

SUPPLEMENTARY DATA (Matthaei et al)

Supplementary Methods

Patients and biospecimens

This study was approved by the Johns Hopkins University Institutional Review Board. The prospectively maintained Johns Hopkins Surgical Pathology Database was scrutinized to identify formalin-fixed, paraffin-embedded (FFPE) tissue specimens of patients who underwent pancreatectomy for IPMN between January 1st, 2000 and August 31st, 2010 at Johns Hopkins hospital. Hematoxylin and eosin stained (H+E) reference slides were used to identify samples for subsequent molecular studies.

Histologic diagnoses were reconfirmed by two pathologists (AM, RHH) according to the latest World Health Organization recommendations (WHO). Briefly, IPMNs had to display a papillary epithelium with abundant extracellular mucin and measure per definition >1cm in maximum diameter. Main duct IPMNs were distinguished from branch duct IPMNs, and a third category of a mixed type was assigned whenever the lesion was located in both main and branch duct. In each case, the final diagnosis referred to the most severe grade of dysplasia observed in the neoplastic epithelium, including low grade (LG), intermediate grade (IG), and high grade (HG) IPMNs. Furthermore, we assessed whether an IPMN had an associated invasive carcinoma.

For unbiased high-throughput (HT) miRNA expression profiling (“FFPE tissue study 1” or “FTS1”, **Figure 1**) we selected 10 LG IPMNs and 12 HG IPMNs. Seven of the latter had an associated invasive adenocarcinoma. For validation of candidate miRNAs in an independent set of specimens (“FFPE tissue study 2” or “FTS2”, **Figure 1**) additional 33 archival IPMNs were

selected (6 LG IPMNs, 14 HG IPMNs, and 13 HG IPMNs with an associated invasive adenocarcinoma). Diagnostic information for all FFPE IPMN specimens is collated in **Supplemental Table 1**.

Furthermore, we selected 65 samples from a prospectively maintained cyst fluid bank at Johns Hopkins, all of which were harvested from pancreatic cysts resected at the Department of Surgery, Johns Hopkins Hospital. Cyst fluid specimens were aspirated immediately after surgical resection in the Department of Surgical Pathology using a sterile syringe, aliquoted in sterile 1.5 ml tubes and stored at -80° C within 30 min after resection. For cysts with multiple locules, the aspirated fluid was pooled, in order to obtain a panoramic representation of the resected lesion. Histologic diagnoses of the 15 CF specimens used in HT miRNA expression profiling (“Cyst fluid study 1” or “CFS1”, **Figure 1**) were as follows: 5 LG IPMNs, 5 HG IPMNs (3 had an associated invasive cancer) and 5 SCAs. For miRNA candidate validation (“Cyst fluid study 2” or “CFS2”, **Figure 1**) we used an independent set of 50 cyst lesions with associated CF specimens, which was composed of 2 LG IPMNs, 12 IG IPMNs, 6 HG IPMNs (one with an associated invasive carcinoma), 20 SCAs, 5 PanNETs and 5 SPNs. Diagnostic information on all lesions from which cyst fluid specimens were harvested is compiled in **Supplemental Table 2**.

Laser microdissection

Sections (6-10 μ m thickness) were embedded onto UV-pretreated PALM[®] membrane slides and stained with H+E prior to laser microdissection. Where extensive papillary epithelium could be grossly identified on the slide, the neoplastic epithelium was microdissected using ultra-fine high-precision tweezers (Electron Microscopy Sciences, Hatfield, PA) under a stereoscopic zoom microscope SMZ1500 (Nikon, Tokyo, Japan). The number of cells harvested by microdissection from FFPE specimens ranged from 5,000–15,000 and 1,250–7,500 for the

FTS1 and FTS2specimens, respectively (**Supplemental Figure 1; Supplemental Table 1**). In contrast to the pooled analysis from multiple locules for cyst fluid specimens, we only microdissected one defined area of cyst epithelium for tissue-based profiling, usually the epithelium most representative and amenable to microdissection.

RNA extraction from FFPE tissue and cyst fluid specimens

Total RNA from microdissected FFPE tissues was extracted using the RecoverAll™ Total Nucleic Acid Isolation Kit for FFPE (Ambion, Austin, TX) according to the manufacturer`s protocol with slight modifications to increase the final RNA concentrationfor downstream applications. This method allows robust and reproducible recovery of RNA from FFPE tissues in sufficient quality and quantity to support miRNA expression profiling studies. The average RNA recovery from cells lifted from FTS1 and FTS2 specimens was 1420ng (range: 245–4650 ng) and 840 ng (range: 159–3552 ng), respectively.

Total RNA from cyst fluid specimens (0.05-1.5 mL) was extracted according to Asuragen`s standard operating procedures using an in-house developed mirVana™ PARIS™ (Ambion) - based protocol. The concentration and purity of RNA were assessed with a NanoDrop 1000 spectrophotometer (NanoDrop Technologies/Thermo Scientific, Wilmington, DE).

MiRNA expression analyses in FFPE tissue and cyst fluid specimens

High-throughput (HT) miRNA expression analyses

10 ng total RNA was converted into cDNA using Megaplex RT Primers (Applied Biosystems) and TaqMan miRNA RT Kits (Applied Biosystems). cDNA was pre-amplified (12

cycles) using Megaplex PreAmp Primers (Pool A and/or Pool B) prior to mixing with TaqMan Universal PCR Master Mix (Applied Biosystems) and loading onto TaqMan human miRNA fluidic cards (Applied Biosystems). Expression of 750 mature miRNAs (Pool A and B) and 377 mature miRNAs (Pool A) was examined for FTS1 and CFS1, respectively. The cards were run using the Applied Biosystems 7900HT real-time PCR instrument equipped with a heating block for the fluidic card (Applied Biosystems). Prior to bioinformatics analysis, raw data were processed using Relative Quantification ($\Delta\Delta\text{Ct}$) and the RQ Manager, with baseline set to “automatic” and Ct threshold set to 0.2.

Bioinformatics Analysis

FFPE tissue study

As a prerequisite to the screen, miRNAs were removed unless they had strong Ct values with both mean and median Ct <30 (based on Megaplex data). Note that this criteria tended to remove miRNAs with multiple non-determined calls. A priority list of candidates was ranked with a simple heuristic: the sum of the negative log of the FDR (False Discovery Rate) values from the t-test and Wilcox-test. To preclude bias that could result from the removal of miRNAs before the FDR estimates are calculated, the final selected candidates had FDR values <0.05 for both tests before any filtering. Finally, additional candidates were selected based on publications or internal studies implicating their roles in pancreatic cancer. The short list of candidates was verified using singleplex RT-qPCR on the original samples profiled by Megaplex (FTS1), plus an additional set of samples (FTS2). For evaluation of batch effects, the microdissected FFPE specimens (**Supplemental Table 1**) were assigned to one of two groups: HG IPMN (including specimens from FTS1 and specimens without and with cancer from the FTS2) and LG IPMN (comprising specimens from FTS1 and FTS2). T-test analysis showed no evidence of significant batch effects for any of the miRNAs selected for further analysis. miRNAs with

average expression values ≥ 35 Ct in the singleplex candidate verification (FTS1) and validation (FTS2) across all the samples were considered to be non-specifically amplified, and therefore were excluded from the final data analysis.

Cyst fluid study

A sequential set of filters was used to generate a final list of miRNA candidates. The filtering was done step-wise in order to minimize cost and time, yet maintain a sufficiently broad panel of candidates. First, the Megaplex analysis was performed using 5 SCA, 5 LG IPMN, and 5 HG IPMN specimens. One LG IPMN specimen, CFS1-4, was retained for the verification of candidate miRNAs in order to preserve RNA. Another, CFS1-1, was unintentionally left out of the analysis, but when it was included towards the end of the study, it had no major impact on the shortlist of miRNA candidates generated without it. One SCA specimen (CFS1-12) was excluded from analysis on the basis of a large number of missing Ct values (~68%). One HG IPMN specimen (CFS1-9) clustered with the LG IPMNs in an unsupervised PCA analysis (**Figure 2B**). Because data analysis performed with and without CFS1-9 showed qualitatively similar results, this specimen was removed as well. As a result, bioinformatics analysis was performed on 11 CF specimens (4 SCA, 3 LG IPMN and 4 HG IPMN). Differentially expressed miRNAs from the Megaplex data were identified using a t-test, while DiffPairs (expression of one miRNA subtracted from another to generate a self-normalizing biomarker) were identified with t-test across diagnostic groups. No miRNA Ct cutoff below the experimental Ct limit of 40 was incorporated into the DiffPair analyses. Based on the Megaplex data analysis of the cyst fluid study, we derived two candidate listings. The first candidate set was derived from the top 10 DiffPairs that included 5 DiffPairs with a FDR value of 0.0475 and the next 5 DiffPairs added on based on comparable *P*-values (**Supplemental Table 3**). This yielded 17 distinct miRNAs. The second candidate set was composed of the top 10 individual miRNAs (with unadjusted *P*-values <0.01), producing 5 miRNAs not present in the top 10 DiffPairs (**Supplemental Table 4**).

These two candidate sets produced a total of 22 miRNA candidates.

The two candidate sets based on Megaplex CF data analysis were merged with candidates from the FFPE tissue study and generated 37 miRNAs, derived from the top 10 cyst fluid DiffPairs (17miRNAs in **Supplemental Table 3**) combined with the top ten individual miRNAs (**Supplemental Table 4**), the top 13 FFPE tissue miRNAs from the 30 DiffPairs (**Table 1, Supplemental Table 3**) and with six miRNAs (let-7b, miR-223, miR-30b, miR-328, miR-532-3p, miR-590-5p) selected based on high expression levels and their performance as individual candidates and in DiffPairs. These 37 miRNAs were verified with singleplex RT-qPCR for all CFS1 samples, for which there was enough RNA available. Eleven specimens were profiled with the full 37 miRNA panel. CFS1-2, CFS1-4, CFS1-11 and CFS1-14 were interrogated with 13 miRNA candidates only (**Table 2**). CFS1-9 and CFS1-12 were omitted from analysis for the same reasons as described in the Megaplex analysis above.

Candidates from the CFS1 singleplex RT-qPCR data set were evaluated as DiffPairs. Expression values above 32 Ct were treated as missing in order to filter out low-signal miRNAs from further consideration. Candidate DiffPairs were assessed by t-test for significant differential expression using FDR-adjusted P -values <0.05 . CFS1 singleplex analysis reduced the candidate 37 miRNA set to 27 DiffPairs composed of 18 miRNAs (**Table 3**).

Expression of those 18 miRNAs together with miR-21 was further evaluated by singleplex RT-qPCR in a second set of 50 CF specimens (CFS2). One IG IPMN specimen (CFS2-8) was excluded due to high Ct values (median Ct >32), leaving 49 samples for further analysis. IG IPMNs were only included in the training set. These samples, along with the CFS1 specimens, were used to develop and assess a logistic regression model for sample classification as described below.

Of particular note, correlation between mean Ct and diagnosis was observed for cyst fluid samples as shown by ANOVA, yielding R^2 values of 0.81 for Megaplex CFS1, 0.63 for singleplex CFS1, and 0.49 for CFS2 (all significant at $P < 0.05$ level) (**Supplemental Figure 2, Supplemental Figure 5**).

Logistic Regression Model to Guide Resection

The logistic regression model was trained and tested on data for the 18 cyst fluid miRNAs (**Table 3**) and miR-21 from training and testing subsets drawn from the 9 CFS1 and 49 CFS2 specimens (see **Supplemental Table 2** for included specimens). No apparent batch effects were observed between the two specimen sets (**Supplemental Figure 6**). The CFS1 and CFS2 specimens were merged together and then split into training and test sets as detailed in **Supplemental Table 2**. Of note, the 4 specimens with incomplete RT-qPCR data for 13 miRNAs (CFS1-2, CFS1-4, CFS1-11 and CFS1-14) were not used in training or testing of the logistic regression model in order to mimic prospective validation. CFS1-9 and CFS1-12 were also not used for the same reasons as described above. The separation of samples into training and test sets was based on order of isolation.

The first round of feature selection was conducted through pair-wise comparison of DiffPair ΔCt values in benign (LG IPMN/SCA) as compared to malignant (HG IPMN/PanNET/SPN) samples on the 21 training set specimens. Despite the lack of apparent batch effects, a two-way ANOVA model (without interactions) was fit to the data for each gene, with discovery/validation batch as one factor and malignancy status as the other. P -values were obtained for benign *versus* malignant ANOVA contrasts, testing null hypotheses of no difference between benign and malignant groups (**Supplemental Table 5**). The 20 DiffPairs found to be most differentially expressed in this analysis were used as predictors for a logistic model for

distinguishing malignancy status (**Supplemental Table 5**). The logistic model was fit to the 20 selected DiffPairs using L1-penalized regression, with the penalty parameter manually optimized (final value $\lambda_1=0.1$ using variance standardization for predictors; cross-validated log-likelihood was essentially constant near maximum value for λ_1 between 0.05 and 0.15) through leave-one-out cross-validation using the `cvl` and `penalized` functions in the R `penalized` package. Manual optimization of CVL was used because of numerical difficulties encountered with the automatic `opt L1` function for penalty parameter estimation. The parameters of the resulting model are shown in **Supplemental Table 5**. Note that only 7 of the 20 predictor DiffPairs received non-zero weights due to the L1-penalty applied during the fitting process.

Not surprisingly, the weights of the DiffPairs in the regression model are correlated with the ANOVA contrast p-values. Note in particular that the DiffPair with the lowest ANOVA contrast *P*-value, Diff(miR-24, miR-30a-3p), also received the largest weight in the regression model, with 5 of the 7 non-zero regression weighted DiffPairs appearing in the top 10 DiffPairs by ANOVA contrast *P*-value.

Supplemental Figures

Supplemental Figure 1: (A) Laser microdissection in an IPMN (hematoxylin and eosin; original magnification: 10x); (B) After precise marking and microdissection the neoplastic epithelium is catapulted into a sterile tube for subsequent RNA extraction. (C) The section after dissection.

Supplemental Figure 2: Detailed diagram of the study design, describing experimental setup of FFPE (FTS) and cyst fluid (CFS) studies, including number of specimens used and number of miRNA candidates identified in the course of each study. Legend: SCA: serous cystadenoma, LG: low grade, IG: intermediate grade, HG: high grade, IPMN: intraductal papillary mucinous neoplasm, HT: high throughput, PanNET: pancreatic neuroendocrine tumor, SPN: solid papillary neoplasm, ¹: indicates that all specimens described in the heading were analyzed, ²: indicates that a subset of the specimens described in the heading were analyzed; more details on excluded specimens can be found **Supplemental Table 1** and **2**.

Supplemental Figure 3: Mean Cts for all FTS and CFS specimens. (A) Megaplex RT-qPCR data for FTS. (B) and (C) Singleplex RT-qPCR data for FTS. (D) Megaplex RT-qPCR data for CFS. (E) and (F) Singleplex RT-qPCR data for CFS.

Supplemental Figure 4: Boxplots showing raw Ct values for Megaplex (A) and singleplex (B) RT-qPCR expression analyses of FTS1 and FTS1 plus FTS2, respectively. The diagnosis associated with each sample is indicated by color, while the FTS1 and FTS2 specimens are depicted as separate panels in the singleplex boxplot (B).

Supplemental Figure 5: Boxplots showing raw Ct values for Megaplex (A) and singleplex (B) RT-qPCR expression analyses of cyst fluid specimens from the CFS1 and CFS1 plus CFS2,

respectively. Diagnosis is indicated by color, the CFS1 and CFS2 specimens by panel.

Supplemental Figure 6: Boxplots of raw Ct values for singleplex RT-qPCR expression analysis of the CFS1 and CFS2 specimens with reassignment to test or training set indicated by separation of panels.

Supplemental Figure 7: Raw Ct values of miRNAs involved in DiffPairs comprising the logistic regression model in the CFS1 and CFS2 specimens. Raw Ct for miR-21, which is not a part of the logistic regression model, are also shown.

Supplemental Figure 8: PCA applied to raw Cts (**A**) and restricted mean-center normalized Cts. (**B**) CFS1 and CFS2 singleplex RT-qPCR data. Note that CFS1 and CFS2 specimens do not separate in either plot.

Supplemental Table 1: Total RNA recovery, demographic and tumor-related information for FFPE tissue specimens. Summary statistics for LG and HG IPMNs are included; specimens excluded from bioinformatics analyses are indicated. Legend: “Excluded: missing values” – more than 10% of miRNAs interrogated in a given specimen were not amplified; “Excluded: mean Ct>30” - the average Ct for those miRNAs amplified was greater than 30 and therefore considered insufficient miRNA recovery.

Supplemental Table 2: Total RNA recovery, demographic and tumor-related information for cyst fluid specimens. Summary statistics for experimental groupings are included; specimens excluded from bioinformatics analyses are indicated. Legend: “Excluded: insufficient material to test all candidates” – not all candidates could be interrogated in a given specimen due to limiting RNA; “Excluded: to preserve RNA for candidate verification” – not used for candidate identification to preserve RNA for downstream analyses; “Excluded: mean Ct>32” - the average Ct for amplified miRNAs was greater than 32, and therefore considered insufficient miRNA recovery; “Excluded: many missing values” - more than 10% of miRNAs interrogated in a given specimen were not amplified.

Supplemental Table 3: Top DiffPairs identified by t-test p-value from the CFS1 Megaplex data set.

Supplemental Table 4: Top 10 miRNA identified by t-test p-value from the CFS1 Megaplex data set.

Supplemental Table 5: Regression coefficients for the 20 DiffPairs selected for use as predictors in the logistic regression model. Because of the use of L1-penalized regression, 13 of these predictors received 0 weights, and were thereby omitted from the final model. The p-values from the ANOVA analysis contrasting higher risk (HG IPMN/SPN/PanNET) samples to low risk (LG IPMN/SCA) samples are also provided.