreatitis. *Annals Of Surgery, 275*(6), e781-e788. doi:10.1097/SLA.0000000000004312

Version:	Publisher's Version
License:	Creative Commons CC BY-NC-ND 4.0 license
Downloaded from:	https://hdl.handle.net/1887/3731065

Note: To cite this publication please use the final published version (if applicable).

OPEN

Whole-exome Sequencing Identifies *SLC52A1* and *ZNF106* Variants as Novel Genetic Risk Factors for (Early) Multiple-organ Failure in Acute Pancreatitis

Fons F. van den Berg, MD,†✉ Yama Issa, MD, PhD,*† Jeroen P. Vreijling,‡ Markus M. Lerch, MD,§ Frank Ulrich Weiss, PhD,§ Marc G. Besselink, MD, PhD,*† Frank Baas, MD, PhD,‡ Marja A. Boermeester, MD, PhD,*† and Hjalmar C. van Santvoort, MD, PhD , for the Dutch Pancreatitis Study Group*

Objective: The aim of this study was to identify genetic variants associated with early multiple organ failure (MOF) in acute pancreatitis.

Summary Background Data: MOF is a life-threatening complication of acute pancreatitis, and risk factors are largely unknown, especially in early persistent MOF. Genetic risk factors are thought to enhance severity in complex diseases such as acute pancreatitis.

Methods: A 2-phase study design was conducted. First, we exome sequenced 9 acute pancreatitis patients with early persistent MOF and 9 case-matched patients with mild edematous pancreatitis (phenotypic extremes) from our initial Dutch cohort of 387 patients. Secondly, 48 candidate variants that were overrepresented in MOF patients and 10 additional variants known from literature were genotyped in a replication cohort of 286 Dutch and German patients.

Results: Exome sequencing resulted in 161,696 genetic variants, of which the 38,333 non-synonymous variants were selected for downstream analyses. Of these, 153 variants were overrepresented in patients with multiple-organ failure, as compared with patients with mild acute pancreatitis. In total, 58 candidate variants were genotyped in the joined Dutch and German replication cohort. We found the rs12440118 variant of *ZNF106* to be overrepresented in patients with MOF (minor allele frequency 20.4% vs 11.6%, $P_{adj=}$0.026). Additionally, *SLC52A1* rs346821 was found to be overrepresented (minor allele frequency 48.0% vs 42.4%, $P_{adj=}$ 0.003) in early MOF. None of the variants known from literature were associated.

Conclusions: This study indicates that *SLC52A1,* a riboflavin plasma membrane transporter, and *ZNF106,* a zinc finger protein, may be involved in disease progression toward (early) MOF in acute pancreatitis.

Keywords: acute pancreatitis, genetics, multiple organ failure

(*Ann Surg* 2022;275:e781–e788)

Persistent (multiple) organ failure is a hallmark of disease progression toward fatal acute pancreatitis and is therefore the main criteria for the classification of severe acute pancreatitis.[1,2] Failure to identify risk factors for multiple organ failure (MOF) in acute pancreatitis has raised the suggestion for genetic variation as a causative factor.[3] Genetic association studies indicate that susceptibility for sepsis is associated with variants in pattern recognition receptors and cytokines.[4] Previous studies in patients with acute pancreatitis showed associations between functional genetic variants and multiple-organ dysfunction syndrome *(LP-PLA2, TNF)*,[3,5] systemic inflammatory response syndrome *(TNFAIP3)*,[6] septic shock (IL-10)[7] and systemic complications *(TNF)*.[8] However, there is a distinction to be made between early MOF that occurs in the first week as a result of systemic inflammation (ie, before local or other infectious complication) and MOF that occurs after the first week as a result of secondary infectious complications, mostly infected pancreatic and/or extrapancreatic necrosis. Clinical and genetic risk factors for the immunological cascade that leads to MOF in the early phase of acute pancreatitis are unknown. The aim of this study was to identify associations of non-synonymous variants with (early) MOF in an international population.

METHODS

Ethics

The protocol was approved by the local ethics committees of the University Medical Center Utrecht (PROPATRIA: 03/169, PANTER: 04/289), Erasmus Medical Center (APEC: 2012–357), and the University of Greifswald (III UV91/03), and the study was conducted in accordance with the declaration of Helsinki. All patients gave consent for the use of their blood for future genetic studies.

Definitions

Acute pancreatitis was defined as abdominal pain in combination with serum amylase or lipase concentrations that were raised to at least three times the institutional upper limit of normal. Mild acute pancreatitis was defined by the absence of local or systemic complications, according to the revised Atlanta criteria.[1] MOF was defined as failure of at least 2 of the following organ systems on the

From the *Department of Surgery, Amsterdam Gastroenterology Endocrinology Metabolism, Amsterdam UMC, University of Amsterdam, Amsterdam, The Netherlands; †Department of Clinical Genetics, Leiden University Medical Center, Leiden, The Netherlands; ‡Department of Medicine A, University Medicine Greifswald, Greifswald, Germany; §Departments of Clinical Chemistry, Genetics and Pediatrics, Amsterdam Gastroenterology & Metabolism, Amsterdam UMC, University of Amsterdam, Amsterdam, The Netherlands; and Department of Surgery, University Medical Center, Utrecht, The Netherlands; Department of Surgery, St. Antonius Hospital, Nieuwegein, The Netherlands.

✉ f.f.vandenberg@amsterdamumc.nl.

Funding: AMC PhD Scholarship (FB) and Next Generation Sequencing grants from the AMC Executive Board.

Author contributions: F.F.B., Y.I., F.B., M.A.B., H.S. contributed in conception and study design; F.F.B., Y.I., M.M.L., F.U.W., M.G.B., H.S. contributed in acquisition of data; F.F.B., Y.I., J.P.V. contributed in data analysis; F.F.B., Y.I. contributed in drafting manuscript; and F.F.B., Y.I., J.P.V., M.M.L., F.U.W., M.G.B., F.B., M.A.B., and H.S. critically revising and final approval manuscript.

The authors report no conflicts of interest.

Supplemental digital content is 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.annalsofsurgery.com).

This is an open access article distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives License 4.0 (CCBY-NC-ND), where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.

Copyright © 2021 The Author(s). Published by Wolters Kluwer Health, Inc.

ISSN: 0003-4932/21/27506-e781

DOI: 10.1097/SLA.0000000000004312

same day: pulmonary (PaO_2<60 mm Hg despite FIO_2 of 30% or the need for mechanical ventilation), renal (serum creatinine >177mmol/L after rehydration or need for hemofiltration, or hemodialysis) and cardiocirculatory (systolic blood pressure <90 mm Hg despite adequate fluid resuscitation or need for vasopressor support. Persistent organ failure was defined as the presence of organ failure on at least 3 consecutive days (lasting more than 48 h).[1] Early MOF was defined by the presence of MOF in the first 2 weeks of admission.

Two-phase Study Design

Our study consists of 2 phases (Fig. 1). First, the exomes of the discovery cohort, formed by patients with acute pancreatitis from the opposite ends of the clinical spectrum (ie, persistent MOF vs mild disease), were sequenced and after bioinformatics analysis a candidate variant list was constructed. Next, the joined cohort, consisting of the discovery cohort and replication cohort, was genotyped for the candidate variants.

Patient Populations

Discovery Cohort

For the discovery phase, we selected patients from 2 phenotypic extreme subgroups: patients with persistent multi-organ failure in the first week of admission on one end of the clinical spectrum, and patients with mild pancreatitis who were discharge within the first

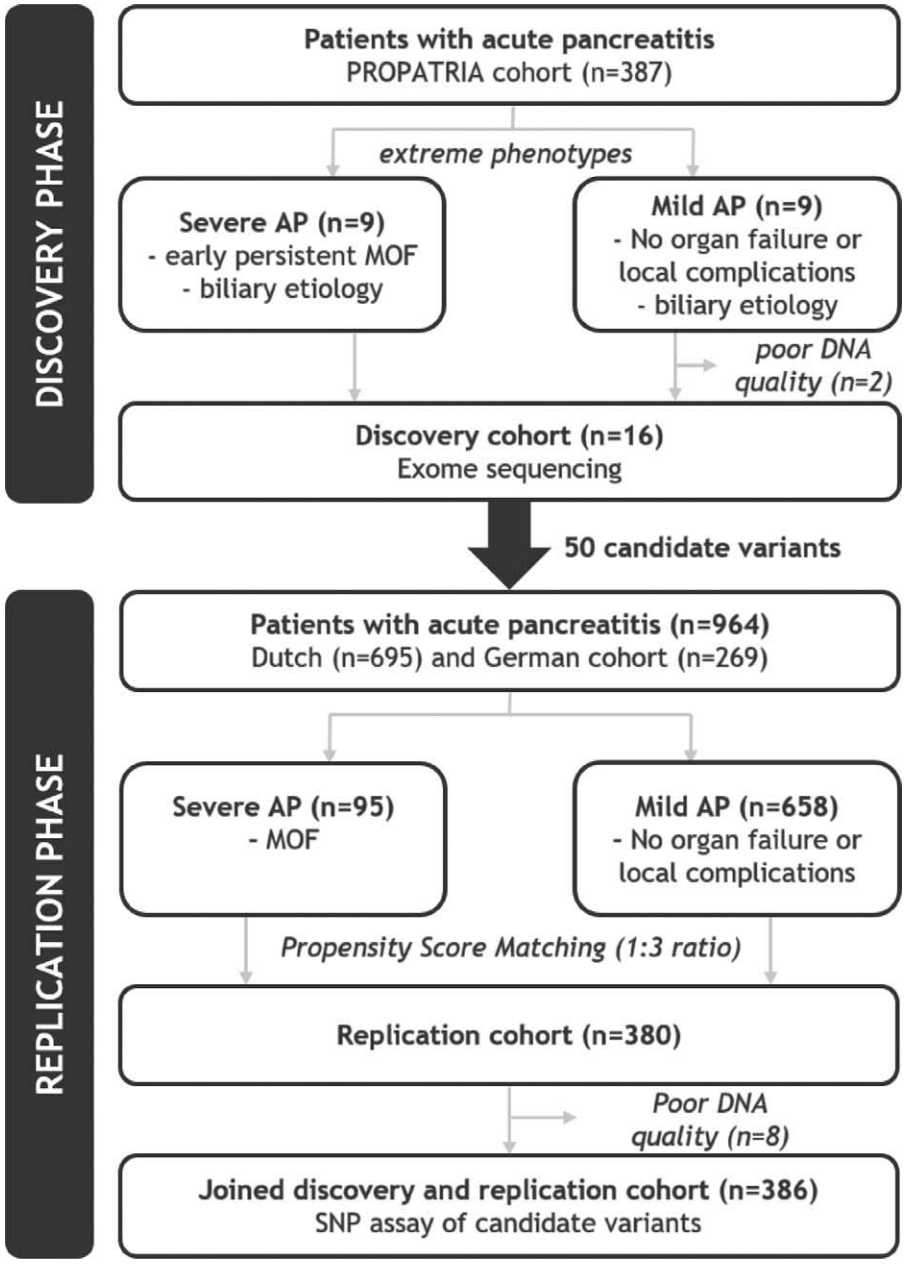

FIGURE 1. Two-phase study design.

© 2021 The Author(s). Published by Wolters Kluwer Health, Inc.

week on the other end of the clinical spectrum. All patients had a biliary etiology (as previously described[9]) and were matched for sex, age, and comorbidity (ie, ASA classification). Patients were selected from a prospectively included cohort of 387 patients who were diagnosed with acute pancreatitis in 15 Dutch hospitals between March 2004 and March 2007.[10]

Replication Cohort

For the replication phase, patients were selected from two cohorts: a Dutch cohort of 695 patients who were screened for eligibility for three studies performed by the Dutch Pancreatitis Study Group: PROPATRIA (2004–2007),[10] PANTER (2005–2008),[11] and APEC (2013–2018, *Lancet, in press),* and a German cohort (2006–2014) of 309 patients with acute pancreatitis. From the Dutch cohort, 89 patients were excluded because no blood sample or clinical data were available and 17 were excluded because patients' MOF was possibly associated with the probiotics (ie, the study intervention in PROPATRIA). After exclusion of these patients, all consecutive patients from the combined Dutch and German cohorts who had MOF were included: 80 patients from the Dutch Cohort, and 15 from the German cohort. Each of these 95 patients were then matched by age, sex, and etiology with 3 patients who had mild acute pancreatitis (ie, no pancreatic/extrapancreatic necrosis or organ failure) of their respectable cohort. Propensity score matching was applied using nearest neighbor matching with R package ''Matchlt''. In total, 380 patient samples were selected for genotyping.

DNA Isolation

For the samples of the discovery phase, DNA was isolated from white blood cells using Qiagen Autopure kit with glycogen according to the manufactures protocol. DNA quantity was measured using Qubit Fluorometric Quantitation before sequencing. DNA was isolated from the blood samples of the German and Dutch validation cohort using respectively Quick gDNA Blood kit (Zymo Research) and DSP DNA Midi Kit (Qiagen; ref: 937255) according to the manual.

Exome Sequencing

Exome capture was performed using the Ion Ampliseq Exome kit according to the manufacturer's manual. The libraries were prepared on the PI chip with the Ion PI Template OT2 200 kit v3 and Ion PI sequencing 200 kit v3 following their respective manuals.

The samples were sequenced on an Ion Proton sequencer.

Alignment, Quality Control, Variant Calling

Base calling, quality control, and alignment were performed using the Torrent Server Suite (V4.0.2) with default settings. The output of the software was as Binary Alignment Map file, a binary representation of reads aligned against the Hg19 reference genome. Variant Calling Format files were also automatically created. The variants were compressed and annotated using KGGSeq[12] and an in-house script.

Bioinformatics Analyses

For the downstream gene discovery analyses we selected only those variants that affected the amino acid sequence (non-synonymous). We used the following bioinformatics strategies to filter out possible biological relevant variants. All calculations were done with R statistical software.[13] Nominal significance was defined as a P value <0.05.

Allele Distribution

Variants that differ in allele frequency between the 2 phenotypes were identified using the ''Fisher Exact'' function of the ''stats'' package of R.

Deviation From Hardy-Weinberg Equilibrium

For the discovery phase, we selected patients with extreme phenotypes, that is, if we are assuming a variant has a gene-dose depended effect on the studied phenotype, we would expect to identify only homozygous wildtype or variant genotypes in our patients. We calculated deviation from Hardy-Weinberg equilibrium (HWE) with the HWExact function of the ''genetics'' package of R.

Variants that deviated from HWE and occurred significantly more often in the patients with the severe phenotype were selected.

Genes From Literature

We selected genes that were associated in the literature with disease susceptibility or progression and cross-matched with our list of significant variants based on allele distribution.[14] Additionally, we selected genetic variants (both intronic and exonic) that are positively associated with severe acute pancreatitis or systemic complications, but that have not been replicated. These were rs1799964 and rs361525 in TNF,[3,8] rs3775291 in *TLR3*,[15] rs7543795 in *TLR6*,[15] rs5744455 in *CD14*,[16] rs696 in *NFKBIA*,[17] rs1805017 and rs76863441 in *PLA2G7*,[5] rs1801282 in *PPARG*,[18] and rs5029924 in *TNFAIP3*.[6]

Candidate Gene Selection

Since the Ion-Proton machine has a high error rate in homopolymer stretches (ie, false-positive variant calls) we inspected each variant manually with the Integrative Genomics Viewer (V2.3, Broad Institute).[19] The variants that were obviously based on an artefact were removed from the analyses. Reasons for exclusion were low coverage, a stretch of $\geq$4 repetitive nucleotides, missing nucleotide calls, rare or novel variant present in all samples, and unequal distribution of nucleotides. The most significant variants from each bioinformatics strategy were selected, thereby creating a candidate variant list of genes that are potentially associated with the severe disease phenotype.

Single-Nucleotide Polymorphism Genotyping

The single-nucleotide polymorphism (SNP) genotyping assay is done by loading genomic DNA samples on custom Taqman OpenArray plates (ThermoFisher Scientific) with a QuantStudio 12k Flex AccuFill System. Eleven of 51 candidate SNPs were technically incompatible with the OpenArray system and therefore replaced by other candidate variants based on statistical significance of allele distribution. Genotypes were determined by running the plates on a QuantStudio 12k Flex Real Time PCR System, following the manufacturer's instructions. Genotype distribution of groups was compared using the co-dominant genetic model. Odds ratio and the corresponding 95% confidence interval were calculated with a likelihood ratio test and the most frequent genotype as reference, using the ''SNPassoc'' R package.

RESULTS

Exome Sequencing for Candidate Variant Selection

Patient characteristics of the case-matched discovery cohort are summarized in Table 1. There were no differences in baseline characteristics such as age, sex, and ASA physical status classification. Quality control measures and summary sequence statistics of the exome sequencing step are presented in Supplementary Table 1, http://links.lww.com/SLA/C402. Two samples of the mild acute pancreatitis group were excluded before sequencing due to insufficient DNA quality. On average 56 million reads per sample (range 32–83) were created with an average read length of 142 (range 134–150). Ninety-three percent of the reads were on targetand the average

© 2021 The Author(s). Published by Wolters Kluwer Health, Inc.

TABLE 1. Clinical Characteristics of Discovery Patient Cohort

		MAP	SAP	*P*
N		7	9	
Sex	Male n (%)	3 (42.9)	4 (44.4)	1
Age, y	Mean (SD)	64.1 (13.57)	66.4 (12.1)	0.726
ASA		2 (0.58)	2.2 (0.67)	0.495
Severity of pancreatitis (mean, SD)				
	APACHE-II score*	6.43 (4.58)	7.78 (4.24)	0.551
	Imrie score	3 (1.83)	4.50 (2.12)	0.414
	CRP, highest value in first 48 h, mg/L	134 (58.3)	355 (160)	0.055
	Severe acute pancreatitis†	0	9	<0.001
	Necrotizing pancreatitis‡	0	7	0.009
Complications				
	Infections	0	8	0.002
	Infected pancreatic necrosis	0	5	0.067
	Positive blood culture	0	5	0.067
	Organ failure during admission	0	9	<0.001
	Multiple-organ failure during admission	0	9	<0.001
	Mortality	0	5	0.067

*Highest score on day of admission.
†Organ failure and/or necrosis.
‡Defined as: pancreatic parenchymal necrosis demonstrated on contrast-enhanced computed tomography scan.
CRP indicates C-reactive protein; MAP, mild acute pancreatitis; SAP, severe acute pancreatitis.

on target depth of sequencing was 129. Figure 2 shows the downstream bioinformatics analyses. After variant calling, 161,696 variants passed the GATK quality control. Filtering for non-synonymous variants resulted in 38,333 variants (24%) for downstream analyses.

When comparing allele frequencies of both patient groups, 135 protein-affecting variants that occurred significantly more in the severe acute pancreatitis group were identified. Twenty-one of these variants did not pass our quality control and were excluded. We ranked the variants based on *P* value and selected the top variants for the candidate list (see Supplementary Table 2A, http://links.lww.com/SLA/C402). Of the 135 identified variants, we found 17 variants that were out of Hardy-Weinberg equilibrium in one of the groups (see Supplementary Table 2B, http://links.lww.com/SLA/C402). All variants were included in the candidate list. When crossreferencing all genes associated with susceptibility and severity for pancreatitis that were identified in a recent systematic review[15] with the 135 significant variants we identified only 2 genes, *CASP8* and *CASP10* (respectively code for caspase-8 and caspase-10). These genes are key components in the extrinsic apoptosis pathway (see Supplementary Table 2C, http://links.lww.com/SLA/C402).

We compiled a candidate variant list for validation by combining these analyses which resulted in 48 non-synonymous variants.

Additionally, 10 SNPs (in TNF, *TLR3, TLR6, CD14, NFKBIA, PLA2G7, PPARG,* and *TNFAIP3* genes) that were positively associated with severity, MOF, or systemic complications in unreplicated studies were genotyped. So, in total 58 variants were genotyped. The variants are presented in Supplementary Table 3, htttp://link-s.lww.com/SLA/C402 with their corresponding reported allele frequency from the ExAC-NFE database.

Replication in an International Cohort

In total, 80 Dutch and 15 German patients with acute pancreatitis with MOF were matched in a 1:3 ratio for sex, age, and etiology with mild acute pancreatitis patients (Fig. 1). Eight patients of the replication cohort were excluded due to low DNA quality. There was no difference in matching variables in the Dutch, German, and joined discovery and replication cohort (See Table 2 for patient characteristics). The average age of the joined discovery and replication cohort was 58.6 [standard deviation (SD) 16.1] for patients with MOF and 59.5 (SD 12.8) for mild acute pancreatitis patients. Biliary and alcoholic were the most prevalent etiologies, accounting for approximately 70% of the patients. From the patients with MOF, this developed in the first week of admission in 56 patients (59%). A total of 88 patients (93%) developed persistent MOF. Seventy-one atients suffered from necrotizing pancreatitis, of whom almost all (94%, n = 67) developed secondary infection of the necrosis.

Twenty-nine patients (30%) in the MOF group died. There was no mortality in the patients with mild acute pancreatitis.

Four of the 58 SNPs were nominally significantly associated with MOF in the joined cohort, with only the variant rs12440118 (*ZNF106)* remaining significant after Bonferroni correction for multiple testing (Table 3) (minor allele frequency 20.4% vs 11.6%, $P<0.001$, $P_{adj} = 0.026$). The other loci correspond to rs743580 at *PML* (52.6% vs 49.5%, $P = 0.041$), rs2272522 at *CHL1* (27.6% vs 19.4%, $P = 0.012$), and rs346821 at *SLC52A1* (45.5% vs 42.4%, $P = 0.003$). The subgroup analysis for early MOF shows that 3 of 50 SNPs reached nominal significance (Table 4).

Only the *SLC52A1* variant remained significant following the Bonferroni correction (48.0% vs 42.4%, $P = <0.001$, $P_{adj} = 0.003$).

Furthermore, the aforementioned *CHL1* variant and rs12440118 in *ZNF106* were nominal significantly associated with early MOF.

Stratification for mortality identified positive associations with the *CHL1, ZNF106,* and *SLC52A1* variants, and 1 additional variant (rs7209474) in *CCDC57*. These associations lost statistical significance after Bonferroni correction. None of the variants that are known from the literature to be associated with severe complications was associated with (early) MOF, infected pancreatic necrosis, or mortality in our analysis (Table 5).

CONCLUSIONS

In this study we identified 2 genetic variants that are overrepresented in patients with pancreatitis and MOF. We found a genetic variant, rs346821, on chr 17:4936972 (HG19), in the solute carrier family 52 member 1 *(SLC52A1)* gene, that was significantly associated with early MOF. According to *in silico* prediction models, this variant represents a benign mutation. This gene codes for a member of the riboflavin (vitamin B2) transporter family, among other expressed in the duodenum and small intestine.[20] Functional

© 2021 The Author(s). Published by Wolters Kluwer Health, Inc.

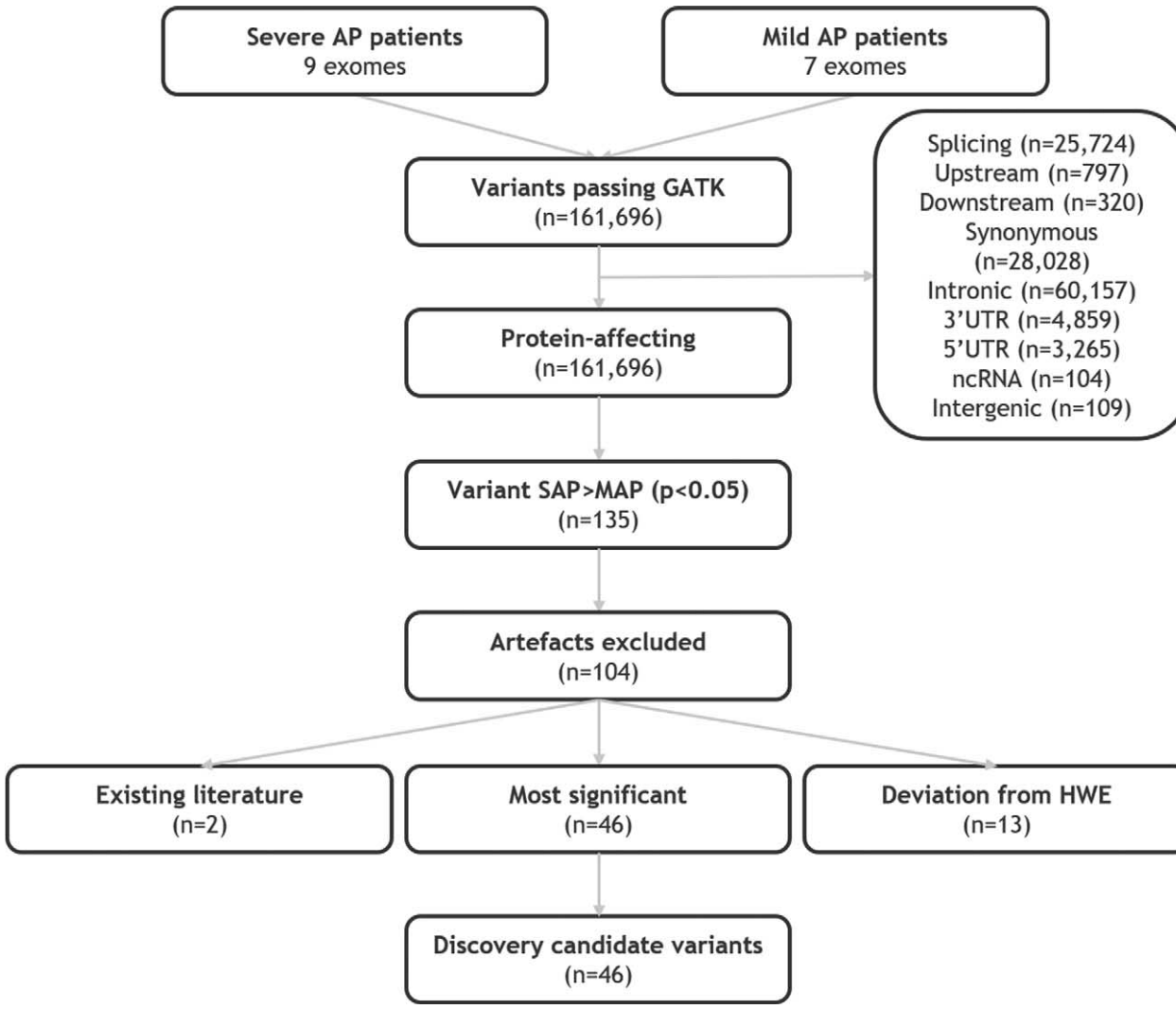

FIGURE 2. Downstream analysis discovery phase.

genetic variants in the riboflavin transport family causing riboflavin deficiency have been described in a few cases.[21,22] Recently, it has been shown that intermediates of riboflavin synthesis, produced by the gut microbiota, interact with mucosal-associated invariant T cells (MAIT), thereby potentially affecting the immune response.[23,24] Also, a high dose of riboflavin protects mice from *E coli*-induced sepsis and septic shock.[25] However, it remains unknown whether this *SLC52A1* variant, or (intronic) variants that are in linkage equilibrium with rs346821, plays a causal role in the immune response leading to early MOF in acute pancreatitis.

Secondly, we found a strong association with MOF of rs12440118 on chr15:42744094 (HG19), a missense variant in the gene that codes for zinc finger protein 106 *(ZNF106). ZNF106* is an RNA-binding protein mostly expressed in skeletal muscle tissue and has been implicated in neurodegenerative disorders.[26] Furthermore, this particular variant has been associated with glycated-hemoglobin levels in humans.[27] Although we found a strong association with MOF, it remains unknown how this relates to the clinical phenotype.

Next to the genome wide discovery of novel associations, we also tested whether variants in genes involved in the innate immune system could be associated with altered immune response leading to MOF. Meta-analyses of genetic association studies revealed variants in *CD14, TLR2,* and *TNF* reported to be associated with sepsis.[28] A common hypothesis states that MOF occurs early in the pathogenesis of acute pancreatitis and is not directly related to infectious complications. A previously published meta-analysis did not show credible associations between SNPs in genes related to the innate immunity with disease susceptibility or severity in acute pancreatitis.[14] This was confirmed by the present study, as we did not find associations with any of the genetic variants in these genes.

We hypothesized that genetic variation in genes that regulate apoptosis, such as caspases, promotes necrosis of acinar cells and progression to MOF. Animal models of acute pancreatitis indicate an important role for apoptosis in the pathogenesis of disease severity.[29,30] Administration of the apoptosis inhibitor XIAB during experimental pancreatitis attenuated pancreatic necrosis and increased mortality.[31] A newly discovered form of programmed cell death called necroptosis has also been implicated in acute pancreatitis.[32] However, variants in *CASP8* and *CASP10* genes which we identified in the discovery phase of our study were not associated with MOF in the replication cohort.

© 2021 The Author(s). Published by Wolters Kluwer Health, Inc.

TABLE 2. Clinical Characteristics of Patients Included in Replication Cohorts

		Replication Cohort (Dutch)			Replicatio n Cohort (Ge rman)			Discovery and Replication Cohorts		
		MAP	SAP	*P*	MAP	SAP	*P*	MAP	SAP	*P*
n		235	74		45	15		287	98	
Sex	Male, n (%)	161 (68.5)	48 (64.9)	0.658	39 (86.7)	13 (86.7)	1	203 (70.7)	65 (66.3)	0.489
Age	Mean (SD)	58.35 (16.5)	58.92 (13.1)	0.788	59.11 (14.3)	58.13 (11.3)	0.810	58.61 (16.1)	59.49 (12.8)	0.625
Etiology				0.409			0.957			0.497
	Biliary, n (%)	110 (46.8)	31 (41.9)		19 (42.2)	6 (40.0)		136 (47.4)	46 (46.9)	
	Alcoholic, n (%)	50 (21.3)	19 (25.7)		16 (35.6)	6 (40.0)		66 (23.0)	25 (25.5)	
	Idiopathic, n (%)	55 (23.4)	17 (23.0)		8 (17.8)	2 (13.3)		63 (22.0)	19 (19.4)	
	Other, n (%)	20 (8.5)	6 (8.1)		2 (4.4)	1 (6.7)		232 (7.7)	7 (7.1)'	
Complications										
	Pancreatic necrosis*	0	49 (66.2)	<0.001	0	15 (100.0)	<0.001	0	71 (72.4)	<0.001
	Infected pancreatic necrosis	0	49 (66.2)	<0.001	0	13 (92.9)	<0.001	0	67 (69.1)	<0.001
	Persistent organ failure	0	67 (91.8)	<0.001	0	13 (92.9)	<0.001	0	89 (92.7)	<0.001
	Early multiple organ failure	0	38 (51.4)	<0.001	0	10 (76.9)	<0.001	0	57 (59.4)	<0.001
	Mortality	0	20 (27.0)	<0.001	0	4 (26.7)	0.003	0	29 (29.6)	<0.001

*Defined as: pancreatic parenchymal necrosis demonstrated on contrast-enhanced computed tomography scan.
CRP indicates C-reactive protein; MAP, mild acute pancreatitis; SAP, severe acute pancreatitis; P, *P* value (Fisher exact test).

In silico prediction models of the functional consequence of non-synonymous mutations have been developed for Mendelian diseases, based on biochemical changes of the protein or degree of conservation. Popular prediction tools such as PolyPhen-2[33] and SIFT[34] have moderate sensitivity but poor specificity and seem to predict loss-of-function mutations better than gain-of-function mutations.[35] The performance is increased through a combination of tools, which is a feature in KGGseq.[36] According to these prediction tools, none of the identified variants were disease-causing (data not shown). However, these tools are primarily designed for diseases with a Mendelian inheritance pattern. Complex diseases that are suspected to be caused or modified by genetics are the result of a combination of variants that interact with environmental factors. Traditional prediction models are therefore not reliable for the functional prediction of variants in complex diseases and gene-set or network enrichment analyses are more suitable. Unbiased ''big

TABLE 3. Significant Associations for Joined Analysis of Patients With Multiple Organ Failure Versus Mild Disease

				AA		Aa		aa				
Gene	SNP (Variant)	MAF	Group	No. (%)	OR (95% CI)	No. (%)	OR (95% CI)	No. (%)	OR (95% CI)	P	P_{adj}	Missing Data (%)
PML	rs743580	0.503	SAP	26 (27.1)	1	39 (40.6)	0.5 (0.3–0.9)	31 (32.3)	0.8 (0.4–1.5)	0.041	1	2.3
	A/G		MAP	65 (23.0)		155 (55.0)		62 (22.0)				
CHL1	rs2272522	0.785	SAP	55 (55.7)	1	32 (33.0)	1.2 (0.7–2.0)	11 (11.3)	4.1 (1.6–10.5)	0.012	0.569	1.6
	C/T		MAP	183 (64.4)		92 (32.4)		9 (3.2)				
ZNF106	rs12440118	0.864	SAP	59 (63.4)	1	30 (32.3)	1.7 (1.0–2.9)	4 (4.3)	–	<0.001	0.026	4.4
	A/G		MAP	214 (77.3)		63 (22.7)		0 (0)				
SLC52A1	rs346821	0.568	SAP	15 (17.0)	1	66 (75.0)	0.9 (0.4–1.6)	7 (8.0)	9.6 (1.8–51.3)	0.003	1	2.1
	G/A		MAP	41 (16.0)		213 (83.2)		2 (0.8)				

AA, indicates homozygous dominant genotype; Aa, heterogenous; aa homozygous recessive; MAF, minor allele frequency; OR, odds ratio; CI, confidence interval; *P, P* value (Fisher exact test), *P*adj, *P* value (Bonferroni corrected); PML, Promyelocytic Leukemia; *CHL1,* Cell Adhesion Molecule L1 like; *ZNF106,* Zinc Finger Protein 106; *SLC52A1,* Solute Carrier Family 52 Member 1; SAP, severe acute pancreatitis; MAP, mild acute pancreatitis.

TABLE 4. Significant Associations for Joined Subgroup Analysis for Patients With Early Multiple Organ Failure

				AA		Aa		aa				
Gene	SNP (variant)	MAF	Group	No. (%)	OR (95% CI)	No. (%)	OR (95% CI)	No. (%)	OR (95% CI)	P	P_{adj}	Missing Data (%)
CHL1	rs2272522	0.785	eMOF	30 (52.6)	1	20 (35.1)	1.3 (0.7–2.4)	7 (12.3)	4.0 (1.5–10.9)	0.038	1	1.6
	C/T		n-eMOF	205 (63.9)		104 (32.4)		12 (3.7)				
ZNF106	rs12440118	0.864	eMOF	34 (61.8)	1	18 (32.7)	1.7 (0.9–3.1)	3 (5.5)	20.8 (2.1–206.0)	0.008	0.346	4.4
	A/G		n-eMOF	236 (75.6)		75 (24.0)		1 (0.3)				
SLC52A1	rs346821	0.568	eMOF	9 (17.6)	1	35 (68.6)	0.7 (0.3–1.6)	7 (13.7)	20.1 (3.5–114.8)	<0.001	0.003	11.1
	G/A		n-eMOF	46 (15.9)		242 (83.4)		2 (0.7)				

AA indicates homozygous dominant genotype; Aa, heterogenous; aa homozygous recessive; MAF, minor allele frequency; OR, odds ratio; CI, confidence interval; *P, P* value (Fisher exact test), P_{adj}, *P* value (Bonferroni corrected); *CHL1,* Cell Adhesion Molecule L1 like; *ZNF106,* Zinc Finger Protein 106; *SLC52A1,* Solute Carrier Family 52 Member 1; eMOF, early multiple organ failure; n-eMOF, non-early multiple organ failure.

© 2021 The Author(s). Published by Wolters Kluwer Health, Inc.

TABLE 5. Significant Associations for Joined Analysis Stratified for Mortality

				AA		Aa		aa				
Gene	SNP (variant)	MAF	Group	No. (%)	OR (95% CI)	No. (%)	OR (95% CI)	No. (%)	OR (95% CI)	*P*	P_{adj}	Missing Data (%)
CHL1	rs2272522	0.785	n-surv	15 (53.6)	1	8 (28.6)	1.0 (0.4–2.5)	5 (17.9)	4.9 (1.6–15.3)	0.041	1	1.6
	C/T		surv	221 (62.8)		116 (33.0)		15 (4.3)				
ZNF106	rs12440118	0.864	n-surv	15 (55.6)	1	10 (37.0)	2.1 (0.9–4.8)	2 (7.4)	17.1 (2.3–130.2)	0.018	0.816	4.4
	A/G		surv	257 (75.1)		83 (24.3)		2 (0.6)				
CCDC57	rs7209474	0.546	n-surv	5 (17.2)	1	12 (41.4)	1.7 (0.6–4.8)	12 (41.4)	3.8 (1.3–11.2)	0.033	1	1.6
	C/T		surv	114 (32.5)		165 (47.0)		72 (20.5)				
SLC52A1	rs346821	0.568	n-surv	3 (12.0)	1	18 (72.0)	1.2 (0.4–4.3)	4 (16.0)	14.1 (2.4–81.8)	0.007	0.316	11.1
	G/A		surv	53 (16.7)		260 (81.8)		5 (1.6)				

AA, indicates homozygous dominant genotype; Aa, heterogenous; aa homozygous recessive; MAF, minor allele frequency; OR, odds ratio; CI, confidence interval; P, *P* value (Fisher exact test), P_{adj}, *P* value (Bonferroni corrected); *CHL1*, Cell Adhesion Molecule L1 like; *ZNF106,* Zinc Finger Protein 106; *CCDC57,* Coiled-Coil Domain 57; *SLC52A1,* Solute Carrier Family 52 Member 1; n-surv, nonsurvivor; surv, survivor.

data'' techniques such as machine learning are using data from genome-wide association studies and have recently been proposed as a genetic prediction model for complex traits.[37] However, all currently used models are moderately reliable and should be used with caution.[38]

A limitation of our study is that only the exons, which represent the part of the genome that codes for protein, were sequenced. Although we used to refer to the noncoding part (introns) as ''junk DNA'', recent insights suggest that approximately 80% of the genomic code has a biochemical function.[39] It is increasingly known that variation in the promotor or enhancer regions of a gene can alter gene expression and can be associated with a disease phenotype. It is therefore possible that the difference in phenotype between the patient groups is based on noncoding variation and cannot be discovered with exome sequencing. Furthermore, we only selected the nonsynonymous variants for further analyses. Although it does not alter the amino acid sequence, synonymous variants are identified that alter the protein function, for example, through altered mRNA stability resulting in changed protein expression levels. The reason we chose this strategy is that we are still largely unable to predict the functional consequence of these variants. Second, although our cohort consists of a rare population of severely ill patients, in terms of genetic research, the sample size of our discovery cohort was relatively small. By applying the extreme phenotype approach, we were able to limit our sample size and still be able to detect true differences between the phenotypes. With the still decreasing costs of high throughput sequencing, it becomes more feasible to sequence whole exomes, or even whole genomes, of the entire cohort, reducing the need for multistage study designs. Third, we performed our study in a predominantly white population. Ethnicity is an important confounder in genetic studies and established associations are often not replicable in other ethnic populations. Therefore, findings should be replicated in different populations to establish a causal relationship with the disease phenotype. Although we carefully matched our discovery population, is it possible that other factors such as ethnicity are responsible for the observed difference in allele frequency.

To our knowledge, this is the first attempt to discover novel genetic associations by whole-exome sequencing in a population of patients with acute pancreatitis. Although, we identified strong associations with *ZNF106* and *SLC52A1* variants, replication in other populations and efforts to determine whether it is a causal risk factor for disease progression are warranted. If so, it might direct the development of highly needed therapies to prevent MOF in the early phase of acute pancreatitis. Especially biological pathways that can easily be targeted, such as the riboflavin synthesis pathway, are of clinical interest for its potential to intervene in the disease process.

ACKNOWLEDGMENTS

The authors thank Prof. Dr. Raoul Hennekam for help with the study design and Dr. Aldo Jongejan for help with the bioinformatics analysis.

REFERENCES

1. Banks PA, Bollen TL, Dervenis C, et al. Classification of acute pancreatitis—2012; revision of the Atlanta classification and definitions by international consensus. *Gut*. 2013;62:102–111.
2. Johnson CD, Abu-Hilal M. Persistent organ failure during the first week as a marker of fatal outcome in acute pancreatitis. *Gut*. 2004;53:1340–1344.
3. Bishehsari F, Sharma A, Stello K, et al. TNF-alpha gene (TNFA) variants increase risk for multi-organ dysfunction syndrome (MODS) in acute pancreatitis. *Pancreatology*. 2012;12:113–118.
4. Lu H, Wen D, Wang X, et al. Host genetic variants in sepsis risk: a field synopsis and meta-analysis. *Crit Care*. 2019;23:26.
5. Ma M, Zhai CX, Sun CX, et al. Correlations between LP-PLA2 gene polymorphisms and susceptibility and severity of acute pancreatitis in a Chinese population. *Gastroenterol Res Pract*. 2017;21:206–212.
6. Liu YG, Dan G, Wu L, et al. Functional effect of polymorphisms in the promoter of TNFAIP3 (A20) in acute pancreatitis in the Han Chinese population. *PLoS One*. 2014;9:e103104.
7. Zhang DL, Zheng HM, Yu BJ, et al. Association of polymorphisms of IL and CD14 genes with acute severe pancreatitis and septic shock. *World J Gastroenterol*. 2005;11:4409–4413.
8. de-Madaria E, Martinez J, Sampere L, et al. Cytokine genotypes in acute pancreatitis: association with etiology, severity, and cytokine levels in blood. *Pancreas*. 2008;37:295–301.
9. van Baal MC, et al. Timing of cholecystectomy after mild biliary pancreatitis: a systematic review. *Ann Surg*. 2012;255:860–866.
10. Besselink MG, van Santvoort HC, Buskens E, et al. Probiotic prophylaxis in predicted severe acute pancreatitis: a randomised, double-blind, placebo-controlled trial. *Lancet (London England)*. 2008;371:651–659.
11. van Santvoort HC, Besselink MG, Bakker OJ, et al. A step-up approach or open necrosectomy for necrotizing pancreatitis. *N Engl J Med*. 2010;362:1491–1502.
12. Li M, Li J, Li MJ, et al. Robust and rapid algorithms facilitate large-scale whole genome sequencing downstream analysis in an integrative framework. *Nucleic Acids Res*. 2017;45:e75.
13. F RC Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing; 2016.
14. van den Berg FF, Kempeneers MA, van Santvoort HC, et al. Meta-analysis and field synopsis of genetic variants associated with the risk and severity of acute pancreatitis. *BJS Open*. 2020;4:3–15.
15. Matas-Cobos AM, Redondo-Cerez E, Alegria-Motte C, et al. The role of Tolllike receptor polymorphisms in acute pancreatitis occurrence and severity. *Pancreas*. 2015;44:429–433.

© *2021 The Author(s). Published by Wolters Kluwer Health, Inc.*

16. Masamune A, Zhang CX, Guo LK, et al. -651C/T promoter polymorphism in the CD14 gene is associated with severity of acute pancreatitis in Japan. *J Gastroenterol*. 2010;45:225–233.
17. Zhang C, Guo L, Qin Y, et al. Interaction between polymorphisms of TLR4 gene G11367C in 3' untranslated region and IkappaB-alpha Hae III in acute pancreatitis and the degree of severity. [Chinese]. *Zhong Nan Da Xue Xue Bao Yi Xue Ban*. 2016;41:272–281.
18. Zhang CX, Guo LK, Zhang LL, et al. Interaction of polymorphisms of TNF-alpha gene promoter -308G/A and PPAR-gamma2 gene -C34G with acute pancreatitis and its severity degree. [Chinese]. *Lancet*. 2017;38:76–82.
19. Robinson JT, Thorvaldsdottir H, Winckler W, et al. Integrative genomics viewer. *Nat Biotechnol*. 2011;29:24–26.
20. Fagerberg L, Hallstrom BM, Oksvold P, et al. Analysis of the human tissuespecific expression by genome-wide integration of transcriptomics and antibody-based proteomics. *Mol Cell Proteomics*. 2014;13:397–406.
21. Mosegaard S, Bruun GH, Flyvbjerg KF, et al. An intronic variation in SLC52A1 causes exon skipping and transient riboflavin-responsive multiple acyl-CoA dehydrogenation deficiency. *Mol Genet Metab*. 2017;122:182–188.
22. Chiong MA, Sim KG, Carpenter K, et al. Transient multiple acyl-CoA dehydrogenation deficiency in a newborn female caused by maternal riboflavin deficiency. *Mol Genet Metab*. 2007;92:109–114.
23. Constantinides MG, et al. MAIT cells are imprinted by the microbiota in early life and promote tissue repair. *Science*. 2019. 366. doi:10.1126/science.aax6624.
24. Kjer-Nielsen L, Patel O, Corbett AJ, et al. MR1 presents microbial vitamin B metabolites to MAIT cells. *Nature*. 2012;491:717–723.
25. Toyosawa T, Suzuki M, Kodama K, et al. Highly purified vitamin B2 presents a promising therapeutic strategy for sepsis and septic shock. *Infection and immunity*. 2004;72:1820–1823.
26. Anderson DM, Cannavino J, Li H, et al. Severe muscle wasting and denervation in mice lacking the RNA-binding protein ZFP106. *Proc Natl Acad Sci U S A*. 2016;113:E4494–E4503.
27. Hebbar P, Elkum N, Alkayal F, et al. Genetic risk variants for metabolic traits in Arab populations. *Sci Rep*. 2017;7:40988.
28. Giamarellos-Bourboulis EJ, Opal SM. The role of genetics and antibodies in sepsis. *Ann Transl Med*. 2016;4:328.
29. Liu Y, Yang L, Chen K-L, et al. Knockdown of GRP78 promotes apoptosis in pancreatic acinar cells and attenuates the severity of cerulein and LPS induced pancreatic inflammation. *PLoS One*. 2014;9:e92389.
30. Fortunato F, Burgers H, Bergmann F, et al. Impaired autolysosome formation correlates with Lamp-2 depletion: role of apoptosis, autophagy, and necrosis in pancreatitis. *Gastroenterology*. 2009;137:350–360.
31. Liu Y, Chen X-D, Yu J, et al. Deletion of XIAP reduces the severity of acute pancreatitis via regulation of cell death and nuclear factor-kappaB activity. *Cell Death Dis*. 2017;8:e2685.
32. Wang G, Qu FZ, Li L, et al. Necroptosis: a potential, promising target and switch in acute pancreatitis. *Apoptosis*. 2016;21:121–129.
33. Adzhubei IA, Schmidt S, Peshkin L, et al. A method and server for predicting damaging missense mutations. *Nat Methods*. 2010;7:248–249.
34. Vaser R, Adusumalli S, Leng SN, et al. SIFT missense predictions for genomes. *Nat Protoc*. 2016;11:1–9.
35. Flanagan SE, Patch AM, Ellard S. Using SIFT and PolyPhen to predict loss-of-function and gain-of-function mutations. *Genet Test Mol Biomarkers*. 2010;14:533–537.
36. Li MX, Kwan JSH, Bao S-Y, et al. Predicting mendelian disease-causing non-synonymous single nucleotide variants in exome sequencing studies. *PLoS Genet*. 2013;9:e1003143.
37. Hu Y, Lu Q, Powles R, et al. Leveraging functional annotations in genetic risk prediction for human complex diseases. *PLoS Comput Biol*. 2017;13:e1005589.
38. Schrodi SJ, Mukherjee S, Shan Y, et al. Genetic-based prediction of disease traits: prediction is very difficult, especially about the future. *Front Genet*. 2014;5:162.
39. Consortium EP. An integrated encyclopedia of DNA elements in the human genome. *Nature*. 2012;489:57–74.

© 2021 The Author(s). Published by Wolters Kluwer Health, Inc.