ORIGINAL ARTICLE

Whole Exome Sequencing Identifies Multiple, Complex Etiologies in an Idiopathic Hereditary Pancreatitis Kindred

Jessica LaRusch[1], M Michael Barmada[2], Sheila Solomon[1], David C Whitcomb[1,2,3]

[1]Division of Gastroenterology, Hepatology and Nutrition, Department of Medicine, [2]Department of Human Genetics, [3]Department of Molecular Cell Biology and Physiology; University of Pittsburgh. Pittsburgh, PA, USA

ABSTRACT

Context Hereditary pancreatitis is the early onset form of chronic pancreatitis that is carried in an autosomal dominant pattern with variable penetrance. While 80% of hereditary pancreatitis has been shown to be due to a single mutation in the trypsinogen gene *PRSS1*, a number of hereditary pancreatitis families have no identified genetic cause for illness; thus no reliable screening options or clear therapy. **Objective** To explore the use of massive parallel DNA sequencing technology to discover the etiology of pancreatitis in a family with idiopathic hereditary pancreatitis. **Design** Candidate gene screening and verification within a kindred. **Setting** Prospective cohort study, university based. **Patients or participants** Kindred with idiopathic hereditary pancreatitis. **Interventions** None. **Main outcome measures** Identification of DNA variants predicted to increase susceptibility to pancreatitis. **Methods** Whole exome sequencing of two distantly related subjects with variant-specific confirmation in the subjects and other family members. **Results** We identified three deleterious genetic changes in the three major pancreatitis associated genes (*PRSS1* CNV, *SPINK1* c.27delC and *CFTR* R117H), two of which were carried by each patient. Individual targeted assays confirmed these variations in the two whole exome sequencing patients as well as affected and non-affected pedigree members. **Conclusion** Whole exome sequencing was useful for rapid screening of candidate genes linked to pancreatitis. This method opens the door for time- and cost-effective screening of multiple disease-associated genes and modifying factors that associate in different ways to generate a complex genetic disorder.

INTRODUCTION

Most hereditary pancreatitis families have gain-of-function mutations in the cationic trypsinogen gene (protease serine 1 (trypsin I precursor); *PRSS1*) [1, 2, 3, 4]. Familial clustering of pancreatitis occurs with various combinations of pancreatic secretory trypsin inhibitor gene (serine protease inhibitor, Kazal type 1; *SPINK1*) mutations [5] and/or severe or atypical cystic fibrosis transmembrane conductance regulator (*CFTR*) mutations (*CFTR*-CF) [6, 7, 8]. Heterozygous or complex *SPINK1*, *CFTR*, chymotrypsinogen C (*CTRC*), and calcium sensing receptors (*CASR*), variants are also seen in idiopathic pancreatitis [8, 9].

To add to the complexity, there is increasing recognition that rare or private mutations in known pancreatitis susceptibility genes are associated with risk of pancreatitis, as demonstrated for *CTRC* [10]. Finally, copy number variants in *PRSS1* increase risk of pancreatitis and must also be considered [11].

As the number of genes and mutations involved in the onset and progression of pancreatitis becomes higher, the time and cost of screening and sequencing specific exons continues to increase. However, a new approach that uses massive parallel sequencing called next generation sequencing (NGS) is becoming standardized, and the cost per patient is rapidly dropping. NGS includes whole genome sequencing, whole exome sequencing (WES) and other methods. Because the cost of WES in our hands is now less than the cost of sequencing *CFTR* use of this technology is becoming an attractive alternative to classic targeted gene sequencing or mutation specific genotyping for a genetic counseling workup. The work presented in this article explores the feasibility and utility of WES to identify the cause of disease in a hereditary pancreatitis family.

The J family was composed of three generations, including four living pancreatitis patients and was typical for hereditary pancreatitis (Figure 1). Direct
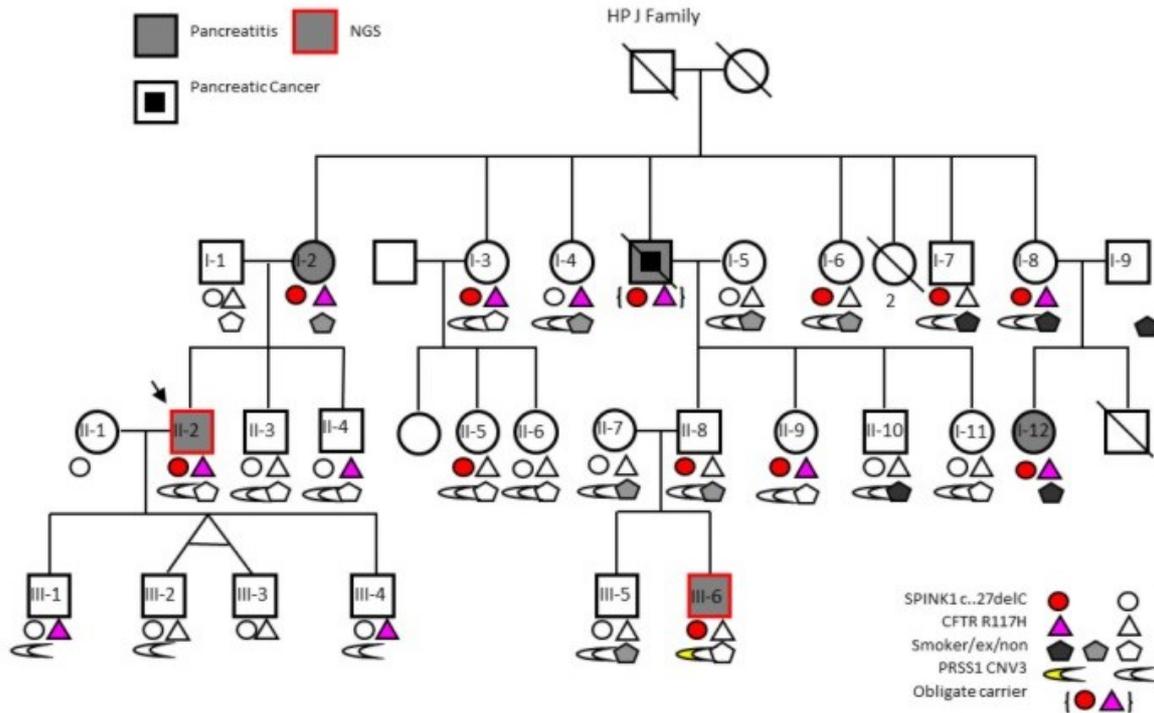
Figure 1. J family pedigree and genotyping/smoking data.

sequencing of all exons of *PRSS1* and exons 2-3 of *SPINK1* sequencing resulted in no apparent familial mutation. Screening of *CFTR* revealed a non-CF causing p.R117H variant (T7), but this did not segregate exclusively with disease. Linkage analysis suggested a possible shared locus on chromosome 12, but the wide variety of phenotypic features (age of onset and symptom severity) confounded definitive phenotyping in key family members. Onset of pancreatitis symptoms was in the teens for both subjects chosen for WES. To determine whether this family had undetected variants in known pancreatitis susceptibility genes, we conducted WES, focusing our analysis on *PRSS1* (5 exons), *SPINK1* (4 exons), *CFTR* (27 exons), *CTRC* (8 exons), and *CASR* (7 exons).

METHODS

Patients

Family J from the Hereditary Pancreatitis (HP) study [12] was selected for NGS using the following criteria: direct sequencing of all exons in *PRSS1* and exons 2-3 in *SPINK1* were negative for any potential disease-causing mutation; large family with multiple affected family members in different generations; and a pedigree that suggests a complex inheritance pattern or incomplete penetrance. DNA of cases II-1 and III-6 were sent for NGS.

Whole Exome Sequencing

DNA (3 µg) was enriched using the Agilent (Santa Clara, CA, USA) SureSelect Human All Exon kit optimized for the SOLiD Sequencing System (AB Life

Technologies, Foster City, CA, USA). This genome-wide approach targets the exons of about 18,000 known genes, covering 38 Mb of the human genome. The prepared exome library was emulsion amplified following the manufacturer's directions for the EZBead Amplifier (AB Life Technologies) at a 0.3 pM library concentration. Each sample was sequenced on one quadrant (1/4) of a SOLiD 4 sequencing slide, and we obtained 3.1 and 2.2 Gb of mappable sequence. Color space reads were mapped to the reference human genome (NCBI GRCh37.2) and variants (both single nucleotide polymorphisms (SNP) and indels) were called using the CLCbio (Germantown, MD, USA) Genomics Workbench software (version 4.6).

Confirmation of WES Data

WES results were confirmed by individual assay, testing all samples of family J with available DNA as well subsets of Hereditary Pancreatitis Study (HP: Dr Whitcomb, principal investigator) and North American Pancreatitis Study II (NAPS2; Dr Whitcomb, principal investigator) cohorts for *SPINK1* c.27delC, *PRSS1* CNV and *CFTR* R117H.

SPINK1 c.27delC

An individual Taqman® (AB Life Technologies) assay was designed for the c.27delC mutation. Data were analyzed using SDS2.3 software (AB Life Technologies).

PRSS1 CNV

A pre-designed and validated CNV assay located in exon 5 of PRSS1 (AB Life Technologies;

Hs03192321_cn) was used. *PRSS1* copy number was calculated in proportion to RNAseP reference assay (AB Life Technologies cat#4403326): The assays were performed in triplicate and repeated at least once for each sample. Data were analyzed using Copy Caller Software (version 1.0; AB Life Technologies).

CFTR R117H

A multiplex PCR-based SNP assay (Sequenom®, AB Life Technologies) was used. Primer sequences available on request.

ETHICS

The study was approved by the Institutional Review Board of the University of Pittsburgh prior to patient contact or ascertainment. All subjects provided written informed consent that conformed to the ethical guidelines of the "World Medical Association Declaration of Helsinki - Ethical Principles for Medical Research Involving Human Subjects".

Figure 2. WES and confirmatory CNV of PRSS1. **a.** PRSS1 coverage of case III-6. Observed coverage of sequencing fragments from WES in reference region (T-cell receptor beta locus: *TRB*) and *PRSS1*. Mean coverage of *PRSS1* (38.67) was significantly higher than in *TRB* (20.60), P=0.004 using Welch two sample t-test (95% CI of the difference: 6.04-30.09). Box-whisker plot: boxes are defined by the second and third quartile sections of the data; whiskers indicate the first and 4th quartile sections of the data. The cross bar is the mean and the notches are indicators of the standard error of the mean. Non-overlap of notches is an indicator of significant differences in distribution using non-parametric approaches. Circles are considered outliners. **b.** *PRSS1* copy number variant assay of pedigree 21. *PRSS1* copy number is plotted for family members, and an unrelated hereditary pancreatitis *PRSS1* R122H case. All samples had the expected copy number of 2, except III-6, with CNV3. The mean and min/max results for each assay are shown.

STATISTICS

Statistical analyses (Welch two sample t-test) were done using the R statistical package [13] (http://www.R-project.org/). Boxplots (figure 2a) were also produced in R, and follow the conventions of the boxplot function (from the default graphics library). Copy number variants were compared using CopyCaller 1.0 softward (Applied Biosystems, Carlsbad, CA, USA).

RESULTS

No significant mutations in *CTRC*, *CASR* or *PRSS1* were identified. We identified a *SPINK1* c.27delC (a one-base deletion in exon 1 of *SPINK1*), shared in a heterozygous state, in both patients. This variant was confirmed by Taqman® (AB Life Technologies) in these two patients and 11 additional family members and was implicated in one obligate carrier. The apparent penetrance of c.27delC for pancreatitis in this family was 42%. This frameshift mutation is expected to cause nonsense mediated decay of the mRNA message, reducing cellular Spink1 protein levels to 50% of normal. *SPINK1* c.27delC was previously reported in 2 French hereditary pancreatitis families with variable penetrance [14]. Upon testing 713 subjects from other hereditary pancreatitis families and 1,550 subjects from the North American Pancreatitis Study 2 cohort [15], we found a single additional carrier, a 47-year-old female with recurrent acute pancreatitis, no other identified genetic mutations, and no family history of pancreatic disease.

The *CFTR* variant R117H was identified in case II-2 but not III-6. R117H is a frequent variant in the general population and has been associated with cystic fibrosis but only when *in cis* with a T5 allele [16]; however, the R117H in this family was not associated with T5. This CFTR variant was confirmed via a multiplex Sequenom® (AB Life Technologies) assay and was subsequently identified in 10 additional family members and implicated in one obligate carrier. The apparent penetrance of R117H alone for pancreatitis was 44% in this family.

Patient III-6 was also found to carry a *PRSS1* copy number variant, as demonstrated by the increased average coverage of this gene in the WES platform: coverage across *PRSS1* was significantly elevated with respect to other genes in the region (P=0.004, Figure 2a). A pre-designed and validated copy number variant assay targeting *PRSS1* (AB Life Technologies; Hs03192321_cn) confirmed a statistically significant increased copy number in this patient but in no other family member (Figure 2b); it estimated a copy number of 3 for III-6 (confidence score of 0.67 and a |z| score of 0.0 for 3 copies of *PRSS1*), and a copy number of 2 for all other family members. *PRSS1* copy number variants have been reported in a small number of hereditary pancreatitis families around the world and a 50% increase in basal trypsin expression is expected to be causal for pancreatitis, but this is the first report of a single apparent *de novo* duplication within an

established hereditary pancreatitis family with no other copy number variants. Testing the rest of the hereditary pancreatitis and NAPS2 subjects for *PRSS1* CNV revealed a single additional family with two affected twin sons and a copy number of 3 for each.

Heterozygous loss-of-function mutations in *CFTR, SPINK1, CTRC*, and *CASR* are commonly identified in control populations, with epistasis seen between *CFTR* and *SPINK1* variants [8] and *CASR* and *SPINK1* variants [17]. In this family, combined penetrance of 2 or more of these variants was 62.5%.

Smoking and drinking are strong independent risk factors for pancreatitis [18, 19]. All but one of the c.27delC/R117H carriers who were affected also self-reported being current or former smokers, while 2/3 of the unaffected c.27delC/R117H carriers were non-smokers. However, alcohol usage did not correlate with disease status: 2/3 of the c.27delC/R117H unaffected carriers were light drinkers, and 2/3 of the affected carriers were never-drinkers, placing all of the subjects well below the risk-threshold for alcohol-associated pancreatitis [19]. Thus, each affected patient carried unique genetic and environmental factors that individually would not necessarily cause pancreatitis but combined to confer very high risk.

DISCUSSION

These data demonstrate an application of WES in evaluating the entire coding region of 5 confirmed pancreatitis susceptibility genes. This includes discovery of a rare *SPINK1* mutation, a common *CFTR* mutation and a single *PRSS1* copy number mutation in a single assay. In addition, other mutations in *PRSS1, CTSR* and *CASR* were excluded. Furthermore, any newly reported variant could be quickly checked in these individuals. Finally, the use of this technology in an idiopathic hereditary pancreatitis family was chosen because the risk of rare variants could be validated in other family members. This approach will also be useful in sporadic patients where known variants are identified, or the functional effect of a rare mutation can be determined.

An important application of WES is in the guidance of clinical decision-making, particularly in managing complex chronic inflammatory disorders. In pancreatitis, clinicians are currently limited to broad interventions, such as minimizing pancreatic stimulation or improving drainage. Knowledge of contributory genetic variants could allow them to take advantage of etiology-based treatment strategies being developed that target the acinar (e.g., *PRSS1* variants [20]) or duct cells (e.g., *CFTR* variants [21]). Our use of WES with validation indicated different etiologies in different patients, and therefore different approaches. Also, WES saved significant time at about the same cost compared with direct sequencing and copy number analysis of individual candidate genes (more details are reported in the Note statement below). In addition, individual III-6 now presents completely different risks for disease progression and inheritance as compared to

that of the rest of the family, and the genetic risk for pancreatitis for the rest of the third generation of this family can be reported as substantially reduced.

While the standard practice in genotyping hereditary pancreatitis families is presupposing a single variant is fully responsible for disease, this would have failed in the case of the J family. There is growing evidence the hereditary pancreatitis families with an absence of a standard (R122H) *PRSS1* mutation (and possibly even those with) require much more consideration and have a more complex pathology, with multiple genetic variants and environmental pressures converging into a small family set. Care must be taken in interpreting *PRSS1* variants since there are nine trypsinogen-like genes and pseudo-genes with nearly identical sequence. Thus, all sequence variants must be verified and custom genotyping methods are required to detect the major *PRSS1* gain-of-function mutations if hereditary pancreatitis is suspected.

As insight into genetic susceptibility to pancreatitis grows the importance of genotyping patients will also grow. We are evaluating the benefit of genotyping patients at the beginning of their clinical evaluation because a positive test immediately completes the diagnostic work-up, and may provide direction for etiology-based treatment strategies in the future (e.g., aimed at the duct cells or aimed at the acinar cell).

In summary, using WES in two members of a family with hereditary pancreatitis, we identified complex etiologies with epistasis between common and rare mutations in *CFTR* and *SPINK1*, copy number variants in *PRSS1*, and a relationship to smoking but not alcohol use. We have thus far limited the study to basic candidate genes for pancreatitis, and even with this temporarily restricted approach, we have located significant genetic lesions that have not been identified through basic techniques. There is a notable decrease in time needed for NGS of our two patients as opposed to direct sequencing and copy number analysis of individual candidate genes, the cost is comparable, and now we have an extensive database in which to mine for additional candidate lesions in this family in the near future. Moving forward, as new susceptibility factors and modifier genes are identified, the patients evaluated by WES can be rapidly and inexpensively screened.

Note The cost of candidate gene sequencing in our research laboratory for individual patients is driven by laboratory cost of sequencing each exon in both directions, and verifying the quality of each sequencing read, in addition to the cost of materials for DNA sequencing. Commercial clinical laboratories (e.g., Ambry Genetics, Aliso Viejo, CA, USA) evaluate for PRSS1, SPINK1, CFTR and CTRC plus CFTR del/dup about $3,500. WES cost through our core lab is less than $2,000 per person, but this data is analyzed by our group and cannot be used for clinical decision making without verification of the variants in a clinical laboratory

Conflict of interest There are no conflicts of interest to disclose

Reference

1. Whitcomb DC, Gorry MC, Preston RA, Furey W, Sossenheimer MJ, Ulrich CD, et al. Hereditary pancreatitis is caused by a mutation in the cationic trypsinogen gene. Nature Genetics. 1996;14(2):141-5. [PMID: 8841182]

2. Howes N, Lerch MM, Greenhalf W, Stocken DD, Ellis I, Simon P, et al. Clinical and genetic characteristics of hereditary pancreatitis in Europe. Clin Gastroenterol Hepatol. 2004;2(3):252-61. [PMID 15017610]

3. Rebours V, Boutron-Ruault MC, Schnee M, Ferec C, Le Marechal C, Hentic O, et al. The natural history of hereditary pancreatitis: a national series. Gut. 2009 Jan;58(1):97-103. [PMID: 18755888]

4. Joergensen MT, Brusgaard K, Cruger DG, Gerdes AM, Schaffalitzky de Muckadell OB. Genetic, epidemiological, and clinical aspects of hereditary pancreatitis: a population-based cohort study in Denmark. The American journal of gastroenterology. 2010 Aug;105(8):1876-83. [PMID: 20502448]

5. Witt H, Luck W, Hennies HC, Classen M, Kage A, Lass U, et al. Mutations in the gene encoding the serine protease inhibitor, Kazal type 1 are associated with chronic pancreatitis. Nature Genetics. 2000;25(2):213-6. [PMID 10835640]

6. Sharer N, Schwarz M, Malone G, Howarth A, Painter J, Super M, et al. Mutations of the cystic fibrosis gene in patients with chronic pancreatitis. New England Journal of Medicine. 1998;339(10):645-52. [PMID 9725921]

7. Cohn JA, Friedman KJ, Noone PG, Knowles MR, Silverman LM, Jowell PS. Relation between mutations of the cystic fibrosis gene and idiopathic pancreatitis. New England Journal of Medicine. 1998;339(10):653-8. [PMID 9725922]

8. Schneider A, Larusch J, Sun X, Aloe A, Lamb J, Hawes R, et al. Combined Bicarbonate Conductance-Impairing Variants in CFTR and SPINK1 Variants Are Associated With Chronic Pancreatitis in Patients Without Cystic Fibrosis. Gastroenterology. 2011 Jan;140(1):162-71. [PMID: 20977904]

9. Whitcomb DC. Genetic aspects of pancreatitis. Annu Rev Med. 2010;61:413-24. [PMID: 20059346]

10. Masson E, Chen JM, Scotet V, Le Marechal C, Ferec C. Association of rare chymotrypsinogen C (CTRC) gene variations in patients with idiopathic chronic pancreatitis. Hum Genet. 2008 Feb;123(1):83-91. [PMID: 18172691]

11. Masson E, Le Marechal C, Delcenserie R, Chen JM, Ferec C. Hereditary pancreatitis caused by a double gain-of-function trypsinogen mutation. Hum Genet. 2008 Jun;123(5):521-9. [PMID: 18461367]

12. Applebaum-Shapiro SE, Finch R, Pfützer RH, Hepp LA, Gates L, Amann S, et al. Hereditary Pancreatitis in North America: The Pittsburgh - Midwest Multi-Center Pancreatic Study Group Study. Pancreatology. 2001;1(5):439-43. [PMID 12120221]

13. R Development Core Team (2011). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/

14. Le Marechal C, Chen JM, Le Gall C, Plessis G, Chipponi J, Chuzhanova NA, et al. Two novel severe mutations in the pancreatic secretory trypsin inhibitor gene (SPINK1) cause familial and/or hereditary pancreatitis. Hum Mutat. 2004 Feb;23(2):205. [PMID: 14722925]

15. Whitcomb DC, Yadav D, Adam S, Hawes RH, Brand RE, Anderson MA, et al. Multicenter approach to recurrent acute and chronic pancreatitis in the United States: the North American Pancreatitis Study 2 (NAPS2). Pancreatology. 2008;8(4-5):520-31. [PMID: 18765957]

16. Thauvin-Robinet C, Munck A, Huet F, Genin E, Bellis G, Gautier E, et al. The very low penetrance of cystic fibrosis for the R117H mutation: a reappraisal for genetic counselling and newborn screening. J Med Genet. 2009 Nov;46(11):752-8. [PMID: 19880712]

17. Felderbauer P, Klein W, Bulut K, Ansorge N, Dekomien G, Werner I, et al. Mutations in the calcium-sensing receptor: a new genetic risk factor for chronic pancreatitis? Scand J Gastroenterol. 2006 Mar;41(3):343-8. [PMID: 16497624]

18. Talamini G, Bassi C, Falconi M, Frulloni L, Di Francesco V, Vaona B, et al. Cigarette smoking: an independent risk factor in alcoholic pancreatitis. Pancreas. 1996;12(2):131-7. [PMID 8720658]

19. Yadav D, Hawes RH, Brand RE, Anderson MA, Money ME, Banks PA, et al. Alcohol consumption, cigarette smoking, and the risk of recurrent acute and chronic pancreatitis. Arch Intern Med. 2009 Jun 8;169(11):1035-45. [PMID: 19506173]

20. Morinville VD, Lowe ME, Elinoff BD, Whitcomb DC. Hereditary pancreatitis amlodipine trial: a pilot study of a calcium-channel blocker in hereditary pancreatitis. Pancreas. 2007 Nov;35(4):308-12. [PMID: 18090235]

21. Accurso FJ, Rowe SM, Clancy JP, Boyle MP, Dunitz JM, Durie PR, et al. Effect of VX-770 in persons with cystic fibrosis and the G551D-CFTR mutation. N Engl J Med. 2010 Nov 18;363(21):1991-2003. [PMID: 21083385]