



Published in final edited form as:

*Gastroenterology*. 2015 November ; 149(6): 1501–1510. doi:10.1053/j.gastro.2015.07.041.

## A Combination of Molecular Markers and Clinical Features Improve the Classification of Pancreatic Cysts

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

### Abstract

**Background & Aims**—The management of pancreatic cysts poses challenges to both patients and their physicians. We investigated whether a combination of molecular markers and clinical information could improve the classification of pancreatic cysts and management of patients.

**Methods**—We performed a multi-center, retrospective study of 130 patients with resected pancreatic cystic neoplasms (12 serous cystadenomas, 10 solid-pseudopapillary neoplasms, 12 mucinous cystic neoplasms, and 96 intraductal papillary mucinous neoplasms). Cyst fluid was analyzed to identify subtle mutations in genes known to be mutated in pancreatic cysts (*BRAF*, *CDKN2A*, *CTNNB1*, *GNAS*, *KRAS*, *NRAS*, *PIK3CA*, *RNF43*, *SMAD4*, *TP53* and *VHL*); to identify loss of heterozygosity at *CDKN2A*, *RNF43*, *SMAD4*, *TP53*, and *VHL* tumor suppressor loci; and to identify aneuploidy. The analyses were performed using specialized technologies for

---

**Address correspondence to.** Dr. Lennon, Dr. Vogelstein, Dr. Karchin or Dr. Hruban, The Ludwig Center and Howard Hughes Medical Institute at the Sidney Kimmel, Cancer Center, The Johns Hopkins Medical Institutions, Baltimore, MD, USA, Tel: 410-955-5800, Fax: 410-614-8337, amlennon@jhmi.edu; vogelbe@jhmi.edu; rachel.karchin@gmail.com; rhruban@jhmi.edu.

\*Drs. Springer, Wang, Dal Molin and Masica contributed equally as first authors

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

The terms of these arrangements are being managed by the university in accordance with its conflict of interest policies.

### Disclosures

Drs. Springer, Wang, Dal Molin, Masica, Jiao, Kinde, Blackford, Raman, Wolfgang, Tomita, Niknafs, Douville, Ptak, Dobbyn, Allen, Klimstra, Schattner, Schmidt, Yip-Schneider, Cummings, Brand, Zeh, Singhi, Scarpa, Salvia, Malleo, Zamboni, Falconi, Jang, Kim, Kwon, Hong, Song, Kim, Swan, Murphy, Geoghegan, Brugge, Fernandez-Del Castillo, Mino-Kenudson, Schulick, Edil, Adsay, Paulino, van Hooft, Yachida, Nara, Hiraoka, Yamao, Hijioka, van der Merwe, Canto, Ahuja, Hirose, Makary, Weiss, Cameron, Pittman, Eshleman, Karchin, and Lennon have no disclosures.

**Writing assistance:** None

### Author contributions

Drs. Springer, Wang, Dal Molin, Jiao, Kinde, Raman, Tomita, Niknafs, Douville, Pittman, Ptak, Dobbyn, Hruban, Lennon and Vogelstein, were involved in the study conception, performed data acquisition, and were involved in editing and final approval of the manuscript.

Drs. Springer, Wang, Dal Molin, Jiao, Kinde, Raman, Tomita, Niknafs, Douville, Pittman, Ptak, Dobbyn, Hruban, Lennon and Vogelstein, were involved in the study conception, performed data acquisition, and were involved in editing and final approval of the manuscript. Drs. Springer, Wang, Dal Molin, Masica, Jiao, Kinde, Blackford, Raman, Wolfgang, Tomita, Niknafs, Douville, Ptak, Dobbyn, Allen, Klimstra, Schattner, Schmidt, Yip-Schneider, Cummings, Brand, Zeh, Singhi, Scarpa, Salvia, Malleo, Zamboni, Falconi, Jang, Kim, Kwon, Hong, Song, Kim, Swan, Murphy, Geoghegan, Brugge, Fernandez-Del Castillo, Mino-Kenudson, Schulick, Edil, Adsay, Paulino, Hooft, Yachida, Nara, Hiraoka, Yamao, Hijioka, van der Merwe, Goggins, Canto, Ahuja, Hirose, Makary, Weiss, Cameron, Pittman, Eshleman, Diaz, Papadopoulos, Kinzler, Karchin, Hruban, Vogelstein, and Lennon were involved in the conception and design of the study, data interpretation, editing and final approval of the manuscript.

Drs. Masica, Karchin and Blackford were involved in the study design, performed the statistical analysis, edited and approved the final manuscript.

Drs. Lennon, Hruban, Karchin and Vogelstein developed the study design, generated and interpreted the data, wrote and approved the final version of the manuscript.

implementing and interpreting massively parallel sequencing data acquisition. An algorithm was used to select markers that could classify cyst type and grade. The accuracy of the molecular markers were compared with that of clinical markers, and a combination of molecular and clinical markers.

**Results**—We identified molecular markers and clinical features that classified cyst type with 90%–100% sensitivity and 92%–98% specificity. The molecular marker panel correctly identified 67 of the 74 patients who did not require surgery, and could therefore reduce the number of unnecessary operations by 91%.

**Conclusions**—We identified a panel of molecular markers and clinical features that show promise for the accurate classification of cystic neoplasms of the pancreas and identification of cysts that require surgery.

### Keywords

molecular; IPMN; diagnosis; pancreatic cyst

---

### Introduction

Pancreatic cysts have been reported as incidental findings in 3% to 13% of individuals undergoing computed tomography (CT) or magnetic resonance imaging (MRI).<sup>1, 2</sup> The four most common types of neoplastic cysts of the pancreas are serous cystadenomas (SCA), solid-pseudopapillary neoplasm (SPN), mucinous cystic neoplasm (MCN) and intraductal papillary mucinous neoplasms (IPMN). SCAs have a very small risk of malignant transformation, and surveillance is usually recommended for these cysts in asymptomatic patients.<sup>3</sup> SPNs are low-grade malignant neoplasms<sup>4</sup>, and should be surgically resected when possible.<sup>5</sup> MCNs have the potential to progress to malignancy<sup>6</sup>, and current guidelines recommend surgical resection if possible.<sup>7</sup> IPMNs can progress from low, to intermediate, to high-grade dysplasia, and ultimately to invasive adenocarcinoma.<sup>8</sup> Ideally, it is recommended that IPMNs with high-grade dysplasia, or an associated invasive adenocarcinoma, should undergo resection, while IPMNs with low- or intermediate-grade dysplasia should undergo surveillance.<sup>9, 10</sup>

The clinical management of patients with pancreatic cysts is currently based on clinical presentation, imaging and cyst fluid analysis.<sup>7</sup> However, this approach is imperfect. For example, an evaluation of surgically resected pancreatic cysts at a high-volume center found that over 20% of the cysts resected due to concerns about their malignant potential, were entirely benign on histopathologic examination; in hind-sight, these cysts could have been safely observed.<sup>11</sup> Similarly, over 75% of resected IPMNs harbor only low-, or intermediate-grade dysplasia, and these also could have been safely observed.<sup>10</sup> Thus, better diagnostic tools are required to determine which patients truly benefit from surgical resection and which patients can be safely observed.

We previously performed whole exome sequencing of a well-characterized series of SCAs, SPNs, MCNs and IPMNs, and identified a distinct mutational profile in each cyst type.<sup>12, 13</sup> For example, *VHL* alterations were characteristic of SCAs,  $\beta$ -catenin gene (*CTNNB1*)

mutations were found in SPNs, *GNAS* mutations in IPMNs, and *KRAS* and *RNF43* alterations were observed in both IPMNs and MCNs.

We have previously used an algorithm for analyzing multi-parametric features (MOCA, for Multivariate Organization of Combinatorial Alterations) to identify composite clinical markers of pancreatic cyst type and grade from a 1026-patient cohort.<sup>14</sup> These composite markers, which included only clinical features such as age, sex, symptoms, and radiologic appearance, identified the common cyst types with high but imperfect accuracy (84% to 92%), and also correctly identified which cysts needed surgical resection with 82% accuracy.

The aim of the current study was to determine whether the molecular genetic features of pancreatic cyst fluid could be used to classify cysts and identify those that require surgical resection. Furthermore, we wished to determine whether combining the molecular markers identified here with the clinical markers identified by Masica et al, would increase the accuracy of diagnosis over either one alone.<sup>14</sup>

## Methods

### Patients

The study was approved by the Institutional Review Boards for Human Research (IRB numbers: 00001584, 00-032, 1011003217, MOD07030072-52), and complied with Health Insurance Portability and Accountability Act.

Patients with SCAs, SPNs, MCNs or IPMNs who had undergone surgical resection at the participating institutions between September 2004 and September 2013 were included in the study. General demographics, the presence of symptoms, CT, MRI, endoscopic ultrasound (EUS) features, cytology and cyst fluid carcinoembryonic antigen (CEA) levels were documented. When available, cross sectional imaging studies were reviewed by a single, experienced abdominal CT/MRI radiologist (S.R.), and the EUS studies reviewed by an experienced pancreatic endosonographer (A.M.L.). Multiple pancreatic cysts were defined as the presence of more than one cyst within the pancreas, which were anatomically separate from each other. The pathology of the surgically resected lesions was reviewed by one of two pancreatic pathologists (R.H.H. or D.K.).<sup>15</sup> The decision to resect a pancreatic cyst is multifactorial, and includes not only an assessment of the risk of the presence of high-grade dysplasia or invasive cancer within a cyst, but also the presence of symptoms secondary to the cyst, the age of the patient, and patient co-morbidities. For this study, cysts were considered as appropriately resected if they were found on histopathologic examination to be SPNs, MCNs, or IPMNs that had high-grade dysplasia or were associated with adjacent invasive adenocarcinoma.

### Cyst fluid collection

Pancreatic cyst fluid was collected at the time of EUS or from the resected specimen in the surgical pathology laboratory. In the majority of instances, the results presented are from the analyses of cyst fluid aspirated post-operatively from the resected specimen. In 24 cases, paired samples, one of which was obtained at the time of EUS and the second at the time of

surgical resection, were available from the same patient for mutation assessment, loss of heterozygosity (LOH) and aneuploidy analyses.

### DNA purification

DNA was purified from cyst fluid (0.25 to 1.0 mL) by adding 3 ml of RLTM buffer (Qiagen) and then binding to an AllPrep column (Qiagen) according to the manufacturer's instructions. DNA amounts were assessed by qPCR using the primers conditions described in Rago et al.<sup>16</sup> A subset of the cysts reported here had been previously analyzed for *KRAS* and *GNAS* mutations.<sup>13</sup>

### Assessment of mutations

Because of their tremendous throughput, massively parallel sequencing instruments are highly cost effective instruments for DNA mutation analysis. However, sample preparation and sequencing steps introduce artifactual mutations into analyses at a low but significant frequency of approximately  $9.1 \times 10^{-6}$ .<sup>17</sup> In most clinical settings, this is irrelevant, as mutations of interest occur in a high fraction of alleles: 50% for germline mutations and greater than 10% for most tumors. Mutations in cysts occur much less frequently, often at 1% or less, and conventional sequencing cannot detect such low frequency mutations in a confident fashion. To better discriminate genuine mutations from artifactual sequencing variants introduced during these processes, we used Safe-SeqS, a technique which decreases the error rate from  $9.1 \times 10^{-6}$  to  $4.5 \times 10^{-7}$ , and allows detection of mutations present in as few as 0.01% of alleles, depending on sequencing depth and position of the mutation.<sup>17, 18</sup>

Safe-SeqS amplification primers were designed to amplify 60-bp to 262-bp segments each containing a region of interest. These regions of interest were derived from the following genes; *BRAF*, *CDKN2A*, *CTNNB1*, *GNAS*, *KRAS*, *NRAS*, *PIK3CA*, *RNF43*, *SMAD4*, *TP53*, *VHL* with primer sequences described in Supplemental Table 1. These primers were used to amplify DNA in 25- $\mu$ L multiplex PCRs as described in Kinde et al.<sup>17, 18</sup> For each sample, eight multiplex PCRs were performed, with each multiplex PCR containing 9 to 28 primer pairs. Reactions were purified with AMPure XP beads (Beckman Coulter) and eluted in 100  $\mu$ L of Buffer EB (Qiagen, cat. no. 19086). Five  $\mu$ L of purified PCR products were then amplified in a second round of PCR, as described in Kinde et al.<sup>17, 18</sup> The PCR products were purified with AMPure and used for sequencing on a MiSeq instrument.

High quality sequence reads were analyzed as previously described.<sup>17</sup> Briefly, we selected reads that contained high quality base calls in their first 14 cycles as assessed by the quality scores generated by the sequencing instrument, which indicate the probability that an individual base call was made in error.<sup>19</sup> Reads in which each of these 14 cycles had a quality score  $\geq 15$  were retained for further analysis. The template-specific portion of the reads that contained the sequence of an expected amplification primer was matched to the reference sequences using Bowtie (Bowtie 0.12.8). The unique identifier sequences (UIDs) that were incorporated as molecular barcodes into each template were used to group reads from a common template, as described in Kinde et al.<sup>17</sup> Artifactual mutations introduced during the sample preparation or sequencing steps were reduced by requiring that  $>90\%$  of reads sharing the same UID contained the identical mutation (a "supermutant"). Normal

peripheral blood DNA was used as a control to identify potential false positive mutations. Only supermutant frequencies in cysts that far exceeded supermutant frequencies in the control DNA samples (i.e.,  $> \text{mean} + 5$  standard deviations) were scored as mutations.

### LOH analysis

This was performed in a fashion similar to that described above for mutations, but different primer sets were used. The primer sets amplified genomic regions of ~120 bp that contained common single nucleotide polymorphisms (SNPs) that were within or closely surrounding (within 1 Mb) the tumor suppressor genes *CDKN2A*, *RNF43*, *SMAD4*, *TP53*, or *VHL*. Analogously to the mutation protocol, each DNA sample was used for six multiplex PCRs, each containing 11 to 32 primer pairs (Supplemental Table 2). The analysis was also carried out similarly, with the goal of identifying independent template molecules, defined by their UIDs, that were informative for the analyzed SNPs. The primer pairs used in this analysis were chosen from a large number of amplicons in the same region after extensive experimentation. The fraction of template molecules containing either allele in each of the chosen SNP amplicons was found to be 50%  $\pm$  2% (mean  $\pm$  1 SD) in the analyses of 20 samples of peripheral blood DNA. A sample was scored as having LOH if  $>80\%$  of the informative SNPs at one of the five loci assessed (*CDKN2A*, *RNF43*, *SMAD4*, *TP53*, or *VHL*) had allelic fractions lower or higher than 45% or 55%, respectively (i.e., lower or higher than 2.5 standard deviations from the mean).

### Assessment of aneuploidy

Aneuploidy across all non-acrocentric autosomal arms was assessed with a technology called FastSeqS.<sup>20</sup> With this technology, a single PCR is used to amplify ~20,000 loci scattered throughout the genome. After massively parallel sequencing, the fractional representation of each chromosomal arm can be determined by summing the reads that correspond to the loci in each arm. The PCR conditions and analytic methods used for these assays are described in Kinde et al.<sup>20</sup>

### Identification of individual molecular features of interest and composite molecular markers

The MOCA algorithm for analysis of multi-parametric data sets has previously been described.<sup>21, 22</sup> Briefly, MOCA first selects features of interest, and then selects collections of features using Boolean logic operations. The composite features (termed "composite markers") are compared to the phenotypes under consideration (i.e., cyst type or need for surgery), and the corresponding Fisher's exact two-tailed P-value, sensitivity, and specificity recorded. Leave-one-out cross validations are used to identify the composite molecular markers that perform best. As the algorithm progresses, an optimization strategy is implemented, resulting in algorithmic convergence on sets of composite markers with optimal performance for predicting the phenotype under consideration. For each cross-validation calculation, P-values are corrected for multiple testing using the Benjamini and Hochberg false discovery rate (FDR); composite markers were only considered if they had an FDR-corrected P-value of  $< 0.05$ . Furthermore, composite markers were only considered if they were selected in each of the cross validations for discriminating a particular type or grade of cyst.

MOCA was applied to an independent set of patients with pancreatic cysts, but without molecular data, to generate composite clinical markers.<sup>14</sup> In the current study, we tested these composite clinical markers on a set of 130 patients in whom molecular data was obtained as described above. None of the data from the patients in the current study were used to develop the original composite markers described in Masica et al.<sup>14</sup>

Finally, we used MOCA to identify composite markers that incorporated both molecular and clinical features (composite molecular/clinical markers) in the same way as described above. Sensitivity and specificity were used to quantify the performances of the composite molecular, clinical, and clinical/molecular markers, with the post-operative diagnoses based on histopathologic criteria.

## Results

### Basic patient and cyst characteristics

The cohort consisted of 130 patients (12 with SCAs, 10 with SPNs, 12 with MCNs, 96 with IPMNs). Of the IPMNs, 30 were main duct, 10 were mixed duct, 55 were branch duct types, and one was an intraductal tubulopapillary neoplasm (ITPN). The general demographics, symptoms, cyst features and pathological diagnoses of each of the 130 patients are presented in Table 1. Preoperative imaging was not available in 6 (5%) patients (2 IPMNs, 2 MCNs, 2 SPNs). The presumed pre-operative diagnosis, and the presence of high-risk or concerning features, is presented in Supplemental Table 3.

### Cyst fluid

The minimum amount of cyst fluid analyzed was 0.25 mL. The median DNA concentration was 4.9 ng/μl (range 0.05–270 ng/μl).

### Molecular features

Three types of molecular genetic tests were applied to each cyst. The first involved a search for subtle mutations (e.g. missense mutations or small insertions or deletions) of the genes known to be altered in pancreatic cysts. The most frequently mutated regions of six oncogenes *BRAF*, *CTNNB1*, *GNAS*, *KRAS*, *NRAS*, *PIK3CA*, and the great majority of the coding regions of five tumor suppressor genes (*CDK2NA*, *RNF43*, *SMAD4*, *TP53*, and *VHL*) were analyzed.

The second test involved a search for LOH of the same five tumor suppressor genes. We designed a massively parallel sequencing-based test to evaluate LOH events in a quantitative fashion, presented for the first time in this manuscript, and applied it to the DNA of cyst fluids.

Finally, it has been observed that aneuploidy is associated with malignant progression of neoplastic lesions of the pancreas.<sup>23</sup> We suspected that an evaluation of aneuploidy might help to identify high-risk cysts as well as to discriminate cyst types, and implemented a previously described PCR-based method for this purpose.<sup>20</sup>



Table 2 summarizes the mutational, LOH and aneuploidy analyses of the cysts; a detailed description of the aneuploidy data is presented in Supplemental Table 4. One or more intragenic mutations, LOH events, or aneuploid chromosomes was identified in 9 (75%) SCAs, 100 (100%) SPNs, 7 (58%) MCNs, 94 (98%) IPMNs. Overall, at least one molecular genetic alteration was detected in 92% of the cyst fluid samples.

There were distinct mutational profiles associated with each type of cyst. The *VHL* gene was mutated in the cyst fluid of five (42%) of the 12 SCAs; intragenic *VHL* mutations were not found in any of the other type of pancreatic cyst. In addition, 7 (64%) of the SCAs had LOH of chromosome 3 at the *VHL* gene locus, with 1 (8%) of the SCAs revealing aneuploidy at chromosome 3p. This finding is consistent with the rationale for evaluating LOH and aneuploidy described above. Overall, 8 (67%) of the SCAs either had a mutation in *VHL*, LOH of chromosome 3 or aneuploidy of chromosome 3p. There were four SCAs in which no abnormality was detected. Two of these four cases had matching surgical samples which were analyzed and no mutation in *VHL*, LOH or aneuploidy was identified.

The cyst fluid from all ten (100%) SPNs harbored a mutation in *CTNNB1*. One of the ten (10%) also harbored a mutation in *TP53*. Six (60%) patients with SPN had aneuploidy, involving chromosomes 11p (n=2) or 16p (n=6).

Of the 12 cyst fluid samples from MCNs analyzed, *KRAS* was the most commonly mutated gene, with activating mutations found in six samples (50%). *RNF43* was mutated in one MCN cyst fluid sample (8%) and this sample also harbored a *KRAS* mutation. LOH of chromosome 18 at the *SMAD4* gene locus was identified in one of the 12 MCN cyst fluid samples. Two (17%) MCNs had aneuploidy, involving chromosome 5p (n=1) or chromosome 16p (n=1).

Cyst fluid samples from 94 of the 96 (98%) IPMNs contained at least one mutation, LOH or aneuploidy. *KRAS* was the most prevalent altered gene in the cyst fluid samples from IPMNs (78%), with single base substitutions occurring at codons 12, 13 or 61. *GNAS* mutations were identified in 56 (58%) IPMNs, and *GNAS* mutations were not present in any other cyst type. All *GNAS* mutations were single base substitutions at codon 201, resulting in substitution of an arginine with histidine (R201H), cysteine (R201C), or serine (R201S). *GNAS* mutations occurred in 9 (82%) of the IPMNs with intestinal type histology. They were found at a lower prevalence in the gastric (n=39 (61%)) and pancreaticobiliary (n=3 (38%)) type IPMNs, while none of the 3 oncocytic type IPMNs harbored a *GNAS* mutation. Overall, 86 (91%) of the IPMNs had a mutation in *KRAS* or *GNAS*, and 45 (47%) had a mutation in both genes. Mutations in *RNF43*, *TP53*, *SMAD4* and *CDKN2A*, also occurred in IPMNs, but less commonly. Six (6%) IPMNs had a mutation in *CTNNB1*. In contrast to SPNs, where *CTNNB1* mutations occurred in isolation, all of the six IPMNs with a *CTNNB1* mutation also had another characteristic mutation, LOH or aneuploidy. A summary of the mutational, LOH and aneuploidy analyses based on IPMN type is presented in Supplemental Table 5.

### Comparison between cyst fluids obtained during EUS versus those obtained at surgery

For 24 of the 130 cyst fluid specimens collected at the time of surgery, matching cyst fluid collected preoperatively by EUS was also available for analysis. This group of matching samples included cyst fluid collected from 17 IPMNs, 3 MCNs, 2 SCAs, 1 SPN and 1 ITPN. The number of genetic alterations detected in the cyst fluid samples collected at EUS (21 of 24, 87.5%) and at the time of surgery (20 of 24, 83.3%) was similar. Of the 1266 possible genetic alteration results (mutations, LOH, or aneuploidy) for these 24 matched cyst fluid samples, 1198 of the 1266 (94.6%) were concordant.

### Cyst classification via composite molecular markers

We then used MOCA to identify composite molecular markers based on the individual features described above. SCAs were identified with 100% sensitivity and 91% specificity by the absence of a *KRAS*, *GNAS*, *RNF43* mutation, or by the absence of aneuploidy in chromosome 5p or 8p (Table 3). The presence of a *VHL* mutation has previously been shown to be predominantly associated with SCAs.<sup>12, 13</sup> In addition, on examination of the new molecular data presented here, the presence of LOH in chromosome 3 in the absence of LOH in chromosomes 9, 17 or 18, was exclusively identified in patients with SCA. These two features were therefore added to form the SCA composite molecular marker in expectation that this would be a useful feature for future assessments of cyst type; these "manually" added genetic features had no effect on the performance of the composite marker.

SPNs were identified with 100% sensitivity and 100% specificity by the presence of a *CTNNB1* mutation and the absence of *KRAS*, *GNAS*, or *RNF43* mutations or chromosome 18 LOH (Table 3).

MCNs were identified with 100% sensitivity and 75% specificity by the *absence* of *CTNNB1* or *GNAS* mutations, chromosome 3 LOH, or aneuploidy in chromosome 1q or 22q.

Finally, IPMNs were identified with 76% sensitivity and 97% specificity by the presence of a mutation in *GNAS*, *RNF43*, LOH in chromosome 9, or aneuploidy in chromosome 1q or 8p (Table 3).

### Cyst classification via composite clinical markers

In an independent set of cysts, we separately identified composite clinical markers for each cyst type.<sup>14</sup> The new, 130-patient cohort gave us the opportunity to validate these composite clinical markers in an independent cohort. When applied to the 130-patient cohort, the composite clinical markers had high sensitivity for SCAs, SPNs, and MCNs (100%, 89%, and 90%, respectively), and modest sensitivity for IPMNs (75%; Supplemental Table 6). The specificities of the composite clinical markers ranged from 71% to 88% for SCAs, SPNS, MCNs and IPMNs. The sensitivities and specificities estimated by cross-validation in Masica et al<sup>14</sup> (Supplemental Table 6) were in general similar to those estimated by analysis of this new 130-patient cohort. Though there were some differences—such as a higher sensitivity for IPMNs in the Masica study—all the markers were highly significant when



applied to the 130-patient study. Although some of the sensitivities and specificities were outside the 95% confidence intervals of the Masica et al study (see Supplemental Table 6), this is not unexpected given that there is almost an order-of-magnitude difference in sample sizes between the cohorts.

### **Cyst classification via composite molecular and clinical markers**

Intuitively, one would expect that the combination of two different sets of biomarkers could, at least in certain circumstances, provide higher accuracy than either alone. For this purpose, we used MOCA to identify a new composite marker set, called "Composite molecular/clinical markers" that included the composite molecular markers noted above plus the clinical or radiologic features identified by Masica et al as useful for cyst classification.<sup>14</sup>

Because the composite molecular marker was so sensitive for identifying SCAs, sensitivity was not increased by adding clinical or radiologic features (Table 3). However, the *absence* of main pancreatic duct (MPD) dilation, communication with the MPD, or abdominal pain, increased the specificity for identifying SCA from 91% to 98% without compromising the 100% sensitivity.

The sensitivity and specificity of the composite molecular marker for identifying SPNs were both 100%. The addition of the clinical or radiologic features to the molecular markers decreased the sensitivity by 11% and decreased the specificity by 8%, for identifying this cyst type.

MCNs were similar to SCAs in that the composite molecular markers alone had perfect sensitivity (100%) but imperfect specificity (75%). The *presence* of age <75 years, and the *absence* of all three clinical or radiologic features (male gender, multiple cysts, communication with the MPD) increased the specificity to 97%, with a slight decrease in sensitivity to 90%.

In contrast, an increase in sensitivity was realized when any of the following features (age 85 years, abdominal pain, MPD dilation or communication with the MPD) were added to the composite molecular marker for IPMNs. This composite molecular/clinical marker panel increased the sensitivity for having an IPMN from 76% (composite molecular marker alone) to 94%, while slightly decreasing specificity (from 97% to 84%; Table 3).

### **Identification of cysts that require surgical resection**

From a practical perspective, the most important question in the management of cyst patients is the decision to perform surgical resection. We wished to determine the accuracy of the various composite markers described here and in Masica et al<sup>14</sup> for determining this need for resection. As noted in Materials and Methods, we considered IPMNs characterized post-operatively as containing high-grade dysplasia or associated with invasive carcinoma to have been the most appropriate IPMNs for surgical excision. We also considered cysts that were histopathologically (i.e., post-operatively) diagnosed as SPNs or MCNs to have required surgical excision. SCAs were considered to have not required surgical excision. IPMNs with low-grade or intermediate-grade dysplasia were also considered in retrospect to have not required surgical resection at that time.

### Predicting which cysts need surgery using the Composite Molecular Markers

To further characterize the molecular characteristics of IPMNs with high-grade dysplasia or associated with invasive carcinoma we analyzed the results from the 96 patients with a resected IPMN. The features of IPMNs that best predicted high-grade dysplasia or associated with invasive carcinoma are presented in Supplemental Table 7. These were the presence of a mutation in *SMAD4*, chromosome 17q LOH (the region containing *RNF43*), or aneuploidy in chromosome 5p, 8p, 13q or 18q (Fig. 1). We manually added a mutation in *TP53*, or chromosome 17p LOH (the region containing *TP53*), as these features have previously been described to occur in IPMNs with high-grade dysplasia or associated with invasive carcinoma.<sup>24</sup> A composite marker for IPMNs with high-grade dysplasia and/or associated with an invasive carcinoma based on these features, together with the composite markers for SCAs, MCNs, and SPNs described in Table 3, were then used to analyze the entire set of 130 patients. These composite molecular markers correctly identified patients requiring surgery with a sensitivity of 75% and a specificity of 92% (Table 4).

### Predicting which cysts need surgery using Composite Clinical Markers, with or without Composite Molecular Markers

The composite clinical markers<sup>14</sup>, when applied alone to the set of 130 patients, were able to identify cysts that required surgery with a sensitivity of 77% and a specificity of 75%. This provides a validation of the clinical composite markers, which were predicted to have a sensitivity of 84% and a specificity of 81% based on the cross-validation among the separate, 1026 patients in the original study (Masica et al).<sup>14</sup> When both the clinical and molecular features were combined into a composite molecular/clinical marker to predict which cysts required surgery and applied to the 130 patients, the sensitivity increased to 89%, but at the expense of specificity, which fell to 69% (Table 4). The composite molecular markers provide information about the risk of high-grade dysplasia or an associated invasive cancer in IPMNs at the point in time the analysis is performed, and do not predict the risk of developing high-grade dysplasia or an associated invasive cancer in the future.

### Potential to avoid unnecessary surgery

All cysts included in this study underwent surgical resection, however those patients with SCA, or IPMNs with low- or intermediate-grade dysplasia in retrospect, may not have required surgery at this point in time. In this study, had the molecular analysis been performed prior to surgery, many unnecessary surgeries could have potentially been prevented. For example, the composite molecular/clinical marker correctly identified all 12 SCAs. If this had been realized prior to surgery, surgical resection would likely have been avoided in most. Similarly, many IPMNs were resected because of concern for the presence of high-grade dysplasia or an associated invasive cancer; however 62 of these IPMNs had only low- or intermediate-grade dysplasia. Fifty-six (90%) of these 62 patients would have been correctly identified as not needing surgery at the time of their evaluation using the composite molecular marker. Of the 74 patients with SCAs or IPMNs that in retrospect did not meet the histopathological criteria for surgical resection, the composite molecular

marker correctly identified 67 of these cases, thus potentially decreasing the number of unnecessary operations by 91%.

## Discussion

The results described above lead to several important conclusions. First, over half of the patients who underwent surgical resection were subsequently found to have a benign SCA, or an IPMN with low-, or intermediate-grade dysplasia, and could potentially have continued surveillance of their cyst rather than undergoing surgical resection. This highlights the difficulties of identifying those cysts that require surgery, versus those in whom surveillance is safe, using the modalities currently available to clinicians. These results are in accord with the numerous publications showing that current diagnostic criteria for managing cyst patients are inadequate.<sup>10, 11</sup>

Second, we show that the use of composite clinical or molecular markers could substantially increase diagnostic accuracy. When either the composite clinical marker or the composite molecular marker was used alone, the sensitivity for identifying cysts that required resection reached ~75%; in contrast, when used together, sensitivity increased to 92%. It is difficult to estimate the number of cyst patients who would have surgery, and therefore would not develop PDAC, if these markers were widely applied in patient management. Similarly, we hesitate to calculate how many needless surgeries might be avoided if these new markers were broadly applied. However, the data strongly suggest that the combination of clinical and molecular features will be more accurate for assessing cyst type and need for surgical resection than either alone.

The clinical and radiologic findings incorporated into our composite clinical marker are the result of decades of careful study by clinicians.<sup>3, 25</sup> What was added in Masica et al<sup>14</sup> was a rigorous and quantitative assessment of the most predictive features and combinations of features. Similarly, the molecular analysis we employed did not require the discovery of new genetic alterations present in pancreatic cysts. Guided by prior studies, we have developed assays employing massively parallel sequencing to robustly detect these genetic alterations, even when present in relatively low fractions of template molecules. We then used these data to identify the most predictive molecular features and combinations of molecular features in a rigorous fashion.

A balance between sensitivity and specificity nearly always must be made during the development of biomarkers for any disease. The current study was no exception. The most obvious example was in the determination of the need for surgery (Table 4). The composite molecular/clinical marker provided an excellent sensitivity (89%), considerably higher than either the composite molecular or composite clinical marker alone (75% or 77%, respectively). However, this increase in sensitivity with the composite molecular/clinical marker compromised specificity, reducing it from 92% with the composite molecular marker to 69%. Note that we purposefully designed these algorithms to reach maximum sensitivity, sacrificing specificity if necessary, as we considered it worse to "miss" a cyst that should be surgically excised than to unnecessarily perform surgery on a cyst that should have not been excised. However, this example illustrates that it is not always possible to

increase sensitivity without decreasing specificity, even with the aid of combinatorial approaches such as MOCA.

The summary data in Table 4, once validated independently, can be used in various ways, depending on the clinical situation. In a young and otherwise healthy patient, it would make sense to use the most sensitive method available to determine the need for surgery. The composite molecular/clinical marker might therefore be used to evaluate such a patient, as it is the most sensitive. In an elderly patient with significant comorbidities, however, it might make more sense to use the less sensitive, but more specific composite molecular marker to determine the need for surgery. Use of the composite molecular marker would largely avoid unnecessary surgery (specificity of 92%), while preserving a reasonable sensitivity for high-risk cysts (75%).

Our study has several limitations. The number of cyst fluid samples from some cyst types was relatively limited, thereby limiting confidence in our estimates of sensitivity and specificity (as indicated by the confidence intervals provided in all Tables). Another limitation is that our composite molecular marker was validated through cross-validation rather than through experimental validation of an independent cohort. Though cross-validation is statistically sound, it is not as reliable as the evaluation of a distinct cohort. It is important to note that this limitation does not apply to the composite clinical marker; the 130 patients evaluated here were distinct from the patients used to define the composite clinical markers.<sup>14</sup> It was therefore gratifying that the sensitivities and specificities estimated from the cross-validation in Masica et al. were similar to those found in the current study (Supplemental Table 4).

The possibility of false negative results is always a concern with any sequencing technique. This study was designed to detect all known mutations in oncogenes and most mutations in tumor suppressor genes that occur in cysts, based on genome wide sequencing.<sup>12</sup> All missense mutations and small insertions or deletions present at allele frequencies of greater than 1% are easily detectable at the sequencing depth employed in this study.<sup>17</sup> Moreover, the approach used in this study yields not just the presence or absence of a mutation; it reveals the precise fractional representation of the mutation in the DNA sample. Another source of false negative results could occur in tumor suppressor genes. Large deletions or insertions, as well as translocations, will not be detected upon sequencing. However, such changes are often associated with LOH or copy number changes, and these would be identified by the other assays employed in this study.

Obtaining large volumes of cyst fluid for analysis can be problematic and could potentially affect the ability to detect alterations in the molecular markers described in this study. However unlike analysis of cyst fluid CEA, which requires 0.5ml to 1ml of cyst fluid in most centers, we specifically chose to use methods that can be used on very small amounts of DNA and do not require library preparation. Each of the three methods is based on direct polymerase chain reaction, and 10 ng of DNA is adequate for all of the tests combined. Using 0.25 mL of either EUS or surgically obtained fluids has nearly always yielded sufficient DNA.

This was a retrospective study. While this has the advantage that all patients underwent surgical resection and therefore had defined pathology, it is possible that these markers will not perform as well in the general population of cyst patients. It was comforting in this respect that our analysis of paired fluids obtained from EUS-guided cyst aspiration and subsequent surgery revealed very similar molecular genetic alterations. In the future, the optimum study design will incorporate examination of cyst fluids taken at routine EUS sessions over time, then comparing the results to those obtained at surgery. In addition, the optimum role of the composite molecular markers, and which patients will benefit from their use, should be addressed in these studies. It will be enticing in such research studies to use the composite molecular and clinical markers described in this work to help guide the decision about surgery. But because of the limitations of our study described above, any such guidance should be performed only in a research study. Use of our composite molecular or clinical markers in common practice, outside of a research study, is not yet warranted.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Authors

Simeon Springer<sup>1,2,\*</sup>, Yuxuan Wang<sup>1,2,\*</sup>, Marco Dal Molin<sup>2,3,\*</sup>, David L. Masica<sup>2,4,10,\*</sup>, Yuchen Jiao<sup>1,2</sup>, Isaac Kinde<sup>1,2</sup>, Amanda Blackford<sup>5</sup>, Siva P. Raman<sup>6</sup>, Christopher L. Wolfgang<sup>2,7,8</sup>, Tyler Tomita<sup>4,10</sup>, Noushin Niknafs<sup>4,10</sup>, Christopher Douville<sup>4,10</sup>, Janine Ptak<sup>1,2</sup>, Lisa Dobbyn<sup>1,2</sup>, Peter J. Allen<sup>11</sup>, David S. Klimstra<sup>12</sup>, Mark A. Schattner<sup>13</sup>, C. Max Schmidt<sup>14</sup>, Michele Yip-Schneider<sup>15</sup>, Oscar W. Cummings<sup>15</sup>, Randall E. Brand<sup>16</sup>, Herbert J. Zeh<sup>17</sup>, Aatur D. Singhi<sup>18</sup>, Aldo Scarpa<sup>19,20</sup>, Roberto Salvia<sup>21</sup>, Giuseppe Malleo<sup>21</sup>, Giuseppe Zamboni<sup>20,22</sup>, Massimo Falconi<sup>23</sup>, Jin-Young Jang<sup>24</sup>, Sun-Whe Kim<sup>24</sup>, Wooil Kwon<sup>24</sup>, Seung-Mo Hong<sup>25</sup>, Ki-Byung Song<sup>26</sup>, Song Cheol Kim<sup>26</sup>, Niall Swan<sup>27</sup>, Jean Murphy<sup>27</sup>, Justin Geoghegan<sup>28</sup>, William Brugge<sup>29</sup>, Carlos Fernandez-Del Castillo<sup>30</sup>, Mari Mino-Kenudson<sup>31</sup>, Richard Schulick<sup>32</sup>, Barish H. Edil<sup>32</sup>, Volkan Adsay<sup>33</sup>, Jorge Paulino<sup>34</sup>, Jeanin van Hooft<sup>35</sup>, Shinichi Yachida<sup>36</sup>, Satoshi Nara<sup>36</sup>, Nobuyoshi Hiraoka<sup>36</sup>, Kenji Yamao<sup>37</sup>, Susuma Hijioka<sup>37</sup>, Schalk van der Merwe<sup>38</sup>, Michael Goggins<sup>2,8,9</sup>, Marcia Irene Canto<sup>9</sup>, Nita Ahuja<sup>7</sup>, Kenzo Hirose<sup>7</sup>, Martin Makary<sup>7</sup>, Matthew J. Weiss<sup>7</sup>, John Cameron<sup>7</sup>, Meredith Pittman<sup>2,3</sup>, James R. Eshleman<sup>1,2</sup>, Luis A. Diaz Jr.<sup>1,2,8</sup>, Nickolas Papadopoulos<sup>1,2</sup>, Kenneth W. Kinzler<sup>1,2</sup>, Rachel Karchin<sup>2,4,8,10</sup>, Ralph H. Hruban<sup>1,2,3,8</sup>, Bert Vogelstein<sup>1,2</sup>, and Anne Marie Lennon<sup>2,7,9</sup>

## Affiliations

<sup>1</sup>The Ludwig Center and Howard Hughes Medical Institute at the Sidney Kimmel Cancer Center, The Johns Hopkins University, Baltimore, MD, USA <sup>2</sup>The Sol Goldman Pancreatic Cancer Research Center, The Johns Hopkins University, Baltimore, MD, USA <sup>3</sup>Department of Pathology, The Johns Hopkins University, Baltimore, MD, USA <sup>4</sup>Department of Biomedical Engineering, The Johns Hopkins University, Baltimore, MD, USA <sup>5</sup>Department of Biostatistics and Bioinformatics,

The Johns Hopkins University, Baltimore, MD, USA <sup>6</sup>Department of Radiology, The Johns Hopkins University, Baltimore, MD, USA <sup>7</sup>Department of Surgery, The Johns Hopkins University, Baltimore, MD, USA <sup>8</sup>Department of Oncology, The Johns Hopkins University, Baltimore, MD, USA <sup>9</sup>Department of Medicine, The Johns Hopkins University, Baltimore, MD, USA <sup>10</sup>The Johns Hopkins Medical Institutions and the Institute for Computational Medicine, The Johns Hopkins University, Baltimore, MD, USA <sup>11</sup>Department of Surgery, Memorial Sloan-Kettering Cancer Center <sup>12</sup>Department of Pathology, Memorial Sloan-Kettering Cancer Center <sup>13</sup>Department of Gastroenterology, Memorial Sloan-Kettering Cancer Center <sup>14</sup>Department of Surgery, University of Indiana <sup>15</sup>Department of Pathology, University of Indiana <sup>16</sup>Department of Medicine, University of Pittsburgh <sup>17</sup>Department of Surgery, University of Pittsburgh <sup>18</sup>Department of Pathology, University of Pittsburgh <sup>19</sup>ARC-Net Research Centre and Department of Pathology and Diagnostics, University and Hospital Trust of Verona, Italy <sup>20</sup>Department of Pathology, General Surgery B, University and Hospital Trust of Verona, Italy <sup>21</sup>Department of Surgery, University and Hospital Trust of Verona, Negrar, Italy <sup>22</sup>Department of Pathology, Ospedale Sacro Cuore-Don Calabria, Negrar, Italy <sup>23</sup>Division of Pancreatic Surgery, Department of Surgery, IRCCS San Raffaele Scientific Institute, Milan, Italy <sup>24</sup>Department of Surgery and Cancer Research Institute, Seoul National University College of Medicine, Seoul, Korea <sup>25</sup>Department of Pathology, Asan Medical Center, University of Ulsan College of Medicine, Seoul, Korea <sup>26</sup>Department of Hepatobiliary and Pancreas Surgery, Asan Medical Center, University of Ulsan College of Medicine, Seoul, Korea <sup>27</sup>Department of Histopathology, St. Vincent's University Hospital, Dublin, Ireland <sup>28</sup>Department of Surgery, St. Vincent's University Hospital, Dublin, Ireland <sup>29</sup>Department of Gastroenterology, Massachusetts General Hospital <sup>30</sup>Department of Surgery, Massachusetts General Hospital <sup>31</sup>Department of Histopathology, Massachusetts General Hospital <sup>32</sup>Department of Surgery, University of Colorado <sup>33</sup>Department of Pathology, Emory University <sup>34</sup>Department of Pathology, Centro Hepatobiliopancreático e Transplantação – Hospital Curry Cabral, Lisbon, Portugal <sup>35</sup>Department of Gastroenterology and Hepatology, Amsterdam Medical Center, Netherlands <sup>36</sup>Department of Hepatobiliary and Pancreatic Surgery, Pathology and Cancer Genomics, National Cancer Center Hospital and National Cancer Center Research Institute, Tokyo, Japan <sup>37</sup>Department of Gastroenterology, Aichi Cancer Center Hospital, Nagoya, Japan <sup>38</sup>Department of Hepatology, University Hospitals KU Leuven, Belgium

## Acknowledgments

The authors would like to acknowledge Natalie Silliman and Joy Schaefer for their expert technical assistance.

### Grant Support:

Supported by The Lustgarten Foundation for Pancreatic Cancer Research, The Sol Goldman Center for Pancreatic Cancer Research, The Virginia and D.K. Ludwig Fund for Cancer Research, Susan Wojcicki and Dennis Troper,



The Michael Rolfe Foundation, National Institutes of Health grant P50 CA62924, Associazione Italiana Ricerca Cancro (12182) Italian Ministry of Research (FIRB RBAP10AHJB) and Health (FIMP J33G13000210001).

Drs. Diaz, Papadopoulos, Kinzler, and Vogelstein are founders of Personal Genome Diagnostics, Inc. and PapGene, Inc. Drs. Vogelstein and Kinzler are also members of the Scientific Advisory Board of Syxmex-Inostics. These companies and others have licensed technologies from Johns Hopkins, of which Drs. Diaz, Hruban, Goggins, Papadopoulos, Kinzler and Vogelstein are inventors and receive royalties from these licenses.

## Abbreviations

<b>CEA</b>	carcinoembryonic antigen
<b>Chr</b>	chromosome
<b>CT</b>	computer axial tomography
<b>EUS</b>	endoscopic ultrasound
<b>IPMN</b>	intraductal papillary mucinous neoplasm
<b>ITPN</b>	intraductal tubulopapillary neoplasm
<b>LOH</b>	loss of heterozygosity
<b>MCN</b>	mucinous cystic neoplasm
<b>MOCA</b>	Multivariate organization of a combinatorial alterations
<b>MPD</b>	main pancreatic duct
<b>MRI</b>	magnetic resonance imaging
<b>SCA</b>	serous cystadenoma
<b>SPN</b>	solid-pseudopapillary neoplasm
<b>UID</b>	unique identifier sequences

## References

1. Laffan TA, Horton KM, Klein AP, et al. Prevalence of unsuspected pancreatic cysts on MDCT. *AJR Am J Roentgenol.* 2008; 191:802–807. [PubMed: 18716113]
2. Lee KS, Sekhar A, Rofsky NM, et al. Prevalence of incidental pancreatic cysts in the adult population on MR imaging. *Am J Gastroenterol.* 2010; 105:2079–2084. [PubMed: 20354507]
3. Lennon AM, Wolfgang C. Cystic neoplasms of the pancreas. *J Gastrointest Surg.* 2013; 17:645–653. [PubMed: 23340991]
4. Bosman, FT. World Health Organization. WHO classification of tumours of the digestive system. Lyon: International Agency for Research on Cancer; 2010. International Agency for Research on Cancer.
5. Law JK, Ahmed A, Singh VK, et al. A systematic review of solid-pseudopapillary neoplasms: are these rare lesions? *Pancreas.* 2014; 43:331–337. [PubMed: 24622060]
6. Yamao K, Yanagisawa A, Takahashi K, et al. Clinicopathological features and prognosis of mucinous cystic neoplasm with ovarian-type stroma: a multiinstitutional study of the Japan pancreas society. *Pancreas.* 2011; 40:67–71. [PubMed: 20924309]
7. Tanaka M, Fernandez-Del Castillo C, Adsay V, et al. International consensus guidelines 2012 for the management of IPMN and MCN of the pancreas. *Pancreatol.* 2012; 12:183–197. [PubMed: 22687371]

8. Lennon AM, Wolfgang CL, Canto MI, et al. The Early Detection of Pancreatic Cancer: What Will It Take to Diagnose and Treat Curable Pancreatic Neoplasia? *Cancer Res.* 2014; 74:3381–3389. [PubMed: 24924775]
9. Sohn TA, Yeo CJ, Cameron JL, et al. Intraductal papillary mucinous neoplasms of the pancreas: an updated experience. *Ann Surg.* 2004; 239:788–797. discussion 797–9. [PubMed: 15166958]
10. Sahara K, Mino-Kenudson M, Brugge W, et al. Branch duct intraductal papillary mucinous neoplasms: does cyst size change the tip of the scale? A critical analysis of the revised international consensus guidelines in a large single-institutional series. *Ann Surg.* 2013; 258:466–475. [PubMed: 24022439]
11. Valsangkar NP, Morales-Oyarvide V, Thayer SP, et al. 851 resected cystic tumors of the pancreas: a 33-year experience at the Massachusetts General Hospital. *Surgery.* 2012; 152:S4–S12. [PubMed: 22770958]
12. Wu J, Jiao Y, Dal Molin M, et al. Whole-exome sequencing of neoplastic cysts of the pancreas reveals recurrent mutations in components of ubiquitin-dependent pathways. *Proc Natl Acad Sci U S A.* 2011; 108:21188–21193. [PubMed: 22158988]
13. Wu J, Matthaei H, Maitra A, et al. Recurrent GNAS mutations define an unexpected pathway for pancreatic cyst development. *Sci Transl Med.* 2011; 3:92ra66.
14. Masica DL, Dal Molin M, Blackford A, et al. Systematic analysis of pancreatic cyst classification using clinical features from a large retrospective cohort. *Clin Gastroenterol Hepatol.* 2015 Submitted for review.
15. Bosman FT, Carneiro Ft, Hruban RH, et al. WHO classification of tumours of the digestive system: World Health Organization. 2010
16. Rago C, Huso DL, Diehl F, et al. Serial assessment of human tumor burdens in mice by the analysis of circulating DNA. *Cancer Res.* 2007; 67:9364–9370. [PubMed: 17909045]
17. Kinde I, Wu J, Papadopoulos N, et al. Detection and quantification of rare mutations with massively parallel sequencing. *Proc Natl Acad Sci U S A.* 2011; 108:9530–9535. [PubMed: 21586637]
18. Kinde I, Bettgowda C, Wang Y, et al. Evaluation of DNA from the Papanicolaou test to detect ovarian and endometrial cancers. *Sci Transl Med.* 2013; 5:167ra4.
19. Ewing B, Green P. Base-calling of automated sequencer traces using phred. II. Error probabilities. *Genome Res.* 1998; 8:186–194. [PubMed: 9521922]
20. Kinde I, Papadopoulos N, Kinzler KW, et al. FAST-SeqS: a simple and efficient method for the detection of aneuploidy by massively parallel sequencing. *PLoS One.* 2012; 7:e41162. [PubMed: 22815955]
21. Masica DL, Karchin R. Correlation of somatic mutation and expression identifies genes important in human glioblastoma progression and survival. *Cancer Res.* 2011; 71:4550–4561. [PubMed: 21555372]
22. Masica DL, Karchin R. Collections of simultaneously altered genes as biomarkers of cancer cell drug response. *Cancer Res.* 2013; 73:1699–1708. [PubMed: 23338612]
23. Bocsi J, Bercei C, Balazs G, et al. Flow cytometric studies on the relationship between DNA content and clinicopathologic features of pancreatic cancers. *Anticancer Res.* 1998; 18:1839–1843. [PubMed: 9673413]
24. Kanda M, Sadakari Y, Borges M, et al. Mutant TP53 in Duodenal Samples of Pancreatic Juice from Patients with Pancreatic Cancer or High-Grade Dysplasia. *Clin Gastroenterol Hepatol.* 2012
25. Lennon AM, Manos LL, Hruban RH, et al. Role of a Multidisciplinary Clinic in the Management of Patients with Pancreatic Cysts: A Single-Center Cohort Study. *Ann Surg Oncol.* 2014

Cyst and patient characteristics

Table 1

	All Samples N = 130	IPMN* N = 96	MCN N = 12	SCA N = 12	SPN N = 10
<b>Sex (n=130<sup>^</sup>)</b>					
Female - no. (%)	83 (64)	52 (54)	12 (100)	9 (75)	10 (100)
<b>Race (n=130<sup>^</sup>)</b>					
African American - no. (%)	7 (5)	4 (4)	1 (8)	1 (8)	1 (10)
White - no. (%)	114 (88)	87 (91)	10 (83)	8 (67)	9 (90)
Other - no. (%)	9 (7)	5 (5)	1 (8)	3 (25)	0 (0)
<b>Age at surgery (n=130<sup>^</sup>)</b>					
Years - mean (SD)	62.3 (17)	69.2 (10)	45.5 (14)	54.9 (15)	24.1 (4)
<b>Symptoms (n=130<sup>^</sup>)</b>					
Abdominal pain - no. (%)	23 (18)	21 (22)	2 (17)	0 (0)	0 (0)
Pancreatitis - no. (%)	16 (12)	14 (15)	2 (17)	0 (0)	0 (0)
Jaundice - no. (%)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Weight loss - no. (%)	6 (5)	5 (5)	1 (8)	0 (0)	0 (0)
Diabetes - no. (%)	25 (19)	21 (22)	1 (8)	2 (17)	1 (10)
<b>Cyst size (n=130<sup>^</sup>)</b>					
cm - median (IQR)	3.4 (2)	3 (2)	5.2 (4)	4 (2)	4 (2)
<b>Cyst location (n=130<sup>^</sup>)<sup>+</sup></b>					
Head or Uncinate - no. (%)	56 (45)	51 (55)	0 (0)	1 (8)	4 (40)
Neck - no. (%)	16 (13)	13 (14)	0 (0)	2 (17)	1 (10)
Body or Tail - no. (%)	65 (52)	42 (45)	12 (100)	9 (75)	5 (50)
<b>Multiple cysts (n=124<sup>^</sup>)</b>					
Yes - no. (%)	31 (25)	29 (31)	0 (0)	1 (8)	1 (11)
<b>Communication with MPD (n=107<sup>^</sup>)</b>					
Yes - no. (%)	39 (36)	38 (46)	1 (14)	0 (0)	0 (0)
<b>Mural Nodule (n=123<sup>^</sup>)</b>					
Yes - no. (%)	32 (26)	19 (20)	3 (30)	2 (17)	8 (100)

	All Samples N = 130	IPMN* N = 96	MCN N = 12	SCA N = 12	SPN N = 10
<b>CEA &gt; 192 ng/mL (n=51<sup>^</sup>)</b>					
Yes - no. (%)	30 (59)	24 (60)	6 (86)	0 (0)	0 (0)
<b>IPMN Histotype (n=95<sup>^</sup>)</b>					
Gastric - no. (%)	-	64 (67)	-	-	-
Intestinal - no. (%)	-	11 (11)	-	-	-
Pancreatobiliary - no. (%)	-	9 (9)	-	-	-
Oncocytic - no. (%)	-	3 (3)	-	-	-
Mixed - no. (%) <sup>o</sup>	-	8 (8)	-	-	-
<b>Grade of Dysplasia/Invasive Cancer in IPMNs, and MCNs (n=108<sup>^</sup>)</b>					
Low - no. (%)	25 (23)	15 (16)	10 (83)	-	-
Intermediate - no. (%)	49 (45.5)	47 (49)	2 (17)	-	-
High - no. (%)	22 (20.5)	22 (23)	0 (0)	-	-
Invasive Cancer - no. (%)	12 (11)	12 (12)	0 (0)	-	-

\* Includes one ITPN.

<sup>^</sup> Indicates the number of patients in whom data on this variable was available.

<sup>+</sup> All sites where a pancreatic cyst was located were documented. Since some of the cysts extended to more than one site, this resulted in a number of locations that was greater than the number of cysts.

<sup>o</sup> Indicates more than one histologic subtypes in the same lesion.

SD = standard deviation. IQR = Interquartile range.

	<b>IPMN*</b> <b>N = 96</b>	<b>MCN</b> <b>N = 12</b>	<b>SCA</b> <b>N = 12</b>	<b>SPN</b> <b>N = 10</b>
<i>KRAS</i> - no. (%)	75 (78)	6 (50)	0 (0)	0 (0)
<i>GNAS</i> - no. (%)	56 (58)	0 (0)	0 (0)	0 (0)
<i>RNF43</i> - no. (%)	36 (38)	1 (8)	0 (0)	0 (0)
<i>CDKN2A</i> - no. (%)	3 (3)	0 (0)	0 (0)	0 (0)
<i>CTNNB1</i> - no. (%)	6 (6)	0 (0)	0 (0)	10 (100)
<i>SMAD4</i> - no. (%)	5 (5)	0 (0)	0 (0)	0 (0)
<i>TP53</i> - no. (%)	9 (9)	0 (0)	0 (0)	1 (10)
<i>VHL</i> - no. (%)	0 (0)	0 (0)	5 (42)	0 (0)
<i>BRAF</i> - no. (%)	1 (1)	0 (0)	0 (0)	0 (0)
<i>NRAS</i> - no. (%)	0 (0)	0 (0)	0 (0)	0 (0)
<i>PIK3CA</i> - no. (%)	0 (0)	0 (0)	0 (0)	2 (20)
LOH chr3 ( <i>VHL</i> ) - no. (%)	4 (4)	0 (0)	7 (64)	0 (0)
LOH chr9 ( <i>CDKN2A</i> ) - no. (%)	8 (8)	0 (0)	0 (0)	0 (0)
LOH chr17 ( <i>RNF43</i> ) - no. (%)	11 (11)	0 (0)	1 (9)	0 (0)
LOH chr17 ( <i>TP53</i> ) - no. (%)	5 (5)	0 (0)	0 (0)	0 (0)
LOH chr18 ( <i>SMAD4</i> ) - no. (%)	10 (10)	1 (8)	0 (0)	0 (0)
Aneuploidy <sup>^</sup>	48 (50)	2 (17)	6 (50)	6 (60)

\* Includes one ITPN patient.

<sup>^</sup> Aneuploidy of at least one chromosome observed. Details are provided in Supplemental Table 3.

LOH = Loss of heterozygosity. chr = chromosome.

Table 3

## Identification of Pancreatic Cyst Type

Type of Cyst	Composite Molecular Markers				Composite Molecular AND Clinical <sup>1</sup> Markers			
	Any of these Present		Any of these Absent		Any of these Present		Any of these Absent	
	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
<b>SCA</b>	<i>VHL</i> <sup>#</sup> chr3 LOH <sup>*,#</sup>	100% (74–100%)	<i>KRAS</i> <i>GNAS</i> <i>RNF43</i> chr5p aneu chr8p aneu	91% (84–95%)	Age <25 years	Abdominal pain Communication with MPD MPD dilation	100% (74–100%)	98% (94–99.8%)
<b>SPN</b>	<i>CTNNB1</i>	100% (69–100%)	<i>KRAS</i> <i>GNAS</i> <i>RNF43</i> chr18 LOH	100% (97–100%)	Age <55 years	Jaundice Multiple cysts Weight loss	89% (52–99.7%)	92% (85–96%)
<b>MCN</b>	None	100% (74–100%)	<i>CTNNB1</i> <i>GNAS</i> chr3 LOH chr1q aneu chr22q aneu	75% (66–82%)	Age <75 years	Sex (male) Communication with MPD Multiple cysts	90% (56–99.8%)	97% (91–99%)
<b>IPMIN</b>	<i>GNAS</i> <i>RNF43</i> <sup>^</sup> chr9 LOH chr1q aneu chr8p aneu	76% (66–84%)	None	97% (85–99.9%)	Age >85 years Communication with MPD MPD dilation Abdominal pain	None	94% (86–98%)	84% (64–95%)

<sup>#</sup>These features were not identified by MOCA but were manually added.

<sup>\*</sup> LOH of chromosome 3 but no LOH of chromosomes 9, 17 or 18.

<sup>^</sup> Mutations in *RNF43* have also been reported in MCNs, and in a larger cohort this feature may not be helpful for identifying IPMNs.

<sup>1</sup>The composite clinical marker is described in Masica et al<sup>[14]</sup>.

MPD = main pancreatic duct. chr = Chromosome. CI=confidence intervals. aneu = aneuploidy.



	<b>Sensitivity</b>	<b>Specificity</b>	<b>PPV</b>	<b>NPV</b>
Composite molecular marker	75%	92%	88%	83%
Composite clinical marker	77%	75%	70%	81%
Composite molecular/clinical marker	89%	69%	69%	88%

\* For this study, cysts were considered as appropriately resected if they were found on histopathologic examination to be SPNs, MCNs, or IPMNs that had high-grade dysplasia or were associated with adjacent invasive adenocarcinoma. Composite molecular and clinical markers are defined in Supplemental Table 5.

PPV=Positive predictive value. NPV=Negative predictive value.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript