Salivary miRNAs as non-invasive biomarkers of hepatocellular carcinoma: a pilot study

Arshiya Mariam[1], Galen Miller-Atkins[1], Amika Moro[2],
Alejandro I. Rodarte[2], Shirin Siddiqi[2], Lou-Anne Acevedo-Moreno[2],
J. Mark Brown[3,4], Daniela S. Allende[5], Federico Aucejo[2] and
Daniel M. Rotroff[1,6]

[1] Department of Quantitative Health Sciences, Cleveland Clinic, Cleveland, Ohio, United States
[2] Department of General Surgery, Cleveland Clinic, Cleveland, Ohio, United States
[3] Department of Cardiovascular and Metabolic Sciences, Cleveland Clinic, Cleveland, Ohio, United States
[4] Center for Microbiome and Human Health, Cleveland Clinic, Cleveland, Ohio, United States
[5] Department of Pathology, Cleveland Clinic, Cleveland, Ohio, United States
[6] Endocrinology and Metabolism Institute, Cleveland Clinic, Cleveland, Ohio, United States

ABSTRACT

Background: Improved detection of hepatocellular carcinoma (HCC) is needed, as current detection methods, such as alpha fetoprotein (AFP) and ultrasound, suffer from poor sensitivity. MicroRNAs (miRNAs) are small, non-coding RNAs that regulate many cellular functions and impact cancer development and progression. Notably, miRNAs are detectable in saliva and have shown potential as non-invasive biomarkers for a number of cancers including breast, oral, and lung cancers. Here, we present, to our knowledge, the first report of salivary miRNAs in HCC and compare these findings to patients with cirrhosis, a high-risk cohort for HCC.

Methods: We performed small RNA sequencing in 20 patients with HCC and 19 with cirrhosis. Eleven patients with HCC had chronic liver disease, and analyses were performed with these samples combined and stratified by the presence of chronic liver disease. P values were adjusted for multiple comparisons using a false discovery rate (FDR) approach and miRNA with FDR $P < 0.05$ were considered statistically significant. Differential expression of salivary miRNAs was compared to a previously published report of miRNAs in liver tissue of patients with HCC vs cirrhosis. Support vector machines and leave-one-out cross-validation were performed to determine if salivary miRNAs have predictive potential for detecting HCC.

Results: A total of 4,565 precursor and mature miRNAs were detected in saliva and 365 were significantly different between those with HCC compared to cirrhosis (FDR $P < 0.05$). Interestingly, 283 of these miRNAs were significantly downregulated in patients with HCC. Machine-learning identified a combination of 10 miRNAs and covariates that accurately classified patients with HCC (AUC = 0.87). In addition, we identified three miRNAs that were differentially expressed in HCC saliva samples and in a previously published study of miRNAs in HCC tissue compared to cirrhotic liver tissue.

Conclusions: This study demonstrates, for the first time, that miRNAs relevant to HCC are detectable in saliva, that salivary miRNA signatures show potential to be

highly sensitive and specific non-invasive biomarkers of HCC, and that additional studies utilizing larger cohorts are needed.

INTRODUCTION

Liver cancer is the most rapidly increasing cancer among men and women in the United States and is estimated to have resulted in 31,780 deaths in 2019 (*American Cancer Society, 2019*). Hepatocellular carcinoma (HCC) accounts for 80% of all primary liver cancers, and the global incidence of HCC is expected to increase to 78 million by 2030 (*Petrick et al., 2016*) due to increasing nonalcoholic steatohepatitis (NASH), hepatitis C, and excessive alcohol consumption. Early detection of HCC has shown to improve reception of curative therapy and overall survival (*Singal, Pillai & Tiro, 2014*). However, current HCC serum biomarkers, such as alpha fetoprotein (AFP) lack prognostic and diagnostic value (39–65% sensitivity) (*Daniele et al., 2004*). Although resection is the treatment of choice for many patients, as many as 54% will recur within 2 years, and ~70% within 5 years (*Forner et al., 2006*; *Tabrizian et al., 2015*). As early detection is associated with improved prognosis and overall survival, reliable biomarkers are needed to detect tumor in early stages of development.

MicroRNAs (miRNAs) are small, non-coding RNA molecules that regulate gene expression by translational inhibition or mRNA degradation. Although the majority of miRNA regulation occurs intracellularly, in some cases select miRNAs are found in the circulation and have been associated with a plethora of human diseases. Several studies have shown miRNA expression is associated with cancer (*Eslam, Valenti & Romeo, 2018*), NASH/NAFLD and alcoholic hepatitis (*Blaya et al., 2016*; *Eslam, Valenti & Romeo, 2018*), and patient survival (*Pineau et al., 2010*; *Dongiovanni et al., 2018*). In addition, differentially expressed miRNA in blood have been detected in HCC (*Murakami et al., 2012*; *Qi et al., 2013*). However, it is not known if detectable alterations in salivary miRNAs are present in patients with HCC. Previously, miRNA abundance in saliva has been associated with oral, head and neck, lung, and breast cancers (*Zhang et al., 2010*, *2012*; *Elashoff et al., 2012*; *Salazar, Calvopiña & Punyadeera, 2014*; *Salazar et al., 2014*; *Wan et al., 2017*). Saliva is an attractive biospecimen for biomarkers because (1) it is non-invasive, (2) less expensive than more invasive biospecimens, (3) can be collected without requiring a patient to go to the doctor's office, and (4) saliva collection tubes can now stabilize samples at room temperature for months (*Nunes et al., 2012*; *Lim et al., 2016*), allowing for the possibility that patients may be able to submit their sample through the mail.

Here, we conducted a pilot-study to evaluate if miRNAs are differentially expressed in the saliva of patients with HCC compared to the saliva of patients with cirrhosis. Furthermore, we evaluate how these miRNAs compare to previously reported tissue- and

blood-based miRNAs in HCC, and determine the potential for saliva miRNAs to serve as a non-invasive, diagnostic biomarker for HCC.

MATERIALS AND METHODS

Patient recruitment and sample collection

Saliva samples were collected from 20 individuals with HCC and 19 from individuals with cirrhosis seen at the Cleveland Clinic (Cleveland, OH). Participants were adult patients (>18 years of age) who underwent liver transplantation for HCC, surgical resection for liver tumors or liver biopsy. Some patients with HCC had previously received treatment but all had active HCC at the time of collection. Prior treatments and clinical staging for patients with HCC can be found in Table S1. A description of the cohort is provided in Table 1, including the frequencies of chronic liver disease (CLD) between the two groups. Twelve out of the twenty patients with HCC had CLD, and analyses were performed with these samples combined and stratified by liver disease, as described below. Initial disease diagnoses were made from a combination of clinical presentation, imaging and laboratory techniques. Subsequently, these diagnoses underwent a secondary confirmatory pathological diagnosis. All participants provided written informed consent and the study was approved by the Cleveland Clinic IRB (IRB #10-347).

Small RNA-seq extraction and sequencing

Small RNA library preps were prepared using the QIAseq miRNA Library Kit (QIAGEN, Hilden, Germany). Adapters are first ligated sequentially to the 3′ and 5′ end of the miRNAs followed by cDNA synthesis with UMI assignment, cDNA cleanup, library amplification and final library cleanup. All protocol steps were followed based on the use of the miRNeasy Serum/Plasma kit used upstream for purification of RNA, which has been shown to be effective for saliva samples (Zahran et al., 2015). The recommended starting amount of total RNA is 5 μl of the RNA eluate when 200 μl of sample has been processed using the miRNAeasy Serum/Plasma Kit. Adapter dilutions throughout the protocol followed the serum/plasma recommendations. Cycles of library amplification followed that of a 10 ng input sample with 19 total cycles, consistent with manufacturer recommendations. Two RNA control samples (Human XpressRef RNA, QIAGEN, Hilden, Germany) with an input of 10 ng RNA were processed alongside the saliva samples. Final libraries were validated by Qubit Fluorometer (Invitrogen, Waltham, MA, USA) and Fragment Analyzer (Agilent Technologies, Inc, Santa Clara, CA, USA), and quantified *via* qPCR using NEBNext Library Quant Kit for Illumina (New England BioLabs, Inc, Ipswich, MA, USA). Pooled libraries were diluted, denatured and loaded onto the Illumina NextSeq 550 System, following the NextSeq User Guide. All 42 libraries were sequenced on one NextSeq High Output flow cell, single read 75 cycle run. FastQ files were developed for downstream analysis.

Data processing

FastQ files were first evaluated for read quality, using FastQC v0.11 (Andrews, 2010). Individual reads were then trimmed for phred quality scores (Q > 20) and presence of

Table 1 Summary statistics for the study cohort.

Characteristic		Cirrhosis without HCC	HCC
Total (*N*)		19	20
Mean age (min-max)		57.2 (33–80)	67.9 (53–89)
Sex			
	Male (%)	9 (47%)	14 (70%)
	Female (%)	10 (53%)	6 (30%)
Race			
	Caucasian (%)	18 (95%)	10 (50%)
	Black (%)	0 (0%)	0 (0%)
	Hispanic (%)	1 (5%)	1 (5%)
	Unspecified (%)	18 (95%)	9 (45%)
Mean BMI (min-max)		33.12 (21.07–57.96)	29.44 (19.53–41.8)
Chronic liver disease		19 (100%)	12 (60%)
Fibrosis (%)		0 (0%)	2 (10%)
Cirrhosis (%)		19 (100%)	10 (50%)
NASH (%)		7 (37%)	5 (25%)
EtOH[a] (%)		7 (37%)	7 (35%)
HCV[b] (%)		0 (0%)	7 (35%)
HBV[c] (%)		0 (0%)	2 (10%)
Primary biliary cholangitis (%)		2 (11%)	0 (0%)
Primary sclerosing cholangitis (%)		1 (5%)	0 (0%)
Autoimmune hepatitis (%)		1 (5)	0 (0%)
Other (%)		0 (0%)	0 (0%)
Child-pugh score			
	5–6	8	8
	7–9	10	1
	10–15	1	1
Diabetes mellitus (%)		9 (47%)	10 (50%)
Hypertension (%)		6 (32%)	16 (80%)
Coronary artery disease (%)		2 (11%)	7 (35%)
Hyperlipidemia (%)		5 (26%)	14 (70%)
Psychiatric disorder (%)		3 (16%)	6 (30%)
Other cancer (%)		1 (5%)	3 (15%)
COPD[d]/Asthma/OSA[e]		3 (16%)	6 (30%)
Thyroid		5 (26%)	0 (0%)
Other PH[f]		0 (0%)	0 (0%)
Ascites		8 (42%)	1 (5%)
Encephalopathy		8 (42%)	0 (0%)
Mean hemoglobin (std.err)		10.68 (0.7)	12.80 (0.5)
Mean platelets (std.err)		116 (16.2)	210 (19.3)
Mean ALP[g] (std.err)		156.45 (29.5)	162.85 (40.0)
Mean AST[h] (std.err)		54.53 (8.4)	56.30 (7.4)
Mean ALT[i] (std.err)		34.79 (6.8)	52.20 (7.0)

Table 1 (continued)		
Characteristic	Cirrhosis without HCC	HCC
Mean bilirubin (std.err)	1.99 (0.4)	0.91 (0.2)
Mean albumin (std.err)	3.42 (0.1)	3.80 (0.1)
Mean INR[j] (std.err)	1.26 (0.06)	1.14 (0.05)
Mean creatinine (std.err)	1.10 (0.1)	1.18 (0.2)

Notes:
[a] Alcoholic hepatitis
[b] Hepatitis C
[c] Hepatitis B
[d] Chronic Obstructive Pulmonary Disease
[e] Obstructive Sleep Apnea
[f] Pulmonary Hypertension
[g] Alkaline Phosphatase
[h] Aspartate Aminotransferase
[i] Alanine Transaminase
[j] International Normalized Ratio.

adapters, using fastp v0.19 software (*Chen et al., 2018*). Processed FastQ files were then aligned to the human genome (hg38) (*Schneider et al., 2017*) and annotated using known miRNA sequences from miRBase v22 (*Griffiths-Jones, 2010*). Alignments and miRNA counts were performed using the Rsubread package v2.2, following the authors' guidelines for miRNA. Mature miRNA and miRNA hairpin precursors were both used for alignment, and are designated "hsa-miR" and "hsa-mir", respectively. Data files, including miRNA counts are located in Files S1 and S2, and R code for the analysis described below can be found at: https://github.com/rotroff-lab/salivary-miRNA-HCC. Raw sequences can be found at: https://www.ncbi.nlm.nih.gov/bioproject/PRJNA755300.

Differential expression analysis

miRNA expression was compared between HCC and cirrhosis samples using the R package DESeq2 v1.28 (*Love, Huber & Anders, 2014*). First, raw read counts were normalized to account for sequencing depth, gene length, and RNA composition. Wald's test was then performed on the normalized counts to identify any differentially expressed miRNA, using the cirrhosis samples as reference. All association tests included age, sex, race, body mass index and smoking status as model covariates to prevent confounding. In addition, patients were stratified to perform comparisons between those with CLD and the cirrhosis reference samples to distinguish miRNA associations specific to HCC *vs* those that may be due to impaired liver function. Log$_2$ fold changes (logFC) were estimated using empirical Bayes procedure and implemented with the R package *apeglm* v1.10 (*Zhu, Ibrahim & Love, 2019*). All *P* values were false discovery rate (FDR) corrected (*Benjamini & Hochberg, 1995*). Any miRNA with an FDR *P* < 0.05 was considered significantly differentially expressed. The level of confidence for significant miRNAs were categorized based on read counts per million (CPM) into low, medium or high categories. miRNAs with <1 CPM in >20% of the samples were considered to be low confidence. miRNAs were classified as medium confidence if at least 20% of the samples had CPM ≥ 1 and high confidence if at least 20% of the samples had CPM ≥ 10. Only miRNAs meeting the criteria for medium or high confidence were considered for inclusion in the predictive models, as

described below. Pathway analysis was performed using QIAGEN Ingenuity Pathway Analysis (IPA) software (*Krämer et al., 2014*). Here, differentially expressed miRNA (FDR $P < 0.05$) were mapped to mRNA and tested for overrepresentation in cancer and disease-specific pathways using the IPA knowledgebase. The analysis was performed (1) restricted to only cancer-related pathways and (2) cancer-related and disease-specific pathways combined.

Comparison to tissue-based miRNA profiles

Differentially abundant miRNAs were compared to microarray expression data from *Martinez-Quetglas et al. (2016)*, GEO Accession Number: GSE74618 (*Martinez-Quetglas et al., 2016*). These data included miRNA expression data from 218 samples from human HCC tumors and 10 samples from cirrhotic non-tumoral tissue. Differential expression was conducted on the normalized expression values between HCC and cirrhosis samples, using the GEO2R tool available from the GEO database (*Barrett et al., 2012*). Overlapping miRNAs detected in both GSE74618 and the salivary miRNA data were compared and were adjusted using a FDR approach, as described above.

Predictive modeling for biomarker development

Model development was performed using the statistical software, R (*Team, 2020*). Significant miRNAs (FDR $P < 0.05$) were assessed for their ability to differentiate HCC and cirrhosis samples. Here, we used support vector machine (SVM) models with a radial basis function using the R *caret* package (*Kuhn, 2008*). Because some patients with HCC had no evidence of CLD, we also performed modeling on the subset of patients with CLD to compare with the cirrhosis group. Hyperparameters for both models with and without covariates are listed in Table S2. miRNA expression was mean centered and scaled by the standard deviation prior to modeling. Leave-one-out cross-validation was used to limit model over-fitting. Recursive feature selection with leave-one-out cross-validation was used to characterize feature importance and 10 miRNAs were selected for inclusion in the final models (*Kuhn, 2008*). Model performance was assessed based on the AUC, sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and balanced accuracy. PPV is the ratio of true positives out of all identified positives, whereas NPV is the ratio of true negatives out of all identified negatives.

RESULTS

Differential expression of miRNAs in the saliva of HCC patients

A total of 4,565 miRNAs were detected and 365 miRNAs were significantly differentially expressed between HCC and cirrhosis samples and met the threshold for medium or high confidence (FDR $P < 0.05$), demonstrating a broad shift in miRNA profiles (Table S3).

The majority of the differentially expressed miRNAs ($N = 283$) were downregulated (logFC < 0) in the HCC samples compared to cirrhosis samples, and 50 of those had a logFC < −2 (Fig. 1A). The five most significantly differentially expressed miRNAs that also met the threshold for high confidence were: has-mir-3198-2 (logFC = −5.00, FDR $P = 6.81 \times 10^{-6}$), hsa-miR-3198 (logFC = −5.01, logFC = 6.80×10^{-6}), hsa-mir-1246
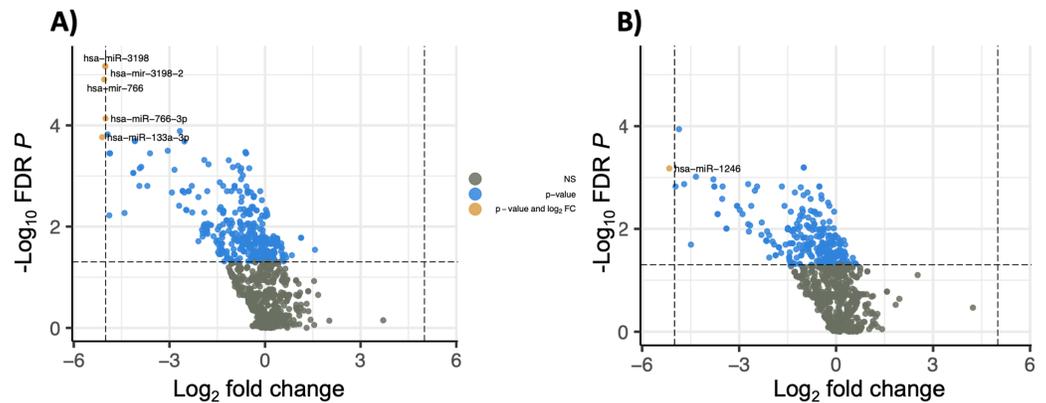
Figure 1 Volcano plots of differential salivary miRNA expression. Volcano plots showing the unadjusted $\log_2$ fold change and log FDR P value for (A) salivary miRNAs in all patients with HCC ($N = 20$) compared to patients with cirrhosis only ($N = 19$), and (B) patients with HCC and chronic liver disease ($N = 11$) compared to patients with cirrhosis only ($N = 19$). Salivary miRNAs with a corresponding log fold change less than -5 and FDR $P < 5 \times 10^{-6}$ are annotated using color. Mature miRNA and hairpin precursors are referred as "miR" and "mir", respectively.

Full-size 🔳 DOI: 10.7717/peerj.12715/fig-1

Table 2 Results for 20 most significantly differentially expressed miRNA.

miRNA	Log₂ FC	Log₂ FC SE	*P* value	FDR *P* value
hsa-miR-3198	-5.01	1.50	5.96×10^{-9}	6.81×10^{-6}
hsa-mir-3198-2	-5.01	1.51	6.56×10^{-8}	6.81×10^{-6}
hsa-miR-1246	-6.86	1.46	1.18×10^{-8}	7.98×10^{-6}
hsa-mir-1246	-6.03	1.57	1.54×10^{-8}	7.98×10^{-6}
hsa-mir-3648-2	-7.10	1.94	3.59×10^{-8}	1.24×10^{-5}
hsa-mir-766	-5.05	1.79	3.50×10^{-8}	1.24×10^{-5}
hsa-mir-1290	-8.24	1.84	1.00×10^{-7}	2.61×10^{-5}
hsa-miR-1290	-8.24	1.84	1.00×10^{-7}	2.61×10^{-5}
hsa-mir-3648-1	-8.94	2.61	3.52×10^{-7}	7.31×10^{-5}
hsa-miR-766-3p	-5.00	2.28	3.22×10^{-7}	7.31×10^{-5}
hsa-mir-10401	-9.85	3.36	7.48×10^{-7}	1.29×10^{-4}
hsa-miR-191-3p	-2.67	1.15	7.47×10^{-7}	1.29×10^{-4}
hsa-mir-133a-2	-4.94	1.21	9.70×10^{-7}	1.48×10^{-4}
hsa-miR-133a-3p	-5.11	1.28	9.96×10^{-7}	1.48×10^{-4}
hsa-mir-3615	-4.08	1.44	1.57×10^{-6}	2.04×10^{-4}
hsa-miR-3615	-4.08	1.44	1.57×10^{-6}	2.04×10^{-4}
hsa-miR-454-5p	-2.53	0.94	1.70×10^{-6}	2.08×10^{-4}
hsa-miR-4449	-3.06	2.09	2.75×10^{-6}	3.17×10^{-4}
hsa-miR-642a-3p	-0.61	1.13	3.25×10^{-6}	3.38×10^{-4}
hsa-miR-642b-5p	-0.61	1.13	3.25×10^{-6}	3.38×10^{-4}

(logFC $= -6.03$, FDR $P = 7.98 \times 10^{-6}$), hsa-miR-1246 (logFC $= -6.86$, FDR $P = 7.98 \times 10^{-6}$) and hsa-mir-3648-2 (logFC $= -7.10$, FDR $P = 1.24 \times 10^{-5}$). The 20 most significantly differentially expressed miRNA are shown in Table 2. A heatmap showing the expression
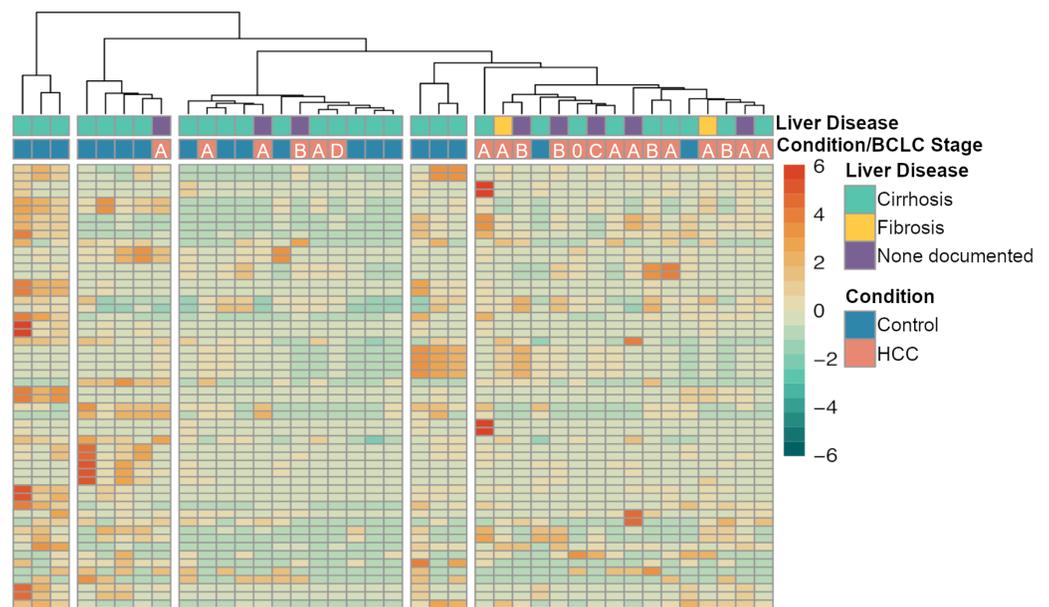
Figure 2 Heatmap showing the expression of the significant miRNA in each sample (FDR *P* < 0.05 and absolute log$_2$ fold change > 1). Heatmap showing the expression of the significant miRNA in each sample (FDR *P* < 0.001 and absolute log$_2$ fold change > 2). Clustering was performed using Ward's D method and Manhattan distance. The dendrogram of samples (columns) was cut to create four clusters, where a distinct cluster with 14 out of 15 samples consisted of HCC samples was observed. The clustering appears to be largely driven by HCC status rather than the presence of chronic liver diseases, such as cirrhosis or fibrosis. The Barcelona Clinic Liver Cancer (BCLC) stage is shown in the annotation bar.

Full-size 🖼 DOI: 10.7717/peerj.12715/fig-2

of the significant miRNAs (FDR *P* < 0.05 and absolute |logFC| > 2) is shown in Fig. 2. Interestingly, a cluster of majority HCC patients was identified (*N* = 14 out of 16, 87.5%). The other four clusters had cirrhosis patients in majority. The number of cirrhosis patients in the four clusters was three (100%), three (75%), four (80%) and seven (58%), respectively. Notably, a cluster of three individuals from the cirrhosis group clustered together with elevated miRNA expression (Fig. 2). The Child-Pugh scores for these individuals were between five and seven, consistent with the rest of the control group, and no demographic or physical characteristics were identified that could explain the observed pattern (Fig. 2).

Pathway analysis revealed 96 (27%) differentially expressed miRNAs that were represented in at least one significant pathway in the IPA metabolic and signaling pathways knowledgebase (Table S4). Approximately a third of these miRNAs were associated with cancer-related diseases and pathways (*N* = 31 of 96) (Table S4). When the analysis was limited to cancer signaling pathways, the top three diseases were "early-stage invasive cervical squamous cell carcinoma" (miRNA *N* = 12, FDR *P* = 8.04 × 10^{-13}), "hypopharyngeal squamous cell carcinoma" (miRNA *N* = 12, FDR *P* = 1.51 × 10^{-11}), and "early stage solid tumor" (miRNA *N* = 13, FDR *P* = 4.59 × 10^{-8}). The cancer pathway with the greatest number of overlapping miRNAs differentially expressed in this cohort was "mammary tumor" (miRNA *N* = 35, FDR *P* = 2.26 × 10^{-2}).
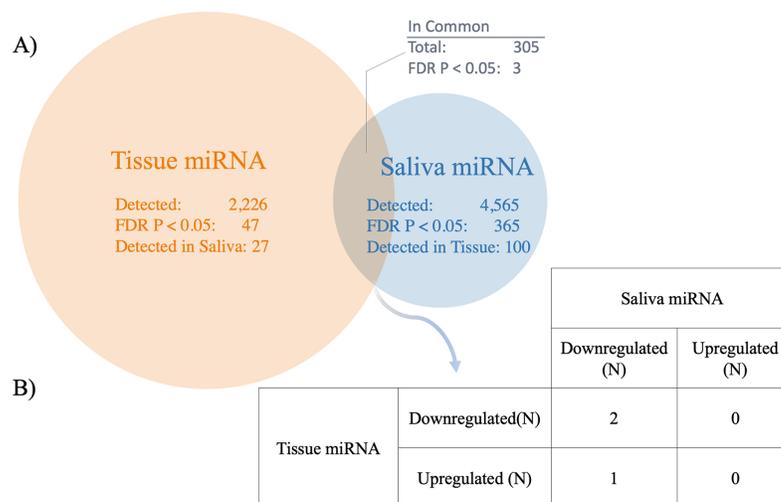
Figure 3 Comparison between salivary miRNAs and previously reported tissue-based miRNA expression. (A) Venn diagram showing the overlap of miRNAs detected in tissue (25) compared to saliva in patients with hepatocellular carcinoma (HCC) *vs* cirrhosis. In addition, the overlap between miRNA determined to be statistically significantly different between patients with HCC and cirrhosis are shown (FDR $P < 0.05$). (B) The direction of association between saliva and tissue miRNA detected to be statistically significantly different between patients with HCC and cirrhosis across both biospecimens (FDR $P < 0.05$). Full-size ◚ DOI: 10.7717/peerj.12715/fig-3

Comparison between patients with and without chronic liver disease (CLD)

Twelve out of 20 patient samples with HCC also had CLD, and these samples were compared with cirrhosis control samples. This resulted in 281 differentially expressed miRNAs, of which 248 (88%) were also significant in the analysis between all HCC samples and those with cirrhosis only (Table S5), suggesting that these differentially expressed miRNAs are likely to be due to the presence of HCC rather than the presence or lack of CLD. Out of 281 differentially expressed miRNA, 240 (85%) were downregulated (logFC < 0) in the HCC samples with CLD compared to only those with cirrhosis, and 49 of these had a logFC < −2 (Fig. 1B). The five most significantly differentially expressed miRNAs were: hsa-mir-122 (logFC = −6.65, FDR $P = 7.54 \times 10^{-5}$), hsa-mir-122b (logFC = −6.64, FDR $P = 7.54 \times 10^{-5}$), hsa-miR-122-5p (logFC = −6.64, FDR $P = 7.54 \times 10^{-5}$), hsa-mir-122b-3p (logFC = −6.64, FDR $P = 7.54 \times 10^{-5}$) and hsa-mir-603 (logFC = −4.86, FDR $P = 1.14 \times 10^{-4}$). These top five miRNAs were also considered to be of high confidence based on the CPM criterion. Pathway analysis after stratifying patients based on CLD status resulted in the same top rankings of cancer-related and disease-specific pathways (Table S6).

Comparison to tissue-based miRNA

Three significant miRNAs detected in saliva were also found to be differentially expressed in HCC tissue samples compared to cirrhotic liver tissue samples (FDR $P < 0.05$) (GSE74618) (Fig. 3 and Table S7). Forty-seven miRNAs were significantly different between HCC and cirrhosis tissue samples (FDR $P < 0.05$), 27 of which were evaluated in

Table 3 Results for 20 most significant common miRNAs in saliva compared to tissue samples from *Martinez-Quetglas et al. (2016)*.

Core miRNA	HCC *vs* cirrhosis liver tissue[a]			HCC *vs* cirrhosis saliva	
	miRNA[b]	logFC	FDR *P* value	logFC	FDR *P* value
hsa-mir-1246	hp_hsa-mir-1246_st	0.051	7.73×10^{-1}	−6.03	7.98×10^{-6}
hsa-mir-1246	hsa-miR-1246_st	0.677	6.24×10^{-1}	−6.03	7.98×10^{-6}
hsa-mir-766	hsa-miR-766_st	0.026	9.53×10^{-1}	−5.05	1.24×10^{-5}
hsa-mir-766	hp_hsa-mir-766_st	−0.019	9.12×10^{-1}	−5.05	1.24×10^{-5}
hsa-mir-1290	hsa-miR-1290_st	0.238	8.01×10^{-1}	−8.24	2.61×10^{-5}
hsa-mir-133a-2	hp_hsa-mir-133a-2_s_st	0.029	8.53×10^{-1}	−4.94	1.50×10^{-4}
hsa-mir-133a-2	hp_hsa-mir-133a-2_st	−0.035	8.52×10^{-1}	−4.94	1.50×10^{-4}
hsa-mir-133a-2	hp_hsa-mir-133a-2_x_st	0.010	9.85×10^{-1}	−4.94	1.50×10^{-4}
hsa-mir-122	hsa-miR-122-star_st	−0.351	6.22×10^{-1}	−4.88	3.57×10^{-4}
hsa-mir-122	hsa-miR-122_st	−0.149	7.94×10^{-1}	−4.88	3.57×10^{-4}
hsa-mir-122	hp_hsa-mir-122_st	−0.038	7.79×1^{-1}	−4.88	3.57×10^{-4}
hsa-mir-1180	hsa-miR-1180_st	0.809	8.5×10^{-2}	−1.14	4.88×10^{-4}
hsa-mir-21	hsa-miR-21_st	1.790	6.97×10^{-5}	0.06	8.79×10^{-1}
hsa-mir-548i-2	hp_hsa-mir-548i-2_st	−0.285	2.45×10^{-4}	−4.88	6.01×1^{-3}
hsa-mir-378c	hsa-miR-378c_st	−1.990	1.40×10^{-3}	0.25	2.55×10^{-1}
hsa-mir-125b-1	hp_hsa-mir-125b-1_x_st	−0.286	2.05×10^{-3}	−1.08	8.13×10^{-2}
hsa-mir-106b	hsa-miR-106b_st	0.876	2.98×10^{-3}	−0.95	1.28×10^{-1}
hsa-mir-548i-2	hp_hsa-mir-548i-2_st	−0.285	2.45×10^{-4}	−4.88	6.01×10^{-3}
hsa-mir-548l	hp_hsa-mir-548l_x_st	−0.220	1.67×10^{-2}	−0.37	9.31×10^{-3}
hsa-mir-92b	hsa-miR-92b-star_st	1.170	3.72×10^{-2}	−1.42	1.33×10^{-3}

Notes:
[a] Data obtained from NIH Gene Expression Omnibus (GSE74618), previously published by *Martinez-Quetglas et al. (2016)*.
[b] miRNA identifier reported in *Martinez-Quetglas et al. (2016)*.

the saliva; whereas out of the 365 significant salivary miRNAs, 100 were also evaluated in tissue. However, only three miRNAs were significantly different in both cohorts (FDR *P* < 0.05) (Fig. 3). Out of the three significant miRNA found in each dataset, all were downregulated in saliva, but two were downregulated and one was upregulated in tissue (Fig. 3, Table 3). The miRNA comparisons are provided in Table S7. The significant miRNAs common to both datasets were hsa-mir-92b, hsa-mir-548i-2 and hsa-mir-548l.

Saliva miRNA predictive modeling

The predictive model trained on HCC +/− CLD *vs* cirrhosis optimized with the following 10 miRNAs: hsa-mir-576, hsa-miR-576-5p, hsa-mir-6727, hsa-mir-27b, hsa-miR-27b-3p, hsa-mir-4664, hsa-mir-125a, hsa-miR-6727-5p, hsa-mir-190b and hsa-miR-125a-5p. All of the above miRNAs met the threshold for high confidence except for hsa-mir-4664 and hsa-miR-6727-5p, which met the threshold for medium confidence. The model classified HCC with salivary miRNA alone (AUC = 0.74), and was improved with the inclusion of demographic and lifestyle variables (AUC = 0.87) (Fig. 4, Table 4). The model trained on the subset of patients with both HCC and CLD *vs* cirrhosis optimized with
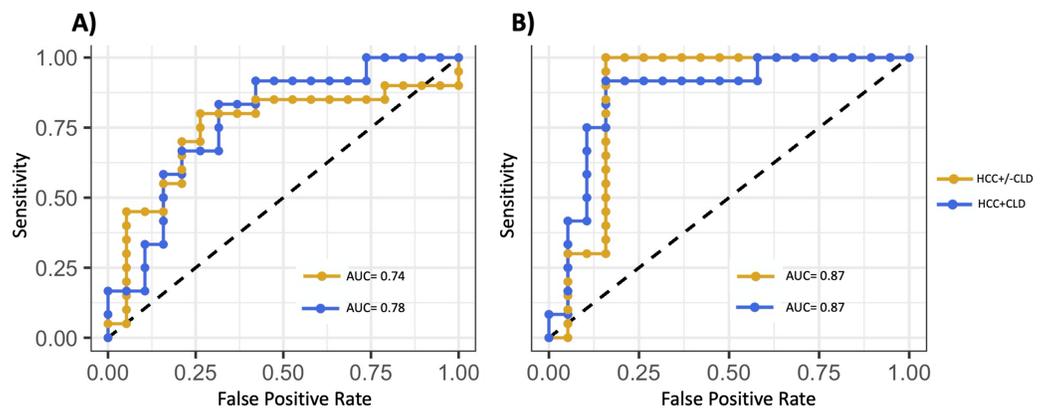
Figure 4 Development of a salivary miRNA signature to predict the presence of hepatocellular carcinoma. (A) ROC curves of the support vector machine (SVM) models fit using the 10 miRNAs selected based on backward selection. (B) ROC curves of the model fit using the 10 selected miRNAs and covariates: age, sex, race, body mass index and smoking status. Gold curves show model performances for hepatocellular carcinoma (HCC) with or without chronic liver disease (CLD) cohort *vs* cirrhosis control samples. Blue curves show model performances for the HCC samples with CLD cohort *vs* cirrhosis control samples. Inclusion of covariates only improved the model performance when restricted to patients with CLD. Full-size ⊡ DOI: 10.7717/peerj.12715/fig-4

Table 4 Accuracy metrics for predicting HCC from cirrhosis.

Covariates included*	Sensitivity	Specificity	Balanced accuracy	PPV	NPV	AUC
Model trained using data from all patients						
miRNA: hsa-mir-576, hsa-miR-576-5p, hsa-mir-6727, hsa-mir-27b, hsa-miR-27b-3p, hsa-mir-4664, hsa-mir-125a, hsa-miR-6727-5p, hsa-mir-190b, hsa-miR-125a-5p						
No	0.80	0.74	0.77	0.75	0.74	0.74
Yes	1.00	0.84	0.92	0.81	0.70	0.87
Model trained using data from patients with chronic liver disease						
miRNA: hsa-miR-1262, hsa-mir-1262, hsa-mir-216a, hsa-mir-484, hsa-mir-30d, hsa-miR-216a-5p, hsa-miR-30d-5p, hsa-miR-484, hsa-mir-10401, hsa-miR-454-3p						
No	0.83	0.68	0.76	0.67	0.68	0.78
Yes	0.92	0.84	0.88	0.82	0.85	0.87

Notes:
PPV, positive predictive value; NPV, negative predictive value; AUC, Area under the receiver operating characteristic curve.
* Covariates include age, sex, race, BMI and smoking.

following miRNAs: hsa-miR-1262, hsa-mir-1262, hsa-mir-216a, hsa-mir-484, hsa-mir-30d, hsa-miR-216a-5p, hsa-miR-30d-5p, hsa-miR-484, hsa-mir-10401 and hsa-miR-454-3p. Five miRNAs above reached the threshold for high confidence: hsa-mir-30d, hsa-miR-30d-5p, hsa-miR-454-3p, hsa-mir-484 and hsa-miR-484. The model performed well, discriminating patients with HCC using salivary miRNA (AUC = 0.78) and with the inclusion of covariates (AUC = 0.87) (Fig. 4). Table 4 shows the AUC, sensitivity, specificity, and balanced accuracy for the predictive models, demonstrating that salivary miRNAs show potential for predicting which patients with CLD are likely to have HCC.

DISCUSSION

Current approaches used to screen patients for HCC lack sensitivity, resulting in too many false negative diagnoses (*Daniele et al., 2004*; *Ayuso et al., 2018*). There is a significant need for improved screening tools for HCC that are also non-invasive and cost-effective. Making screening tools more widely accessible may result in detection of HCC prior to the onset of symptoms, which could improve patient outcomes (*Singal, Pillai & Tiro, 2014*). The relative stability of miRNA (*Fabian, Sonenberg & Filipowicz, 2010*), combined with the non-invasive nature of saliva collection, makes this an attractive option to address these needs. Liver cirrhosis is the primary risk factor for development of HCC, so distinguishing patients with HCC from this cohort of high-risk patients serves as a proof-of-principle for this potential screening approach (*Fattovich et al., 2004*).

Salivary miRNAs have demonstrated biomarker potential for CLD and cancers, such as oral cancer, lung cancer, and breast cancer (*Zhang et al., 2010*, *2012*; *Elashoff et al., 2012*). However, to our knowledge this is the first study to investigate whether miRNAs expressed in saliva can distinguish individuals with HCC. miRNA are important cell signaling molecules affecting functions such as cell proliferation, apoptosis, metastasis, and many others (*Jansson & Lund, 2012*). Although, their importance and biomarker potential are recognized, relatively little is known about specific miRNA functions and targets (*Gebert & MacRae, 2019*).

Here, we observed that a large fraction of the miRNAs detected were differentially expressed in HCC patients *vs* cirrhosis (365 out of 4,565 detected miRNA), and the majority of these (78%, $N = 283$) were significantly down-regulated in patients with HCC *vs* cirrhosis (FDR $P < 0.05$). Importantly, in patients with both CLD and HCC, 248 miRNAs were significantly down-regulated compared to patients with cirrhosis without HCC—of which 85% were also statistically significantly down-regulated when all HCC patients were included (Tables S3, S5). This suggests that most of the miRNAs observed were dysregulated due to the presence of HCC regardless of the presence of CLD. This is further supported in Fig. 2, where the clustering of miRNAs is driven mostly by HCC status rather than the presence of CLD. Other studies have reported significant imbalances between the ratio of upregulated and downregulated salivary miRNAs in gastric cancer and pancreatic cancer (*Machida et al., 2016*; *Li et al., 2018*), and additional research is needed to determine the mechanism responsible for the downregulation of the miRNAs observed in this study. Furthermore, the patients with HCC were heterogeneous with regards to cancer stage, with 13/20 (65%) with very-early or early-stage HCC based on the Barcelona Clinic Liver Cancer (BCLC) stage of 0 or A. However, significant miRNAs clustered HCC samples together regardless of stage (Fig. 2) and the predictive model classified the majority of HCC samples correctly (Table 4) suggesting that the miRNA changes were detected even in early-stage HCC.

Pathway analysis established that salivary miRNAs differentially expressed in HCC are known to regulate genes involved in a range of cancer-related pathways, including those related to breast, and colorectal cancers (*Ren et al., 2014*; *Zheng et al., 2019*; *Peng et al., 2020*; *Fuso et al., 2021*; *Ye et al., 2021*). Furthermore, miRNAs were overrepresented for the

"early stage solid tumor" pathway (FDR $P = 4.59 \times 10^{-8}$). Many of the salivary miRNAs detected in this study have been previously implicated in, or identified as potential biomarkers for, colorectal cancer, HCC, and other cancers. miR-122 has emerged as one of the most promising biomarkers for HCC and treatment response (*Köberle et al., 2013*; *Amr et al., 2017*). It is estimated that approximately 50% of HCC tumors exhibit downregulation of miR-122 compared to surrounding tissue (*Kutay et al., 2006*), and has also been shown to be reduced in the plasma of HCC patients (*Amr et al., 2017*). Here, we found that mir-122 (logFC = −4.88; FDR $P = 3.57 \times 10^{-4}$) and miR-122-5p (logFC = −4.86; FDR $P = 3.57 \times 10^{-4}$) were expressed with high confidence and significantly down-regulated in the saliva of HCC patients compared to those with cirrhosis. Interestingly, studies have shown that the hepatitis C virus (HCV) is dependent on miR-122 for replication, and that the subsequent sequestration of miR-122 by HCV has been linked to decreases in circulating miR-122 (*Luna et al., 2015*). Additional work will be needed to determine if the sequestration of miR-122 in patients with HCV is observed in saliva, and if that is driving the findings observed here. *Zhou et al., 2016*, previously reported that miR-98 was downregulated in HCC tissue compared to adjacent normal tissue, and displayed a negative correlation with tumor size, metastasis, and overall survival through modulation of SALL4. In saliva from patients with HCC, no detectable changes were observed for miR-98 expression. Other studies have also found discordance between salivary miRNA levels and their corresponding intra-tumor expression, although the mechanisms driving these differences are unknown (*Fang et al., 2017*; *Li et al., 2018*). miR-21 has been shown to be significantly upregulated in HCC tissue and commonly used cell lines, and targets PDCD4 and PTEN, resulting in increased cell proliferation (*Meng et al., 2007*; *Frankel et al., 2008*). Interestingly, although miR-21 has been detected in the saliva of colorectal cancer patients, prostate cancer patients, esophageal cancer patients, and head and neck cancer patients (*Rapado-González et al., 2018*), miR-21 was expressed with high confidence, but not significantly different between HCC and cirrhosis patients (logFC = 0.06; FDR $P = 0.88$) (Fig. S2).

The most significant differentially expressed salivary miRNAs were: hsa-mir-3198-2 (logFC = −5.00, FDR $P = 6.81 \times 10^{-6}$), hsa-miR-3198 (logFC = −5.01, logFC = 6.80×10^{-6}), hsa-mir-1246 (logFC = −6.03, FDR $P = 7.98 \times 10^{-6}$), hsa-miR-1246 (logFC = −6.86, FDR $P = 7.98 \times 10^{-6}$) and hsa-mir-3648-2 (logFC = −7.10, FDR $P = 1.24 \times 10^{-5}$). These miRNA have been previously associated with HCC and other cancers (*Du et al., 2009*; *Brunet Vega et al., 2013*; *Nagpal & Kulshreshtha, 2014*; *He et al., 2018*; *You et al., 2018*; *Nuoroozi et al., 2021*). Expression of hsa-miR-3198 has been implicated in HCC recurrence (*Itami-Matsumoto et al., 2019*), epithelial ovarian cancer (*Chong et al., 2015*), metastatic colorectal cancer (*Xiao et al., 2017*) and inhibition of nasopharyngeal carcinoma proliferation (*Yang et al., 2020*). Additionally, upregulation of hsa-miR-1246 is linked with more aggressive HCC (*Huang et al., 2020*), metastasis of non-small-cell lung cancer (*Kim et al., 2016*, p. 12), and chemoresistance in oral carcinomas (*Lin et al., 2018*). Previously, the serum expression of miR-1246 has been proposed as a biomarker for both HCC and early tumor recurrence of HCC (*Chuma et al., 2019*; *Chen et al., 2021*). Expression of hsa-miR-191 in particular has been found to be associated 16 different

cancers, including HCC (*Nagpal & Kulshreshtha, 2014*). Based on these findings, the following miRNAs, miR-122, miR-93, miR-125 and mir-1246 were significantly differentially expressed in patients with HCC, have prior associations with cancer, have putative mechanistic roles in HCC, and therefore present opportunities for future studies in HCC.

Several miRNAs demonstrated discriminatory ability between HCC and cirrhosis, although they were not always the most significantly differentially expressed miRNAs (Table 4). The best model was highly accurate in differentiating HCC and CLD (AUC = 0.87) (Fig. 4B). The addition of demographic and lifestyle variables substantially improved the accuracy of both models (Fig. 4). Although it is unknown whether this accuracy can be achieved in larger prospective cohorts, it does indicate that combinations of salivary miRNAs can successfully discriminate HCC and cirrhosis, warranting additional investigation.

As with any study, there are limitations that should be considered. The patients with HCC were all enrolled post-HCC diagnosis and it is not clear whether these miRNAs would have been detected earlier than with current clinical approaches. The biomarker potential was evaluated using SVM and leave-one-out cross-validation. However, the association results and the predictive model need to be validated in a larger, independent cohort, and ultimately in a prospective cohort, before salivary miRNAs can be considered clinically actionable. This exploratory pilot study utilized a real-world clinical cohort, and some patients have received previous treatments, and there were increases in certain conditions such as ascites, encephalopathy, and thyroid disorders in the cirrhosis cohort. However, these patient numbers were quite small and we were not powered to investigate their influence on miRNA expression. This analysis also included precursor hairpin miRNAs, which require additional processing to form mature miRNAs (*Lee et al., 2008*; *Annese et al., 2020*). Precursor miRNAs have been shown to be dysregulated in certain cancers (*Thomson et al., 2006*), although the mechanisms impacting these precursor miRNAs is not well understood. Furthermore, precursor miRNAs may be less stable and prone to degradation, making them poor biomarker candidates. However, studies have identified precursor miRNA configurations that are stable (*Krol et al., 2004*), indicating that additional research is needed to determine which precursor miRNAs may be stable in saliva and are potential biomarker candidates for HCC. However, because many of the salivary miRNAs detected here have been previously detected in the blood and tissue of patients with HCC, this lends additional support that many of the dysregulated salivary miRNAs observed here are indeed specific to the presence of HCC. The most significant miRNAs appear to cluster individuals based on HCC status rather than the presence of CLD or BCLC stage (Fig. 2), and the pathway analysis supports that most dysregulated miRNAs are enriched for cancer pathways. In addition, although we performed a stratified analysis comparing only patients with CLD, it is still possible that etiology of cirrhosis impacted miRNA expression. Additional studies with larger sample sizes will be needed to investigate the impact of cirrhosis etiology on miRNA expression and biomarker performance.

CONCLUSIONS

Overall, this study provides the first evidence that salivary miRNAs may serve as useful, non-invasive biomarkers for HCC. In addition, many of the identified miRNAs in saliva are concordant with previous findings of miRNAs in both plasma from HCC patients and HCC tissue. Future work should consider whether salivary miRNAs can help to improve detection of HCC either alone or in combination of other non-invasive biospecimens such as breath (*Miller-Atkins et al., 2020*), or in combination with other-omics technologies such as metabolomics or proteomics, to develop a more comprehensive and accurate screening approach for HCC.

ADDITIONAL INFORMATION AND DECLARATIONS

Competing Interests

Daniel M. Rotroff has an equity stake in Clarified Precision Medicine LLC. He holds intellectual property related to the detection of hepatocellular carcinoma.

Author Contributions

- Arshiya Mariam performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the paper, and approved the final draft.
- Galen Miller-Atkins performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the paper, and approved the final draft.
- Amika Moro performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the paper, and approved the final draft.
- Alejandro I. Rodarte performed the experiments, authored or reviewed drafts of the paper, and approved the final draft.
- Shirin Siddiqi performed the experiments, authored or reviewed drafts of the paper, and approved the final draft.
- Lou-Anne Acevedo-Moreno performed the experiments, authored or reviewed drafts of the paper, and approved the final draft.

- J. Mark Brown conceived and designed the experiments, authored or reviewed drafts of the paper, and approved the final draft.
- Daniela S. Allende conceived and designed the experiments, authored or reviewed drafts of the paper, and approved the final draft.
- Federico Aucejo conceived and designed the experiments, authored or reviewed drafts of the paper, and approved the final draft.
- Daniel M. Rotroff conceived and designed the experiments, authored or reviewed drafts of the paper, and approved the final draft.

Human Ethics

Data Availability

The following information was supplied regarding data availability:

The code is available at GitHub: https://github.com/rotroff-lab/salivary-miRNA-HCC and data is available in the Supplemental Files. The sequences described here were aligned to the miRBase v22 database and are available at Genbank: PRJNA755300.

Supplemental Information

Supplemental information for this article can be found online at http://dx.doi.org/10.7717/peerj.12715#supplemental-information.

REFERENCES

American Cancer Society. 2019. *Cancer facts & figures 2019*. Atlanta, GA, USA: American Cancer Society.

Amr KS, Atia HAE, Elbnhawy RAE, Ezzat WM. 2017. Early diagnostic evaluation of miR-122 and miR-224 as biomarkers for hepatocellular carcinoma. *Genes & Diseases* **4(4)**:215–221 DOI 10.1016/j.gendis.2017.10.003.

Andrews S. 2010. *FastQC: a quality control tool for high throughput sequence data*. Cambridge, United Kingdom: Babraham Bioinformatics, Babraham Institute.

Annese T, Tamma R, De Giorgis M, Ribatti D. 2020. microRNAs biogenesis, functions and role in tumor angiogenesis. *Frontiers in Oncology* **10**:e1003018 DOI 10.3389/fonc.2020.581007.

Ayuso C, Rimola J, Vilana R, Burrel M, Darnell A, García-Criado Á, Bianchi L, Belmonte E, Caparroz C, Barrufet M. 2018. Diagnosis and staging of hepatocellular carcinoma (HCC): current guidelines. *European Journal of Radiology* **101(Suppl. 7)**:72–81 DOI 10.1016/j.ejrad.2018.01.025.

Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, Marshall KA, Phillippy KH, Sherman PM, Holko M. 2012. NCBI GEO: archive for functional genomics data sets—update. *Nucleic Acids Research* **41(D1)**:D991–D995 DOI 10.1093/nar/gks1193.

Benjamini Y, Hochberg Y. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)* **57(1)**:289–300 DOI 10.1111/j.2517-6161.1995.tb02031.x.

Blaya D, Coll M, Rodrigo-Torres D, Vila-Casadesús M, Altamirano J, Llopis M, Graupera I, Perea L, Aguilar-Bravo B, Díaz A. 2016. Integrative microRNA profiling in alcoholic hepatitis reveals a role for microRNA-182 in liver injury and inflammation. *Gut* **65(9)**:1535–1545 DOI 10.1136/gutjnl-2015-311314.

Brunet Vega A, Pericay C, Moya I, Ferrer A, Dotor E, Pisa A, Casalots À, Serra-Aracil X, Oliva J-C, Ruiz A. 2013. microRNA expression profile in stage III colorectal cancer: circulating miR-18a and miR-29a as promising biomarkers. *Oncology Reports* **30(1)**:320–326 DOI 10.3892/or.2013.2475.

Chen S, Fu Z, Wen S, Yang X, Yu C, Zhou W, Lin Y, Lv Y. 2021. Expression and diagnostic value of miR-497 and miR-1246 in hepatocellular carcinoma. *Frontiers in Genetics* **12**:790 DOI 10.3389/fgene.2021.666306.

Chen S, Zhou Y, Chen Y, Gu J. 2018. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* **34(17)**:i884–i890 DOI 10.1093/bioinformatics/bty560.

Chong GO, Jeon H-S, Han HS, Son JW, Lee YH, Hong DG, Lee YS, Cho YL. 2015. Differential microRNA expression profiles in primary and recurrent epithelial ovarian cancer. *Anticancer Research* **35**:2611–2617.

Chuma M, Toyoda H, Matsuzaki J, Saito Y, Kumada T, Tada T, Kaneoka Y, Maeda A, Yokoo H, Ogawa K. 2019. Circulating microRNA-1246 as a possible biomarker for early tumor recurrence of hepatocellular carcinoma. *Hepatology Research* **49(7)**:810–822 DOI 10.1111/hepr.13338.

Daniele B, Bencivenga A, Megna AS, Tinessa V. 2004. α-fetoprotein and ultrasonography screening for hepatocellular carcinoma. *Gastroenterology* **127(5)**:S108–S112 DOI 10.1053/j.gastro.2004.09.023.

Dongiovanni P, Meroni M, Longo M, Fargion S, Fracanzani AL. 2018. miRNA signature in NAFLD: a turning point for a non-invasive diagnosis. *International Journal of Molecular Sciences* **19(12)**:3966 DOI 10.3390/ijms19123966.

Du L, Schageman JJ, Subauste MC, Saber B, Hammond SM, Prudkin L, Wistuba II, Ji L, Roth JA, Minna JD. 2009. miR-93, miR-98, and miR-197 regulate expression of tumor suppressor gene FUS1. *Molecular Cancer Research* **7(8)**:1234–1243 DOI 10.1158/1541-7786.MCR-08-0507.

Elashoff D, Zhou H, Reiss J, Wang J, Xiao H, Henson B, Hu S, Arellano M, Sinha U, Le A. 2012. Prevalidation of salivary biomarkers for oral cancer detection. *Cancer Epidemiology and Prevention Biomarkers* **21(4)**:664–672 DOI 10.1158/1055-9965.EPI-11-1093.

Eslam M, Valenti L, Romeo S. 2018. Genetics and epigenetics of NAFLD and NASH: clinical impact. *Journal of Hepatology* **68(2)**:268–279 DOI 10.1016/j.jhep.2017.09.003.

Fabian MR, Sonenberg N, Filipowicz W. 2010. Regulation of mRNA translation and stability by microRNAs. *Annual Review of Biochemistry* **79(1)**:351–379 DOI 10.1146/annurev-biochem-060308-103103.

Fang Z, Yin S, Sun R, Zhang S, Fu M, Wu Y, Zhang T, Khaliq J, Li Y. 2017. miR-140-5p suppresses the proliferation, migration and invasion of gastric cancer by regulating YES1. *Molecular Cancer* **16(1)**:1–11 DOI 10.1186/s12943-017-0708-6.

Fattovich G, Stroffolini T, Zagni I, Donato F. 2004. Hepatocellular carcinoma in cirrhosis: incidence and risk factors. *Gastroenterology* **127(5)**:S35–S50 DOI 10.1053/j.gastro.2004.09.014.

Forner A, Hessheimer AJ, Real MI, Bruix J. 2006. Treatment of hepatocellular carcinoma. *Critical Reviews in Oncology/Hematology* **60(2)**:89–98 DOI 10.1016/j.critrevonc.2006.06.001.

Frankel LB, Christoffersen NR, Jacobsen A, Lindow M, Krogh A, Lund AH. 2008. Programmed cell death 4 (PDCD4) is an important functional target of the microRNA miR-21 in breast cancer cells. *Journal of Biological Chemistry* **283(2)**:1026–1033 DOI 10.1074/jbc.M707224200.

Fuso P, Di Salvatore M, Santonocito C, Guarino D, Autilio C, Mulè A, Arciuolo D, Rinninella A, Mignone F, Ramundo M. 2021. Let-7a-5p, miR-100-5p, miR-101-3p, and miR-199a-3p hyperexpression as potential predictive biomarkers in early breast cancer patients. *Journal of Personalized Medicine* **11(8)**:816 DOI 10.3390/jpm11080816.

Gebert LF, MacRae IJ. 2019. Regulation of microRNA function in animals. *Nature Reviews Molecular Cell Biology* **20(1)**:21–37 DOI 10.1038/s41580-018-0045-7.

Griffiths-Jones S. 2010. miRBase: microRNA sequences and annotation. *Current Protocols in Bioinformatics* **29(1)**:12–19 DOI 10.1002/0471250953.bi1209s29.

He Z-J, Li W, Chen H, Wen J, Gao Y-F, Liu Y-J. 2018. miR-1306-3p targets FBXL5 to promote metastasis of hepatocellular carcinoma through suppressing snail degradation. *Biochemical and Biophysical Research Communications* **504(4)**:820–826 DOI 10.1016/j.bbrc.2018.09.059.

Huang J-L, Fu Y-P, Gan W, Liu G, Zhou P-Y, Zhou C, Sun B-Y, Guan R-Y, Zhou J, Fan J. 2020. Hepatic stellate cells promote the progression of hepatocellular carcinoma through microRNA-1246-RORα-Wnt/β-Catenin axis. *Cancer Letters* **476**:140–151 DOI 10.1016/j.canlet.2020.02.012.

Itami-Matsumoto S, Hayakawa M, Uchida-Kobayashi S, Enomoto M, Tamori A, Mizuno K, Toyoda H, Tamura T, Akutsu T, Ochiya T. 2019. Circulating exosomal miRNA profiles predict the occurrence and recurrence of hepatocellular carcinoma in patients with direct-acting antiviral-induced sustained viral response. *Biomedicines* **7(4)**:87 DOI 10.3390/biomedicines7040087.

Jansson MD, Lund AH. 2012. MicroRNA and cancer. *Molecular Oncology* **6(6)**:590–610 DOI 10.1016/j.molonc.2012.09.006.

Kim G, An H-J, Lee M-J, Song J-Y, Jeong J-Y, Lee J-H, Jeong H-C. 2016. Hsa-miR-1246 and hsa-miR-1290 are associated with stemness and invasiveness of non-small cell lung cancer. *Lung Cancer* **91**:15–22 DOI 10.1016/j.lungcan.2015.11.013.

Krol J, Sobczak K, Wilczynska U, Drath M, Jasinska A, Kaczynska D, Krzyzosiak WJ. 2004. Structural features of microRNA (miRNA) precursors and their relevance to miRNA biogenesis and small interfering RNA/short hairpin RNA design. *Journal of Biological Chemistry* **279(40)**:42230–42239 DOI 10.1074/jbc.M404931200.

Krämer A, Green J, Pollard J Jr, Tugendreich S. 2014. Causal analysis approaches in ingenuity pathway analysis. *Bioinformatics* **30(4)**:523–530 DOI 10.1093/bioinformatics/btt703.

Kuhn M. 2008. Building predictive models in R using the caret package. *Journal of Statistical Software* **28(5)**:1–26 DOI 10.18637/jss.v028.i05.

Kutay H, Bai S, Datta J, Motiwala T, Pogribny I, Frankel W, Jacob ST, Ghoshal K. 2006. Downregulation of miR-122 in the rodent and human hepatocellular carcinomas. *Journal of Cellular Biochemistry* **99**:671–678 DOI 10.1002/(ISSN)1097-4644.

Köberle V, Kronenberger B, Pleli T, Trojan J, Imelmann E, Peveling-Oberhag J, Welker M-W, Elhendawy M, Zeuzem S, Piiper A. 2013. Serum microRNA-1 and microRNA-122 are prognostic markers in patients with hepatocellular carcinoma. *European Journal of Cancer* **49(16)**:3442–3449 DOI 10.1016/j.ejca.2013.06.002.

Lee EJ, Baek M, Gusev Y, Brackett DJ, Nuovo GJ, Schmittgen TD. 2008. Systematic evaluation of microRNA processing patterns in tissues, cell lines, and tumors. *RNA* **14(1)**:35–42 DOI 10.1261/rna.804508.

Li F, Yoshizawa JM, Kim K-M, Kanjanapangka J, Grogan TR, Wang X, Elashoff DE, Ishikawa S, Chia D, Liao W. 2018. Discovery and validation of salivary extracellular RNA biomarkers for noninvasive detection of gastric cancer. *Clinical Chemistry* **64(10)**:1513–1521 DOI 10.1373/clinchem.2018.290569.

Lim Y, Wan Y, Vagenas D, Ovchinnikov DA, Perry CF, Davis MJ, Punyadeera C. 2016. Salivary DNA methylation panel to diagnose HPV-positive and HPV-negative head and neck cancers. *BMC Cancer* **16(1)**:1–12 DOI 10.1186/s12885-016-2785-0.

Lin S-S, Peng C-Y, Liao Y-W, Chou M-Y, Hsieh P-L, Yu C-C. 2018. miR-1246 targets CCNG2 to enhance cancer stemness and chemoresistance in oral carcinomas. *Cancers* **10(8)**:272 DOI 10.3390/cancers10080272.

Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* **15(12)**:550 DOI 10.1186/s13059-014-0550-8.

Luna JM, Scheel TK, Danino T, Shaw KS, Mele A, Fak JJ, Nishiuchi E, Takacs CN, Catanese MT, de Jong YP. 2015. Hepatitis C virus RNA functionally sequesters miR-122. *Cell* **160(6)**:1099–1110 DOI 10.1016/j.cell.2015.02.025.

Machida T, Tomofuji T, Maruyama T, Yoneda T, Ekuni D, Azuma T, Miyai H, Mizuno H, Kato H, Tsutsumi K. 2016. miR-1246 and miR-4644 in salivary exosome as potential biomarkers for pancreatobiliary tract cancer. *Oncology Reports* **36(4)**:2375–2381 DOI 10.3892/or.2016.5021.

Martinez-Quetglas I, Pinyol R, Dauch D, Torrecilla S, Tovar V, Moeini A, Alsinet C, Portela A, Rodriguez-Carunchio L, Solé M. 2016. IGF2 is up-regulated by epigenetic mechanisms in hepatocellular carcinomas and is an actionable oncogene product in experimental models. *Gastroenterology* **151(6)**:1192–1205 DOI 10.1053/j.gastro.2016.09.001.

Meng F, Henson R, Wehbe-Janek H, Ghoshal K, Jacob ST, Patel T. 2007. MicroRNA-21 regulates expression of the PTEN tumor suppressor gene in human hepatocellular cancer. *Gastroenterology* **133(2)**:647–658 DOI 10.1053/j.gastro.2007.05.022.

Miller-Atkins G, Acevedo-Moreno L-A, Grove D, Dweik RA, Tonelli AR, Brown JM, Allende DS, Aucejo F, Rotroff DM. 2020. Breath metabolomics provides an accurate and noninvasive approach for screening cirrhosis, primary, and secondary liver tumors. *Hepatology Communications* **4(7)**:1041–1055 DOI 10.1002/hep4.1499.

Murakami Y, Toyoda H, Tanahashi T, Tanaka J, Kumada T, Yoshioka Y, Kosaka N, Ochiya T, Taguchi YH. 2012. Comprehensive miRNA expression analysis in peripheral blood can diagnose liver disease. *PLOS ONE* **7(10)**:e48366 DOI 10.1371/journal.pone.0048366.

Nagpal N, Kulshreshtha R. 2014. miR-191: an emerging player in disease biology. *Frontiers in Genetics* **5(e10691)**:99 DOI 10.3389/fgene.2014.00099.

Nunes AP, Oliveira IO, Santos BR, Millech C, Silva LP, González DA, Hallal PC, Menezes AM, Araújo CL, Barros FC. 2012. Quality of DNA extracted from saliva samples collected with the OrageneTM DNA self-collection kit. *BMC Medical Research Methodology* **12**:65 DOI 10.1186/1471-2288-12-65.

Nuoroozi G, Mirmotalebisohi SA, Sameni M, Ariyanmehr Y, Zali H. 2021. Deregulation of microRNAs in oral squamous cell carcinoma, a bioinformatics analysis. *Gene Reports* **24(11)**:101241 DOI 10.1016/j.genrep.2021.101241.

Peng Q, Shen Y, Zhao P, Cheng M, Zhu Y, Xu B. 2020. Biomarker roles identification of miR-106 family for predicting the risk and poor survival of colorectal cancer. *BMC Cancer* **20**:1–13 DOI 10.1186/s12885-020-06863-9.

Petrick JL, Kelly SP, Altekruse SF, McGlynn KA, Rosenberg PS. 2016. Future of hepatocellular carcinoma incidence in the United States forecast through 2030. *Journal of Clinical Oncology* **34**:1787 DOI 10.1200/JCO.2015.64.7412.

Pineau P, Volinia S, McJunkin K, Marchio A, Battiston C, Terris B, Mazzaferro V, Lowe SW, Croce CM, Dejean A. 2010. miR-221 overexpression contributes to liver tumorigenesis. *Proceedings of the National Academy of Sciences* **107**:264–269 DOI 10.1073/pnas.0907904107.

Qi J, Wang J, Katayama H, Sen S, Liu S. 2013. Circulating microRNAs (cmiRNAs) as novel potential biomarkers for hepatocellular carcinoma. *Neoplasma* **60(2)**:135–142 DOI 10.4149/neo_2013_018.

Rapado-González Ó, Majem B, Muinelo-Romay L, Álvarez-Castro A, Santamaría A, Gil-Moreno A, López-López R, Suárez-Cunqueiro MM. 2018. Human salivary microRNAs in cancer. *Journal of Cancer* **9(4)**:638–649 DOI 10.7150/jca.21180.

Ren Y, Han X, Yu K, Sun S, Zhen L, Li Z, Wang S. 2014. microRNA-200c downregulates XIAP expression to suppress proliferation and promote apoptosis of triple-negative breast cancer cells. *Molecular Medicine Reports* **10(1)**:315–321 DOI 10.3892/mmr.2014.2222.

Salazar C, Calvopiña D, Punyadeera C. 2014. miRNAs in human papilloma virus associated oral and oropharyngeal squamous cell carcinomas. *Expert Review of Molecular Diagnostics* **14(8)**:1033–1040 DOI 10.1586/14737159.2014.960519.

Salazar C, Nagadia R, Pandit P, Cooper-White J, Banerjee N, Dimitrova N, Coman WB, Punyadeera C. 2014. A novel saliva-based microRNA biomarker panel to detect head and neck cancers. *Cellular Oncology* **37(5)**:331–338 DOI 10.1007/s13402-014-0188-2.

Schneider VA, Graves-Lindsay T, Howe K, Bouk N, Chen H-C, Kitts PA, Murphy TD, Pruitt KD, Thibaud-Nissen F, Albracht D. 2017. Evaluation of GRCh38 and de novo haploid genome assemblies demonstrates the enduring quality of the reference assembly. *Genome Research* **27(5)**:849–864 DOI 10.1101/gr.213611.116.

Singal AG, Pillai A, Tiro J. 2014. Early detection, curative treatment, and survival rates for hepatocellular carcinoma surveillance in patients with cirrhosis: a meta-analysis. *PLOS Medicine* **11(4)**:e1001624 DOI 10.1371/journal.pmed.1001624.

Tabrizian P, Jibara G, Shrager B, Schwartz M, Roayaie S. 2015. Recurrence of hepatocellular cancer after resection: patterns, treatments, and prognosis. *Annals of Surgery* **261(5)**:947–955 DOI 10.1097/SLA.0000000000000710.

Team RC. 2020. R: A language and environment for statistical computing. R foundation for statistical computing, Vienna (2019). *Available at* http://www.R-project.org/.

Thomson JM, Newman M, Parker JS, Morin-Kensicki EM, Wright T, Hammond SM. 2006. Extensive post-transcriptional regulation of microRNAs and its implications for cancer. *Genes & Development* **20(16)**:2202–2207 DOI 10.1101/gad.1444406.

Wan Y, Vagenas D, Salazar C, Kenny L, Perry C, Calvopiña D, Punyadeera C. 2017. Salivary miRNA panel to detect HPV-positive and HPV-negative head and neck cancer patients. *Oncotarget* **8(59)**:99990–100001 DOI 10.18632/oncotarget.21725.

Xiao J, Lv D, Zhou J, Bei Y, Chen T, Hu M, Zhou Q, Fu S, Huang Q. 2017. Therapeutic inhibition of miR-4260 suppresses colorectal cancer via targeting MCC and SMAD4. *Theranostics* **7(7)**:1901–1913 DOI 10.7150/thno.19168.

Yang J, Gong Y, Jiang Q, Liu L, Li S, Zhou Q, Huang F, Liu Z. 2020. Circular RNA expression profiles in nasopharyngeal carcinoma by sequence analysis. *Frontiers in Oncology* **10**:601 DOI 10.3389/fonc.2020.00601.

Ye L-L, Cheng Z-G, Cheng X-E, Huang Y-L. 2021. Propofol regulates miR-1-3p/IGF1 axis to inhibit the proliferation and accelerates apoptosis of colorectal cancer cells. *Toxicology Research* **10(4)**:696–705 DOI 10.1093/toxres/tfab047.

You Y, Que K, Zhou Y, Zhang Z, Zhao X, Gong J, Liu Z. 2018. MicroRNA-766-3p inhibits tumour progression by targeting Wnt3a in hepatocellular carcinoma. *Molecules and Cells* **41**:830 DOI 10.14348/molcells.2018.0181.

Zahran F, Ghalwash D, Shaker O, Al-Johani K, Scully C. 2015. Salivary micro RNA s in oral cancer. *Oral Diseases* **21(6)**:739–747 DOI 10.1111/odi.12340.

Zhang L, Xiao H, Karlan S, Zhou H, Gross J, Elashoff D, Akin D, Yan X, Chia D, Karlan B. 2010. Discovery and preclinical validation of salivary transcriptomic and proteomic biomarkers for the non-invasive detection of breast cancer. *PLOS ONE* **5(12)**:e15573 DOI 10.1371/journal.pone.0015573.

Zhang L, Xiao H, Zhou H, Santiago S, Lee JM, Garon EB, Yang J, Brinkmann O, Yan X, Akin D. 2012. Development of transcriptomic biomarker signature in human saliva to detect lung cancer. *Cellular and Molecular Life Sciences* **69(19)**:3341–3350 DOI 10.1007/s00018-012-1027-0.

Zheng Z, Xie B, Cui B, Ke Y, Lin Y. 2019. miR-106 b expression in breast cancer tissues and mechanism of enhancing cisplatin sensitivity in breast cancer cells. *Chinese Journal of Clinical Pharmacology and Therapeutics* **24**:541 DOI 10.12092/j.issn.1009-2501.2019.05.010.

Zhou W, Zou B, Liu L, Cui K, Gao J, Yuan S, Cong N. 2016. MicroRNA-98 acts as a tumor suppressor in hepatocellular carcinoma via targeting SALL4. *Oncotarget* **7(45)**:74059–74073 DOI 10.18632/oncotarget.12190.

Zhu A, Ibrahim JG, Love MI. 2019. Heavy-tailed prior distributions for sequence count data: removing the noise and preserving large differences. *Bioinformatics* **35(12)**:2084–2092 DOI 10.1093/bioinformatics/bty895.