RNA sequencing distinguishes benign from malignant pancreatic lesions sampled by EUS-guided FNA

Sarah A. Rodriguez, MD,[1,4] Soren D. Impey, PhD,[2] Carl Pelz, PhD,[2] Brintha Enestvedt, MD,[1] Gennadiy Bakis, MD,[1] Michael Owens, MD,[4] Terry K. Morgan, MD, PhD[3]

Portland, Oregon, USA

Background and Aims: EUS-guided FNA (EUS-FNA) is the primary method used to obtain pancreatic tissue for preoperative diagnosis. Accumulating evidence suggests diagnostic and prognostic information may be obtained by gene-expression profiling of these biopsy specimens. RNA sequencing (RNAseq) is a newer method of gene-expression profiling, but published data are scant on the use of this method on pancreas tissue obtained via EUS-FNA. The aim of this study was to determine whether RNAseq of EUS-FNA biopsy samples of undiagnosed pancreatic masses can reliably discriminate between benign and malignant tissue.

Methods: In this prospective study, consenting adults presented to 2 tertiary care hospitals for EUS of suspected pancreatic mass. Tissue was submitted for RNAseq. The results were compared with cytologic diagnosis, surgical pathology diagnosis, or benign clinical follow-up of at least 1 year.

Results: Forty-eight patients with solid pancreatic mass lesions were enrolled. Nine samples were excluded because of inadequate RNA and 3 because of final pathologic diagnosis of neuroendocrine tumor. Data from the first 13 patients were used to construct a linear classifier, and this was tested on the final 23 patients (15 malignant and 8 benign lesions). RNAseq of EUS-FNA biopsy samples distinguishes ductal adenocarcinoma from benign pancreatic solid masses with a sensitivity of .87 (range, .58-.98) and specificity of .75 (range, .35-.96).

Conclusions: This proof-of-principle study suggests RNAseq of EUS-FNA samples can reliably detect adenocarcinoma and may provide a new method to evaluate more diagnostically challenging pancreatic lesions. (Gastrointest Endosc 2016;84:252-8.)

EUS-guided FNA (EUS-FNA) biopsy sampling is a safe and useful tool to diagnose many pancreatic lesions.[1,2] Diagnostic accuracy is limited, however, especially when evaluating cystic lesions.[3-5] Better ancillary testing methods to aid clinicians and pathologists when making these challenging diagnoses are needed. Moreover, prospective studies investigating the neoplastic progression of low-grade intraductal papillary mucinous neoplasms, which may not require surgery, to high-grade lesions that do require resection would be greatly aided by molecular assays requiring no more than an EUS-FNA.[3-5]

Pancreatic carcinogenesis is thought to progress from dysplasia to malignancy as driving mutations accumulate.[5] Mutations in the K-*ras* oncogene, *p53*, and *SMAD4* tumor suppressor genes are common events in the dysplastic process,[5] but these alterations are not sensitive or specific to diagnose. In contrast to directly examining genetic mutations using DNA analysis of tissue, gene-expression levels can be measured by analyzing messenger RNA (mRNA), the precursor to protein synthesis. Examining which genes are over-expressed and under-expressed compared with normal or benign tissue is a method of separating types of tissue as well as a potential way of identifying useful clinical targets. Alterations in gene expression using mRNA can be measured using DNA

Abbreviations: cDNA, complementary DNA; mRNA, messenger RNA; RNAseq, RNA sequencing.

DISCLOSURE: *All authors disclosed no financial relationships relevant to this publication. Research support for this study was provided by a generous endoscopic research award from the American Society for Gastrointestinal Endoscopy (S. Rodriguez).*

0016-5107/$36.00
http://dx.doi.org/10.1016/j.gie.2016.01.042

Received November 11, 2015. Accepted January 14, 2016.

Current affiliations: Division of Gastroenterology (1), Oregon Stem Cell Center, Department of Pediatrics (2), Department of Pathology (3), Oregon Health & Science University, Portland, Oregon, USA; The Oregon Clinic Gastroenterology (4), Portland, Oregon, USA.

Reprint requests: Sarah A. Rodriguez, MD, The Oregon Clinic Gastroenterology-West, 9701 SW Barnes Rd Ste 300, Portland, OR 97239.

microarray platforms and direct mRNA sequencing. DNA microarray is an older method of gene-expression analysis where a known number of gene probes (short segments of DNA) are arranged on a chip. RNA is extracted from tissue and amplified, and then complementary DNA (cDNA) is generated using reverse transcriptase. The cDNA is then hybridized to the chip for analysis. RNA sequencing (RNAseq) is a method of sequencing the mRNA itself; it is a more direct method of measuring expression levels and has become the preferred modality. There has also been recent interest in microRNAs, which are small chains of noncoding RNAs that regulate gene expression post-transcription.

Transcriptome analysis using cDNA microarrays has been shown to reliably distinguish benign from malignant lesions,[6-9] but the major limitation of cDNA microarrays is the amount of RNA required for testing. Alternatively, RNAseq requires only 100 ng of total RNA for reliable and reproducible transcriptome results.[10,11] This concentration may be reasonably obtained from a single EUS-FNA biopsy specimen; therefore, we hypothesized that RNAseq of EUS-FNA biopsy specimens of solid pancreatic lesions may discriminate benign from malignant lesions.

METHODS

Study design

This is a prospective cross-sectional study of RNAseq of EUS-FNA biopsy samples from 48 adults with undiagnosed solid pancreatic masses obtained by EUS-FNA at a tertiary care medical center (Oregon Health and Science University Hospital) and regional medical center (Providence St. Vincent Medical Center, Portland, Ore). This study was approved by the institutional review boards at Oregon Health and Science University Hospital (no. 8317) and Providence St. Vincent (no. 13-009).

Study subjects were excluded if they had prior pancreatic or gastric resection, contraindications to EUS-FNA, inadequate cytology for diagnosis, or no mass found on EUS. Pancreatic endocrine neoplasms (n = 3) and cases with poor-quality RNA (n = 9) also were excluded. The first 13 cases were used to define transcriptome classifiers to predict outcomes for the subsequent 23 cases (Fig. 1) with available outcome data (diagnostic surgical resection specimen or at least 1 year of negative clinical follow-up).

EUS-FNA and clinical diagnoses

One additional EUS-FNA pass was obtained from each study subject using a 22-gauge needle with suction after diagnostic specimens were obtained for cytology and determined to be adequate by an onsite cytopathologist. Final cytopathologist diagnoses were recorded as either benign (pancreatitis) or malignant (adenocarcinoma). There were no "suspicious" or "atypical" cytologic diagnoses in this series. Final surgical pathology diagnoses were recorded when available as either benign (pancreatitis) or ductal adenocarcinoma. Benign clinical follow-up required at least 1 additional imaging study and benign clinical course after 1 year from index EUS-FNA study evaluation.

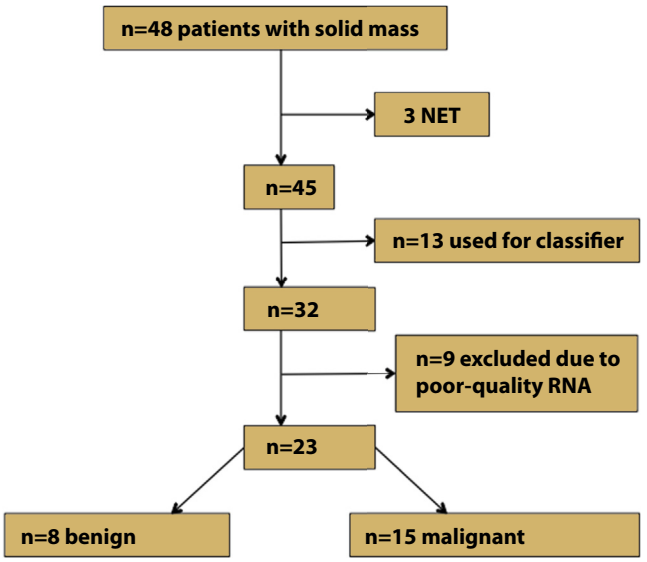

Figure 1. Study flow chart. *NET,* neuroendocrine tumor.

RNAseq

One hundred nanograms of RNA was extracted from a single EUS-FNA biopsy specimen and made into cDNA (Invitrogen, Grand Island, NY) for RNAseq using Illumina (Illumina, Inc, San Diego, Calif) high-throughput sequencing.[12]

RNAseq analysis

Quality control analyses included kmer-analyses of over-represented sequences and assessment of the percentage of sequences mapping to single locations, repetitive regions (RepeatMasker UCSC genome browser), and RefSeq exons. The distribution of exonic sequences relative to the 3′ end of RefSeq transcripts was also analyzed. We excluded samples that did not show a raw read alignment percentage > 40% to single locations or did not show a RefSeq exon annotation percentage > 40% of aligned sequences.

For classifier construction, training data were preprocessed by removing invariant genes and genes with low expression (eg, <10 copies per million; the threshold for analysis was set to ≥10 sequences per gene after normalizing for the average sequencing depth). We iterated through a range of filtering thresholds as we refined our classifier to test robustness. We also filtered based on significance (false discovery rate–adjusted *P* value via EdgeR and DESeq). This linear classifier was intended to be purely diagnostic (not predictive of outcome or response to therapy), and no univariate analyses were conducted.