Plasma microRNAs as biomarkers of pancreatic cancer risk in a prospective cohort study

Eric J. Duell, Leila Lujan-Barroso, Nuria Sala, Samantha Deitz McElvey, Kim Overvad, Anne Tjonneland, Anja Olsen, Elisabete Weiderpass, Lill-Tove Busund, Line Moi, David Muller, Paolo Vineis, Dagfinn Aune, Giuseppe Matullo, Alessio Naccarati, Salvatore Panico, Giovanna Tagliabue, Rosario Tumino, Domenico Palli, Rudolf Kaaks, Verena A. Katzke, Heiner Boeing, H. Bas Bueno-de-Mesquita, Petra H. Peeters, Antonia Trichopoulou, Pagona Lagiou, Domenico Palli, Navarra and the Catalan Institute of Oncology, Barcelona, Spain

Key words: biomarkers, microRNAs, pancreatic cancer, cohort studies

Disclosure: The authors have declared no conflicts of interest

Additional Supporting Information may be found in the online version of this article.

Grant sponsor: the National Cancer Institute (NCI) of the National Institutes of Health; Grant number: CA-075059 (to M. Korc); Grant sponsor: Indiana Economic Development Fund (IEDF) (to M. Korc); Grant sponsors: The European Commission (DG-SANCO) and International Agency for Research on Cancer (for coordination of EPIC); Grant sponsor: Health Research Fund (FIS) of the Spanish Ministry of Health, Regional Governments of Andalucía, Asturias, Basque Country, Murcia; Grant number: no. 6236; Grant sponsor: Navarra and the Catalan Institute of Oncology, Barcelona, Spain; Grant number: BM 06-130; Grant sponsor: Red Temática de Investigación Cooperativa en Cáncer; Grant numbers: RD12/0036/0018; RD06/0020/0091; Grant sponsors: Danish Cancer Society (Denmark); Ligue contre le Cancer, Institut Gustave Roussy, Mutuelle Générale de l’Education Nationale, Institut National de la Santé et de la Recherche Médicale (INSERM) (France); Deutsche Krebshilfe, Deutsches Krebsforschungszentrum (DKFZ) and Federal Ministry of Education and Research (Germany); the Hellenic Health Foundation (Greece); Associazione Italiana per la Ricerca sul Cancro (AIRC) and National Research Council (Italy); Dutch Ministry of Public Health, Welfare and Sports (VWS), the Netherlands Cancer Registry (NKR), LK Research Funds, Dutch Prevention Funds, Dutch ZON (Zorg Onderzoek Nederland), World Cancer Research Fund (WCRF) and Statistics Netherlands (the Netherlands); Nordic Center of Excellence in Food, Nutrition and Health-Helga (Norway); Swedish Cancer Society, Swedish Scientific Council and Regional Government of Skåne and Västerbotten (Sweden); Cancer Research UK, Medical Research Council (United Kingdom)

DOI: 10.1002/ijc.30790

History: Received 6 Feb 2017; Accepted 19 Apr 2017; Online 19 May 2017

Correspondence to: Dr. Eric J. Duell, Unit of Nutrition and Cancer, Catalan Institute of Oncology, Barcelona, Spain, Tel.: +34 93 260 7401, E-mail: eduell@iconologia.net or Dr. Murray Korc, Departments of Medicine, and Biochemistry and Molecular Biology, Indiana University School of Medicine, and the Pancreatic Cancer Signature Center, Indiana University Simon Cancer Center, Indianapolis, IN 46202, Tel.: 317-278-6410, E-mail: mkorc@iu.edu

Noninvasive biomarkers for early pancreatic ductal adenocarcinoma (PDAC) diagnosis and disease risk stratification are greatly needed. We conducted a nested case-control study within the Prospective Investigation into Cancer and Nutrition (EPIC) cohort to evaluate prediagnostic microRNAs (miRs) as biomarkers of subsequent PDAC risk. A panel of eight miRs (miR-10a, -10b, -21-3p, -21-5p, -30c, -106b, -155 and -212) based on previous evidence from our group was evaluated in 225 microscopically confirmed PDAC cases and 225 controls matched on center, sex, fasting status and age/date/time of blood collection. MiR levels in prediagnostic plasma samples were determined by quantitative RT-PCR. Logistic regression was used to model levels and PDAC risk, adjusting for covariates and to estimate area under the receiver operating characteristic curves (AUC). Plasma miR-10b, -21-5p, -30c and -106b levels were significantly higher in cases diagnosed within 2 years of blood collection compared to matched controls (all p-values < 0.04). Based on adjusted logistic regression models, levels for six miRs (miR-10a, -10b, -21-5p, -30c, -106b and -212) were significantly associated with subsequent risk. A score based on the panel showed a linear dose-response trend with risk (p-value = 0.0006). For shorter follow-up (<5 yr), AUC for the score was 0.73, and for individual miRs ranged from 0.73 (miR-212) to 0.79 (miR-21-5p).

Pancreatic cancer has the worst prognosis of all cancers because of late detection, and early noninvasive biomarkers are urgently needed to improve outcome. Here, the authors examined blood levels of a panel of microRNAs previously associated with pancreatic cancer. Samples collected within two years of a cancer diagnosis showed elevated plasma levels of miR-10b, miR-21-5p, miR-30c and miR-106b. In multivariable models, three of these microRNAs (10b, 21-5p, 30c) as well as miR-10a showed significant associations with subsequent risk underscoring the clinical potential of plasma microRNAs in pancreatic cancer screening.
death by 2030.3 Little is understood in terms of preventable risk factors, and known risk factors including common genetic variants are inadequate for risk stratification or early diagnosis; thus, a top research priority in PDAC research has been to identify noninvasive early markers of risk.

MicroRNAs (miRs) are small noncoding RNAs consisting of 18–25 nucleotides that function through several mechanisms. In general, miRs target specific mRNAs for translational repression or degradation, and thus regulate important biological processes including cell proliferation, survival, invasion and metastasis.4–7 In comparison with the normal pancreas, numerous miRs are overexpressed in PDAC, including miR-21, -10b, -30c, -196a, -203, -155, -101b, -205, -221, -222, -223, -486, -744 and -10a.5,8–21 Both miR-10b and miR-21 are frequently up-regulated in PDAC, where the presence of high levels of each of these miRs has been associated with lower patient survival and responsiveness to gemcitabine.9,11,14,19 Moreover, miR-10b enhances the pro-metastatic actions of EGF and TGF-β.5

We selected the miRs examined in the present study because, with the exception of miR-21-3p, they are all over-expressed in PDAC and may contribute to PDAC progression and/or metastasis.5,8–22 We chose to include miR-21-3p which originates from the opposite arm of the same pre-miR as miR-21-5p that is recognized as an oncogenic miR (Onco-miR) in PDAC,10,11 and because miR-21-3p also may have oncogenic properties.23,24 In the present nested study of prediagnostic plasma samples from 225 PDAC cases and 225 matched controls from the EPIC cohort, we evaluated relative expression of a panel of eight miRs (miR-10a, miR-10b, miR-21-3p, miR-21-5p, miR-30c, miR-106b, miR-155 and miR-212) and PDAC risk, adjusting for known risk factors and potential confounders, as well as examination of risk estimates by follow-up time between blood collection and PDAC diagnosis. Importantly, miR-10b, -21-5p, -30c, -106b and -155 are expressed in pancreatic intra-epithelial neoplasia (PanIN) that are the immediate precursor lesions to PDAC.25–27

Materials and Methods
Population
The EPIC cohort includes a total of 521,457 participants (368,010 women and 153,447 men) recruited through 23 research centers in 10 European countries including Denmark (Aarhus, Copenhagen), France, Germany (Heidelberg, Potsdam), Greece, Italy (Florence, Turin, Varese, Naples, Ragusa), The Netherlands (Bilthoven, Utrecht), Norway, Spain (Asturias, Granada, Murcia, Navarra, Guipuzcoa), Sweden (Malmo, Umeå) and the United Kingdom (Oxford, Cambridge).28 Most of the participants were enrolled between 1992 and 1998 between the ages of 35 and 70 years. Participants were recruited from the general population residing in the corresponding geographic areas including towns and provinces, except for the French cohort which recruited participants from a teacher’s organization health insurance program, the cohorts consisting of women attending breast cancer screening programs (Utrecht and Florence), parts of the Italian and Spanish cohorts in which participants were recruited from among blood donors, and most of the Oxford cohort which recruited mainly health-conscious participants, including vegetarians. Eligible participants gave written informed consent and completed questionnaires on diet, lifestyle and medical history at baseline. Anthropometric characteristics were measured by interviews except in France and Norway, and for the majority of participants from EPIC-Oxford (although the accuracy of these self-reported data has been validated29), for whom the data were self-reported in the lifestyle questionnaires. All participants were cancer-free at the time of data collection, including biological samples obtained from approximately 80% of the cohort. Ethical review boards from the International Agency for Research on Cancer (IARC) and local centers approved the study. In addition, approval for obtaining blood samples from patients at Indiana University School of Medicine (IUSM) was granted by the local Institutional Review Board, and each participant from whom blood was collected provided written informed consent prior to sample collection.

Nested case-control study
A nested case-control study within EPIC was conducted for the analyses of miR levels and PDAC risk. Cohort follow-up started at study enrolment and continued until diagnosis, death or last completed follow-up, whichever came first. Cancer incidence in EPIC was determined through population-based registries or through active follow-up. Cases had no prior cancer history (other than nonmelanoma skin cancer) and included first primary adenocarcinoma of the pancreas ICD-Oncology third edition codes C25.0-C25.3 and C25.7-C25.9 which did not include endocrine or neuroendocrine pancreatic tumors. All PDAC cases selected for this study were microscopically confirmed (based on histology of the primary tumor, metastases, cytology or autopsy findings). Control participants were selected randomly among EPIC participants who were alive and free of cancer (except nonmelanoma skin cancer) at the time of diagnosis of each index case and matched to each case based on sex, study center or country, age at blood collection (± 3 years), date of blood collection (± 3 months), time of blood collection (± 2 hrs), fasting status (<3, 3–6, >6 hrs after last meal) and among women, use of hormones (OC, HRT). The final sample size was 225 cases and 225 matched controls. Sweden (Malmo, Umeå) did not participate in the study.

miR selection and assay methodology
EPIC plasma samples were stored in Denmark (-150°C nitrogen vapor), and at the IARC EPIC-biobank for the remaining non-Danish samples (-196°C liquid nitrogen). Total RNA was isolated from plasma samples using Trizol-LS (Life Technologies, Carlsbad, CA, USA) and Direct-zol RNA MiniPrep kit (Zymo Research). cDNA was generated using 10 ng of
RNA per reaction in conjunction with miR-10a, -10b, -30c, -21-3p, -21-5p, 106b, -155, -212 or -425-5p RT primers and a miR reverse transcription kit (Life Technologies), as previously reported.\(^{22}\) Quantitative PCR (qPCR) was performed using miR Taqman\(^ {\oplus}\) probes and TaqMan Fast Advanced Master Mix.\(^ {22}\) Expression levels for candidate miRs were normalized to miR-425-5p, which was expressed at similar levels across all sample types, exhibiting <1 cycle threshold (Ct) difference across the samples.\(^ {22}\) After normalization to miR-425-5p (ΔCt), the ΔCt values for miRs in controls were averaged and subtracted from the ΔCt values of each individual sample (ΔΔCt), and expression levels were calculated using the 2\(^ {−}\Delta\DeltaCt\) method.\(^ {30}\)

All miR measurements were initially blinded to case-control status. To confirm assay reproducibility, samples from seven subjects (four cases, three controls) were reanalyzed with similar results for all miRs. Moreover, since the same PCR instrument (ViiA7) and methodology were used in our current and previous study,\(^ {22}\) we compared plasma samples from a cohort of normal controls at IUSM with the normal values in the current study following data unblinding. In spite of differences between EPIC and IUSM with respect to the duration of sample storage, temperature (−150°C or -196°C for European samples vs -70°C for American samples), there was remarkable concordance between the two studies with respect to ΔCt values in the control samples for each miR (Supporting Information Table S1).

In view of the near identical seed sequence for miR-10a and -10b, and the fact that their mature forms only differ at a single nucleotide, we next confirmed that our qPCR method readily distinguished between miR-10a and -10b (Supporting Information Fig. S1).

To confirm that miRNA-425-5p is a reliable internal control, following completion of the RT-qPCR assays and designation of control and PDAC cases, the mean Ct values for miRNA-425-5p in all 450 EPIC samples were compared with the corresponding miRNA mean Ct values for the 225 control subjects and the 225 individuals who were diagnosed with PDAC (Supporting Information Table S2). There was remarkable uniformity in miRNA-425-5p levels across the entire study, validating its utility as an internal control miRNA.

To validate the efficiency of our RNA extraction procedure, we performed a spike-in experiment using plasma samples from three normal and three PDAC patients from the IU Simon Cancer Center Tissue and Biofluid Bank that were spiked with varying levels of cel-miRNA-39 (Supporting Information Table S3). The mean threshold (Ct) values ± SEM in normal and PDAC samples were then determined for cel-miRNA-39, hsa-miRNA-425-5p, hsa-miRNA-16 or RNU6B, by RT-qPCR. This analysis confirmed the efficiency of our RNA extraction procedure and the uniformity of miRNA-425-5p levels even in plasma from non-EPIC samples. Moreover, miR-16 and RNU6B levels were fairly uniform across normal and PDAC samples.

Danish samples were subjected to a flood in 2011. Therefore, we next evaluated whether miR expression for each miR was systematically different between Danish and non-Danish samples.\(^ {31}\) Among 146 Danish and 79 non-Danish control samples, only two miRs (miR21-3p and miR-106b) had lower median expression in Danish samples, but neither difference was statistically significant (Wilcoxon rank sum test p-values ≥0.19). In general, Danish samples showed little evidence of miR degradation in comparison to non-Danish samples. These observations are consistent with the fact that there was no evidence that water entered the vials storing the Danish samples, and that miRs are resistant to degradation under many conditions, including at least 10 cycles of freeze-thawing.\(^ {22}\)

**Statistical methods**

**Data transformation.** Expression levels for each miR were modelled continuously using log\(_{2}\)-transformation to facilitate interpretation of results and improve normality of the distributions. Spearman correlation coefficients (\(r_s\)) were used to assess correlations among the un-transformed miRs and covariates. A nonparametric Wilcoxon rank sum test was used to compare median untransformed miR levels by case-control status and other dichotomous variables. Nonparametric Kruskal-Wallis tests were used to compare median untransformed miR levels according to nominal variables.

**Logistic regression models.** Conditional logistic regression models for associations between log\(_{2}\)-transformed miR expression level and PDAC risk [where odds ratios (OR) represent relative risk per doubling of miR expression level] were conditioned on matching variables and adjusted for age at recruitment (years), smoking intensity (never, current 1–15 cig/d, current 16–25 cig/d, current 26+ cig/d, former quit ≤10 years, former quit 11–20 years, former quit 20+ years and current users of nicotine tobacco products) or smoking status (never, former, current), baseline alcohol intake (g/day), education (none, primary, technical/professional, secondary, university/graduate school), BMI (kg/m\(^2\)) and physical activity using the Cambridge index (inactive, moderately inactive, moderately active and active).\(^ {33}\) As a result of sample size constraints, conditional logistic regression models that were stratified by the follow-up time between blood collection and PDAC diagnosis were adjusted for age, smoking status and alcohol intake only.

**miR score and interactions.** A miR score was created for each participant by summing each miR (except miR-21-3p which had similar expression levels between cases and controls) that was over-expressed above the 3\(^{rd}\) quartile (75\(^{th}\) percentile) in all 225 controls combined (range: 0–7). Participants with scores of 6–7 (the highest category) were combined because of small sample size. In exploratory analyses, two-fold interactions between each miR were evaluated by categorizing each miR as a dichotomous variable or in quartiles (using cutpoints based on the distribution in controls).
Joint ORs were estimated for PDAC risk and pair-wise overexpression of miRs. Statistical significance of potential interactions was evaluated using a likelihood ratio test (LRT).

**Receiver operating characteristic (ROC) curves.** Predictive performance was assessed with respect to the ability of each miR and the miR score to discriminate between PDAC cases and controls using the Receiver operating characteristic (ROC) curve which evaluates the true positive rate (sensitivity) on the y-axis and the false positive rate (1-specificity) on the x-axis, with excellent accuracy defined as AUC >0.90. Predictive performance for each miR was evaluated based on unconditional logistic regression models adjusted for matching factors, smoking status and alcohol intake at baseline. Models for the miR score were additionally adjusted for BMI. To perform the ROC curve and to assess AUC, functions *prediction* and *performance* from the ROCR R package were used; to assess the 95% confidence intervals for the AUC, function ci.cvAUC implemented in the cvAUC R package was used. All statistical analyses were performed using SAS v.9.4 (Cary, NC, USA) and R v.3.2.5.

**Results**

The median follow-up time between blood collection and PDAC diagnosis in the 225 cases was 7.85 years (range: 10.86 years; inter-quartile range: 4.22–9.47 years). Correlation coefficients ($r_s$) among the eight miRs ranged from 0.09 to 0.6 in control participants ($p$-values <0.0001 to 0.18) with the strongest correlation between miR-10b and miR-212, and the weakest between miR-10b and miR-106b (Supporting Information Table S4). There was little evidence for correlation between miR expression levels and age at recruitment ($r_s$ range -0.037 to 0.057, $p$-values 0.39 to 0.95), except for miR-10b ($r_s$ 0.13, $p$ values 0.049). There was no evidence for correlations between the eight miRs and smoking intensity (correlation range -0.01 to 0.06, $p$-values 0.13 to 0.89), or BMI (correlation range -0.0003 to 0.037, $p$-values 0.54 to 0.99). Individual miR expression levels did not significantly differ by sex (all Wilcoxon rank-sum test $p$-values >0.26).

Only miR-10b was weakly correlated with baseline alcohol intake (correlation coefficient 0.16, $p$ values 0.017), while the remaining miRs were not (correlation range -0.041 to 0.12, $p$-values 0.069 to 0.79). None of the miRs was significantly associated with physical activity level (Kruskal-Wallis test $p$-values 0.06 to 0.8), or education level (Kruskal-Wallis test $p$-values 0.09 to 0.7).

Comparison of expression levels for each miR between patients subsequently diagnosed with localized PDAC ($n = 13$) vs. metastatic PDAC ($n = 96$) did not show any significant differences (all $p$-values >0.18; 116 cases had missing data on stage); however, the possibility that some patients with suspected localized PDAC already had micro-metastases cannot be excluded. Nine cases and 9 controls reported having had a history of diabetes at baseline (16 cases and 5 controls were missing data on diabetes). Median relative expression was statistically significantly lower in diabetics vs. nondiabetics for miR-155 (cases: 0.43 vs. 0.75, $p$ values = 0.016; controls: 0.29 vs. 0.72, $p$ values = 0.024), and miR-21-5p (controls: 1.07 vs. 1.11, $p$ values = 0.029). Because of the low number of participants who reported a positive history of diabetes at baseline, this covariate was not included in conditional logistic regression models.

Levels for each miR were compared between cases and controls according to the follow-up time between blood collection and PDAC diagnosis in the index cases (Fig. 1). For the shortest follow-up times (≤2 years), miR-10b, miR-21-5p, miR-30c and miR-106b were statistically significantly higher in cases than their matched controls (Fig. 1 and Table 1). For longer follow-up times (>5–8 years and >8 years), miR-10a, miR-10b, miR-30c and miR-106b were statistically significantly higher in cases than controls. Overall, for longer follow-up times (>8 years), and for all 225 cases and 225 matched controls (≤12 years), the most consistent and statistically significant differences were observed for miR-10a, miR-10b and miR-30c (Fig. 1 and Table 1). There was no evidence for longitudinal differences in miR expression levels among controls and among cases by follow-up time between blood collection and PDAC diagnosis (Kruskal-Wallis test $p$-values >0.6, data not in tables), except for miR-106b among cases ($p$ values =0.01); however, there was no clear monotonic trend between median miR-106b expression level and follow-up time, and data were relatively sparse, especially at shorter follow-up times.

Multivariable-adjusted, conditional logistic regression models for the association between relative miR expression level and PDAC risk were evaluated for various overlapping follow-up times (i.e., ≤2, ≤5, ≤8 and ≤12 years) between blood collection and PDAC diagnosis (Table 2). When evaluating all 225 cases and 225 matched controls (follow-up time ≤12 years), ORs for the association between miR expression level (as continuous variables) and PDAC risk were statistically significant for six miRs (miR-10a, -10b, -21-5p, -30c, -155 and -212) (Table 2). When these associations were evaluated according to shorter follow-up times between blood collection and PDAC diagnosis (≤2, ≤5, ≤8 years) ORs remained elevated for miR-10b, miR-21-5p and miR-30c. The magnitude of ORs was somewhat stronger with shorter follow-up times for miR-21-5p, and to a lesser extent for miR-30c (Table 2).

Adjusted associations between miR expression levels and PDAC risk were also evaluated according to the nonoverlapping follow-up times that were examined in Figure 1 and Table 1 (>2–5, >5–8 and >8 years). For follow-up times >2–5 years (29 cases, 29 controls), because of small sample size, these estimates are not shown. For >5–8 years (53 cases, 53 controls), positive associations were observed for miR-10b (1.76, 95% CI = 1.05–2.96) and miR-30c (OR = 2.01, 95% CI = 1.11–3.62). Finally, for longer follow-up times >8 years (108 cases, 108 controls), positive associations were observed for miR-10a (OR = 2.05, 95% CI = 1.36–3.10), miR-10b

miR-10b (OR = 2.35, 95% CI = 1.62–3.92), miR-30c (OR = 2.35, 95% CI = 1.54–3.57) and miR-106b (OR = 2.61, 95% CI = 1.57–4.34). Associations for follow-up ≤2 years were statistically significant for miR-10b, miR-21-5p and miR-30c (listed in Table 2).

Fully adjusted conditional ORs for the association between a miR expression score and PDAC risk showed statistically significant evidence for higher risks with higher numbers of over-expressed miRs, as well as a monotonic dose–response trend (Table 3). Adjusted and un-adjusted analyses of the miR score by shorter follow-up (≤8, ≤5 years) produced risk estimates that were unstable because of sparse data (not shown).

Fully adjusted, conditional logistic regression models for each miR analyzed in quintiles and as continuous variables showed monotonic dose–response trends for miR-10a, miR-10b, miR-21-5p and miR-30c (Supporting Information Table S5). A single, mutually adjusted, conditional logistic regression model including all 8 miRs and covariates was also evaluated, and showed higher PDAC risk with increasing over-expression of miR-10a, miR-10b and miR-30c, but with less of a tendency for monotonic trends (Supporting Information Table S5).

In exploratory analyses, we evaluated joint ORs for two-fold (miR-miR) interactions. Although some joint effect ORs

Figure 1. microRNA expression levels by case/control status, and by follow-up time between blood collection and PDAC diagnosis. X-axis represents follow-up time between blood collection and PDAC diagnosis, and the Y-axis represents relative miRNA expression level.
Table 1. microRNA expression levels in PDAC cases and controls according to follow-up time between blood collection and pancreatic cancer diagnosis.

<table>
<thead>
<tr>
<th>microRNA</th>
<th>&lt;2 yr Cases n = 35</th>
<th>&lt;2 yr Controls n = 35</th>
<th>&lt;2–5 yr Cases n = 29</th>
<th>&lt;2–5 yr Controls n = 29</th>
<th>&lt;5–8 yr Cases n = 53</th>
<th>&lt;5–8 yr Controls n = 53</th>
<th>&lt;8 yr Cases n = 108</th>
<th>&lt;8 yr Controls n = 108</th>
<th>&lt;12 yr Cases n = 225</th>
<th>&lt;12 yr Controls n = 225</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>miR-10a</td>
<td>0.78</td>
<td>0.59</td>
<td>0.15</td>
<td>0.72</td>
<td>0.67</td>
<td>0.62</td>
<td>0.76</td>
<td>0.58</td>
<td>0.034</td>
<td>0.0055</td>
<td>0.0002</td>
</tr>
<tr>
<td>miR-10b</td>
<td>1.15</td>
<td>0.80</td>
<td>0.12</td>
<td>0.80</td>
<td>0.57</td>
<td>0.37</td>
<td>1.26</td>
<td>0.93</td>
<td>0.0032</td>
<td>&lt;0.0001</td>
<td>0.90</td>
</tr>
<tr>
<td>miR-21–3p</td>
<td>0.71</td>
<td>0.88</td>
<td>0.17</td>
<td>0.77</td>
<td>0.64</td>
<td>0.17</td>
<td>0.69</td>
<td>0.68</td>
<td>0.076</td>
<td>0.0009</td>
<td>0.23</td>
</tr>
<tr>
<td>miR-21–5p</td>
<td>1.42</td>
<td>0.89</td>
<td>0.044</td>
<td>1.25</td>
<td>1.08</td>
<td>0.062</td>
<td>1.21</td>
<td>1.06</td>
<td>0.11</td>
<td>&lt;0.0001</td>
<td>0.0009</td>
</tr>
<tr>
<td>miR-30c</td>
<td>1.49</td>
<td>1.05</td>
<td>0.0025</td>
<td>1.10</td>
<td>0.83</td>
<td>0.14</td>
<td>1.41</td>
<td>0.85</td>
<td>0.0045</td>
<td>&lt;0.0001</td>
<td>0.85</td>
</tr>
<tr>
<td>miR-106b</td>
<td>0.99</td>
<td>0.83</td>
<td>0.012</td>
<td>0.86</td>
<td>0.79</td>
<td>0.099</td>
<td>0.80</td>
<td>0.80</td>
<td>0.0006</td>
<td>0.83</td>
<td>0.85</td>
</tr>
<tr>
<td>miR-155</td>
<td>0.63</td>
<td>0.55</td>
<td>0.03</td>
<td>0.79</td>
<td>0.68</td>
<td>0.33</td>
<td>0.75</td>
<td>0.73</td>
<td>0.031</td>
<td>0.74</td>
<td>0.69</td>
</tr>
<tr>
<td>miR-212</td>
<td>0.58</td>
<td>0.62</td>
<td>0.02</td>
<td>0.73</td>
<td>0.53</td>
<td>0.32</td>
<td>0.93</td>
<td>0.64</td>
<td>0.022</td>
<td>0.69</td>
<td>0.61</td>
</tr>
</tbody>
</table>

QR, quartile range. 1Wilcoxon rank-sum test on median untransformed miR expression levels (cases vs. matched controls).

Discussion

Our case-control study evaluated the relative expression of a targeted panel of eight microRNAs in plasma samples of patients with PDAC who had not yet been diagnosed with PDAC. We found that miR-21–5p, miR-21–3p, miR-30c, miR-10b, miR-30a, and miR-10b were over-expressed in PDAC cases compared to controls. The expression of these microRNAs was higher in cases than in controls, even when adjusted for smoking and PDAC risk factors.

Both miR-10a and -10b have been implicated in PDAC development and along with miR-221 may contribute to PDAC risk in individuals with two or more risk factors. miR-21 has been shown to be over-expressed in PDAC compared to controls.

The performance of the miR-21 diagnostic model (area under the curve 0.65–0.70, p ≤ 0.05) was also higher than for the miR-30a and miR-30c models.

The results suggest that the expression of miR-21–5p and miR-21–3p may contribute to PDAC development and along with miR-21 may contribute to PDAC risk in individuals with two or more risk factors.
PDAC’s propensity to metastasize. While it is not known whether miR-21-3p has a role in PDAC, this miR may contribute to cancer progression and cisplatin resistance. Moreover, miR-30c is upregulated by the EGF receptor and may contribute to PDAC progression, whereas miR-106b and miR-212 target Rb, thereby potentially inducing loss of negative growth constraints. Finally, miR-155 is expressed in PDAC, but its role in this cancer remains to be delineated. Nonetheless, miR-155 has been shown to enhance cancer cell proliferation and angiogenesis, and is likely to have an important role in PDAC.

It is well established that PDAC is a heterogeneous cancer, and recent studies suggest that there are four major molecular subtypes. However, there is no evidence that each of these different molecular subtypes is associated with a unique miRNA profile. Conversely, it is possible that the kinetics of miRNA expression occurs at different time-points in these subtypes. Moreover, while PDAC may develop and progress over 15 to 17 years prior to developing distant metastases, it is now appreciated that a subgroup of patients have much more rapid PDAC progression because of concurrent and rapid acquisition of mutations, which further complicates the delineation of specific time points for the appearance of diagnostic miRNAs in the circulation.

Importantly, a patient who underwent total pancreatectomy and islet autotransplantation (TPIAT), meaning that he received an infusion of his own islets into the venous portal system, developed hepatic metastases 10 months following the infusion that were histologically confirmed to be metastatic PDAC. The patient’s pancreas that had been removed over 15 to 17 years prior to developing distant metastases, is likely to have an important role in PDAC.

Table 2. Odds ratios (OR) for PDAC risk per doubling of log2-transformed microRNA expression level, by follow-up time between blood collection and PDAC diagnosis.

<table>
<thead>
<tr>
<th>Follow-up time</th>
<th>Cases</th>
<th>Controls</th>
<th>OR1 (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>n, Cases microRNA</td>
<td>≤2 years</td>
<td>35 OR1 (95% CI)</td>
<td>≤5 years</td>
</tr>
<tr>
<td>mir-10a</td>
<td>2.71 (0.96–7.63)</td>
<td>1.77 (1.01–3.09)</td>
<td>1.70 (1.18–2.45)</td>
</tr>
<tr>
<td>mir-10b</td>
<td>1.82 (1.03–3.23)</td>
<td>1.60 (1.08–2.39)</td>
<td>1.66 (1.24–2.25)</td>
</tr>
<tr>
<td>mir-21–3p</td>
<td>1.30 (0.66–2.54)</td>
<td>1.31 (0.81–2.14)</td>
<td>1.21 (0.86–1.69)</td>
</tr>
<tr>
<td>mir-21–5p</td>
<td>4.02 (1.14–14.1)</td>
<td>3.86 (1.64–9.10)</td>
<td>2.09 (1.26–3.46)</td>
</tr>
<tr>
<td>mir-30c</td>
<td>2.64 (1.11–6.27)</td>
<td>2.15 (1.29–3.59)</td>
<td>1.98 (1.39–2.82)</td>
</tr>
<tr>
<td>mir-106b</td>
<td>1.38 (0.38–4.99)</td>
<td>1.63 (0.75–3.52)</td>
<td>0.91 (0.54–1.54)</td>
</tr>
<tr>
<td>mir-155</td>
<td>1.75 (0.80–3.80)</td>
<td>1.44 (0.90–2.31)</td>
<td>1.33 (0.97–1.83)</td>
</tr>
<tr>
<td>mir-212</td>
<td>1.08 (0.72–1.62)</td>
<td>1.05 (0.79–1.41)</td>
<td>1.13 (0.90–1.41)</td>
</tr>
</tbody>
</table>

1 Separate conditional logistic regression models for each log2-transformed microRNA, adjusted for age, smoking status, and alcohol intake at baseline.

Table 3. Odds ratios (OR) for PDAC risk per microRNA expression score.

| Trend test p-value | 0.0006 |

1 Number of miRs (miR-10a, -10b, -21-5p, -30c, -106b, -155, -212) with expression above 75th percentile in EPIC controls.

2 Conditional logistic regression model adjusted for age, smoking intensity, alcohol intake, education, BMI, and physical activity at baseline.

To our knowledge, this is the first published evaluation of miR expression and subsequent pancreatectomy in a prospective cohort study. In previous studies, blood- and tissue-based miR expression signatures have been reported to differentiate various pancreatic lesions (precursor lesions, chronic pancreatitis, neuroendocrine tumors and PDAC) from one another, and from normal tissue, but previous studies have been retrospective, limited in sample size and the specific miRs evaluated and reported have not been consistent across studies [as previously reviewed].
Although our analyses provide support for the hypothesis that the over-expression of miRs may help to distinguish some individuals who may be at a higher risk of eventually developing PDAC, based on area under the ROC curves, our miR biomarker results fall short of being considered “clinically useful” tests for early PDAC detection or diagnosis in the general population. As pointed out by Wentzensen and Wacholder (2013), “to achieve any meaningful risk stratification for a rare disease would require a biomarker test with very high specificity”.48 In the case of pancreatic cancer, a biomarker test that yields ORs of magnitude of approximately 100 or more would likely be needed to yield high enough specificity (>99.5%) for a test to be considered clinically useful.48,49

Weaknesses of our study include limited information on stage of cancer at diagnosis and diabetes at baseline, the absence of information on pancreatitis and new-onset diabetes prior to diagnosis, lack of information on CA19-9 levels and a limited sample size. We chose a targeted approach in evaluating eight highly promising miRs as biomarkers of PDAC risk. Clearly, there may be other known or as yet undiscovered miRs that may improve risk prediction, or that

Figure 2. ROC curves for each microRNA and the score variable in 225 PDAC cases and 225 matched controls for different follow-up times (≤12 years, ≤8 years, ≤5 years). ROC curves for separate unconditional logistic regression models for each log$_2$-transformed miRNA, adjusted for age, smoking status, and alcohol intake at baseline. The miR score was additionally adjusted for BMI. TPR: True positive rate; FPR: False positive rate%.
may be more appropriate as markers of PDAC prognosis and/or treatment response. Strengths of our study design include the collection of exposure data and blood samples in a well-defined cohort many years before pancreatic cancer diagnosis. Although our study has not yet yielded a clear biomarker for early PDAC diagnosis, it provides useful information with respect to PDAC risk stratification, and, therefore, advances the concept that levels of circulating miRs may have the potential to be useful for diagnosing PDAC. Additional work needs to be done to find a set of biomarkers or biochemical tests that can reliably identify high-risk individuals in the general population before their tumors spread beyond the pancreas. The simultaneous consideration of multiple biomarker tests (low abundance mutations, proteins, over-expressed miRs) in biofluids (plasma, secretin-stimulated pancreatic juice, saliva, etc.) and salient clinical information (e.g., family history, smoking history, recent onset diabetes, ABO blood group and recent use of heartburn or acid reflux medications) could be one way forward. Our data thus advocate for the constitution of a much larger cohort together with molecular and imaging studies.

References


