SUPPLEMENTARY MATERIAL

Development and validation of a microRNA based diagnostic assay for primary tumor site classification of liver core biopsies.

Katharina Perell, Martin Vincent, Ben Vainer ,Bodil Laub Petersen, Birgitte Federspiel, Anne Kirstine Møller, Mette Madsen, Niels Richard Hansen, Lennart Friis-Hansen, Finn Cilius Nielsen, Gedske Daugaard.

miRNA Quantitative real-time PCR profiling

TaqMan low density array (TLDA) cards, human MicroRNA array A (Applied Biosystems) were used to determine the expression of 377 microRNAs. Each array contains six positive controls. RT-PCR reactions were performed according to the manufacturer's instructions. All reagents were obtained from Applied Biosystems. Briefly, 30 ng of total RNA was reverse transcribed (RT) with Megaplex RT primer human pool A and the TaqMan miRNA reverse transcription kit in a total volume of 7.5 µl per reaction. The amount of miRNA that can be extracted from core biopsies may be limited, and a 40 round pre-amplification step was therefore included. cDNA of liver core biopsies as well as primary tumor resections was pre-amplified in order to make the analysis from primary tumors and metastases comparable. Pre-amplification was performed using 2.5 µl RT product together with Megaplex PreAmp Primers and TaqMan PreAmp Master Mix in a 25 µl PCR reaction. Given that the FFPE tissue was not collected with RNA preservation in mind, we used the expression of the small nucleolar RNA, RNU44, as a surrogate measure of RNA integrity prior to miRNA quantification. RNU44 expression was examined in triplicate in the pre-amplified cDNA-mix using a 384-well plate. Each reaction consisted of 2.5 µl pre-amplified cDNA, 0.5 µl TaqMAn MicroRNA assay, 5 µl Universal Master Mix No AmpErase UNG and 2 µl Nuclease-free water (all from Applied Biosystems). PCR reactions were run on an Applied Biosystems 7900 HT system, according to the manufacturer`s instructions and analyzed using SDS software (v.2.4, automatic baseline setting). Based on preliminary results, we defined a mean cycle threshold (Ct) value ≤ 20 as a cut-off and only samples with Ct values below this cut-off were further processed. Pre-amplified cDNA was diluted with 375 µl 0.1 TE (ph. 8.0) and transferred to a TaqMan Human MicroRNA A array (v. 2.0). Quantitative real-time PCR was performed using an Applied Biosystems ViiA 7 Real Time PCR and TaqMan Advanced Master Mix with 50 µl input cDNA template per lane. Ct values were calculated using the ViiA 7 software (v. 1.1).

Successful analysis was performed for 333 samples (98.5 %). Five samples were excluded because the RNU44 Ct-values were above the predefined cut-off.

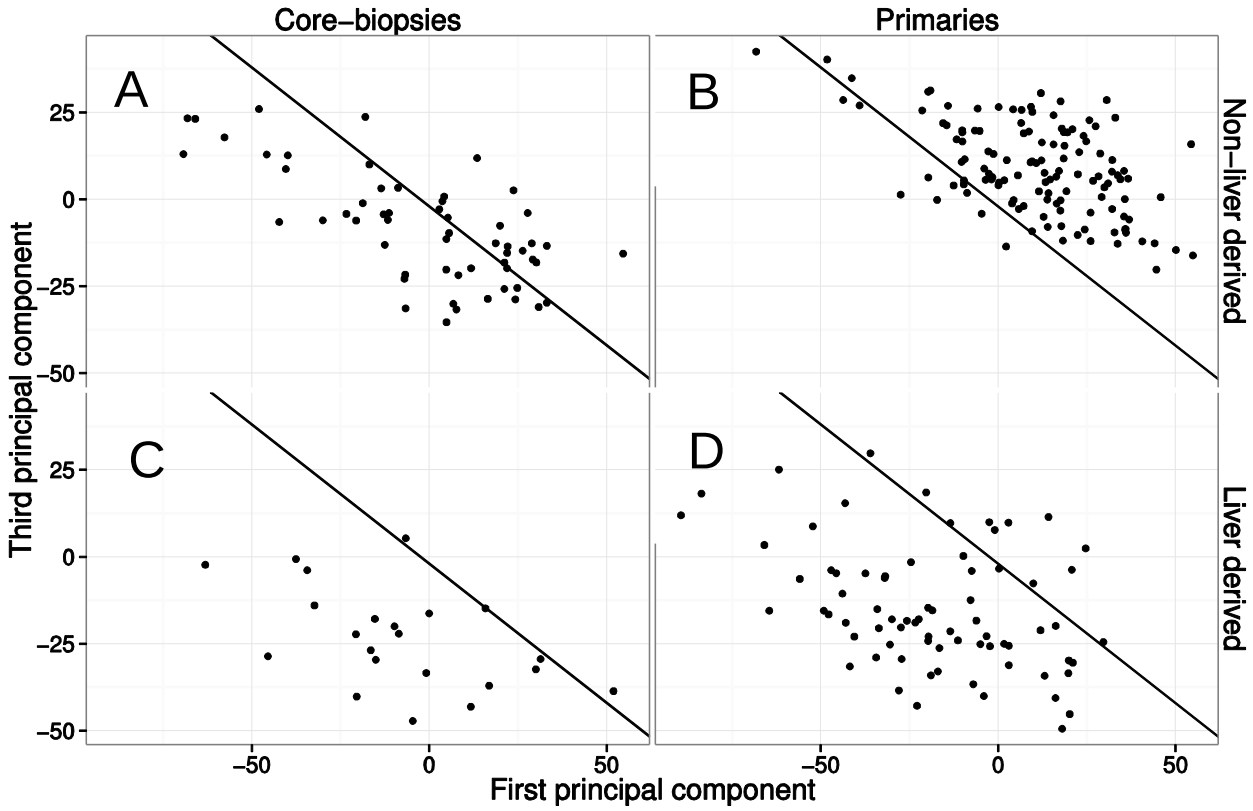
Supplementary Figure S1

		Reference diagnosis									
		Bladder	Breast	CCA	CRC	GC	HCC	Normal liver	Lung	Pancreas	Squamous
Predicted diagnosis	Bladder	17	0	0	0	0	0	0	0	0	1
	Breast	0	16	0	0	0	0	0	2	0	1
	CCA	0	0	14	0	1	1	0	0	0	0
	CRC	0	0	0	19	1	0	0	0	0	0
	GC	0	0	0	1	16	0	0	0	0	0
	HCC	0	0	0	0	0	14	0	0	0	0
	Normal liver	0	0	2	0	0	2	37	0	0	0
	Lung	0	0	1	0	0	0	0	14	1	1
	Pancreas	0	0	3	0	0	0	0	1	19	0
	Squamous	0	1	0	0	0	0	0	0	0	13
Positive percentage agreement		**100%**	**94%**	**70%**	**95%**	**89%**	**82%**	**100%**	**82%**	**95%**	**81%**

Supplementary Figure S1. Confusion matrix showing PRIM classifier predictions upon cross-validation. *Primary tumor site predictions of the PRIM classifier obtained by cross-validation are shown. The PRIM classifier was trained on primary tumor and normal liver resections only. Classes according to the reference diagnosis are shown along the columns and classes according to classifier predictions are shown along the rows. The positive percentage agreement for each assay class was calculated.*

***Squamous**, squamous cell carcinoma (mixed population); **CCA**, cholangiocarcinoma; **CRC**, Colorectal carcinoma; **GC**, gastric or cardia carcinoma; **HCC**, hepatocellular carcinoma.*

Supplementary Figure S2



Supplementary Figure S2. Principal Component Plot illustrating the clustering of training samples. *Training samples were divided according to:* ***A****) Core biopsies from non-liver derived malignancies representing metastases from bladder, breast, colorectal, gastric/cardia, lung, pancreatic cancer and mixed primary tumors of squamous cell morphology (57 samples).* ***B)*** *Resected non-liver derived malignancies constituting primary tumors from the same 7 classes mentioned above (125 samples).* ***C)*** *Core biopsies from liver derived malignancies (hepatocellular carcinoma, cholangiocarcinoma) and normal liver (reactive and cirrhotic) (22 samples).* ***D)*** *Resected liver-derived malignancies and normal liver (74 samples).*

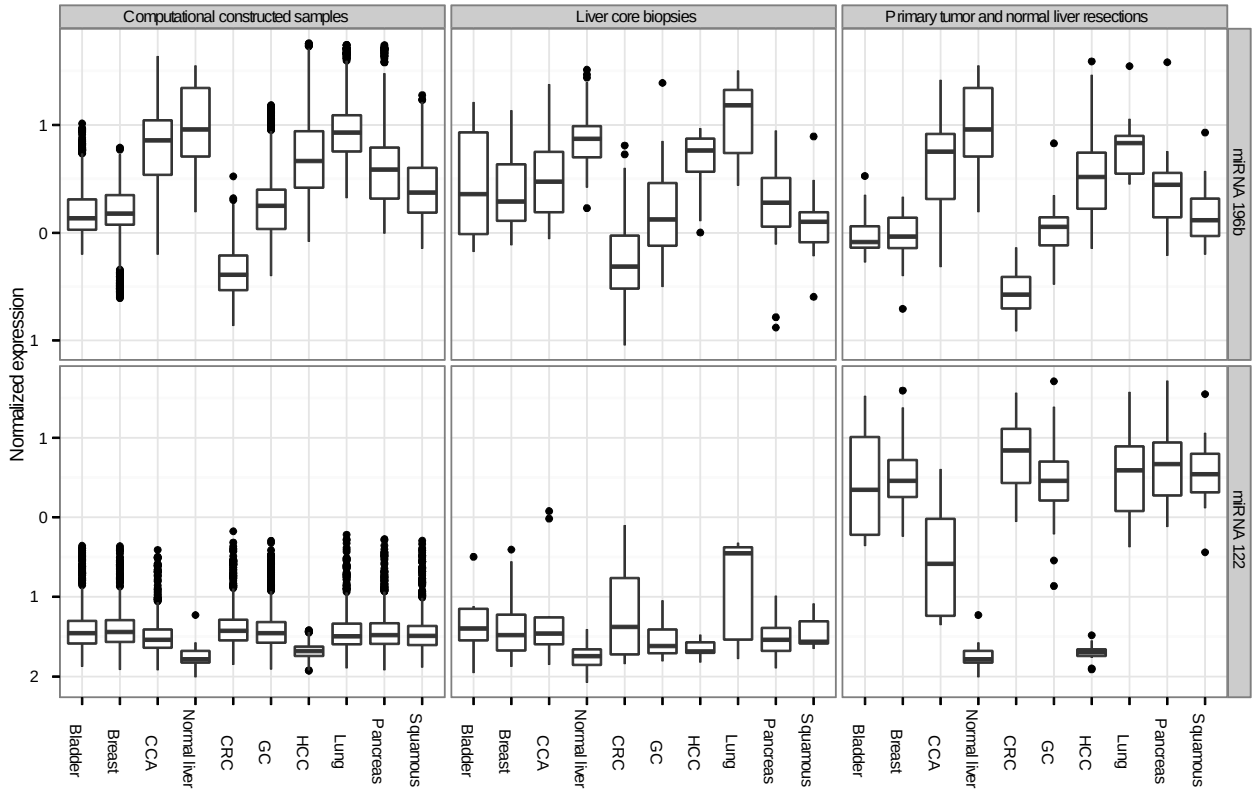
Supplementary Figure S3

Predicted diagnosis \ Reference diagnosis	Bladder	Breast	CCA	CRC	GC	HCC	Normal liver	Lung	Pancreas	Squamous
Bladder	1	0	0	0	0	0	0	0	0	1
Breast	0	4	0	0	0	0	0	0	1	0
CCA	0	1	4	0	1	0	2	1	0	0
CRC	0	0	0	8	2	0	0	0	1	0
GC	0	1	0	3	8	0	0	0	2	2
HCC	0	0	0	0	0	1	0	0	0	0
Normal liver	0	0	0	0	0	2	13	0	1	0
Lung	0	1	0	0	0	0	0	1	0	0
Pancreas	0	0	0	1	1	0	0	0	4	0
Squamous	1	0	0	0	0	0	0	0	1	9
Positive percentage agreement	**50%**	**57%**	**100%**	**67%**	**67%**	**33%**	**87%**	**50%**	**40%**	**75%**

Supplementary Figure S3. ***Confusion matrix showing CCM+CB classifier predictions upon 8-fold cross-validation.*** *Cross-validation was performed on 79 liver core biopsies. Each assay class was represented by 2-12 samples. Each row and column corresponds to one of the assay classes included in the classifier. Columns indicate classes according to the reference diagnosis; rows indicate the diagnosis predicted by the CCM+CB classifier. Numbers on the diagonal indicate cases for which the predicted diagnosis matched the reference diagnosis, whereas off-diagonal numbers were in disagreement and counted as test errors. The positive percentage agreement for each class was calculated.*

Squamous*, squamous cell carcinoma (mixed population);* ***CCA****, cholangiocarcinoma;* ***CRC****, colorectal carcinoma;* ***GC****, gastric or cardia carcinoma;* ***HCC****, hepatocellular carcinoma.*

Supplementary Figure S4



Supplementary Figure S4. Normalized expression profiles of miR-122 and miR-196b in computational constructed samples, liver core biopsies and primary tumor and normal liver resections. *Expression profiles below 0 indicate that the miRNA is up-regulated. Accordingly, miR-196b is shown to be up-regulated in colorectal cancer in all three sample sets. This miR is selected as part of the classification signature in all three classifiers (PRIM, CCM, CCM+CB). In contrast, miR-122 is up regulated in normal liver (reactive liver and cirrhosis) HCC and CCA which makes it an excellent discriminative miRNA in the PRIM classifier, but in samples with liver contamination, the predictive ability diminishes.*

CCA: *cholangiocarcinoma;* ***CRC:*** *colorectal carcinoma;* ***GC:*** *gastric or cardia carcinoma;* ***HCC:*** *hepatocellular carcinoma.*

Supplementary Table S1.

Comparison of classifier predictions for 134 liver core biopsies.

Reference Class	Sample age	Tumor %	T/V	PRIM	CCM	CCM+CB
Bladder	2008	10 %	Test	CCA	Squamous	Squamous
Bladder	2006	30 %	Test	Liver	Liver	Bladder
CCA	2006	35 %	Test	CCA	CCA	CCA
CCA	2010	80 %	Test	CCA	CCA	CCA
CCA	2007	30 %	Test	CCA	Cirrhosis	CCA
CCA	2010	40 %	Test	CCA	Squamous	CCA
Cirrhosis	2077	None	Test	Cirrhosis	Cirrhosis	CCA
Cirrhosis	2010	None	Test	Cirrhosis	CCA	CCA
Cirrhosis	2006	None	Test	CCA	Liver	Liver
Cirrhosis	2008	None	Test	Cirrhosis	Cirrhosis	Cirrhosis
Cirrhosis	2011	None	Test	Cirrhosis	Cirrhosis	Cirrhosis
Cirrhosis	2011	None	Test	Cirrhosis	Cirrhosis	Cirrhosis
Cirrhosis	2011	None	Test	Cirrhosis	Cirrhosis	Cirrhosis
Cirrhosis	2008	None	Test	Liver	Cirrhosis	Cirrhosis
CRC	2010	50 %	Test	CRC	CRC	CRC
CRC	2010	25 %	Test	Liver	CRC	GC
CRC	2010	100 %	Test	CRC	CRC	CRC
CRC	2007	25 %	Test	CRC	CRC	CRC
CRC	2005	40 %	Test	CRC	CRC	CRC
CRC	2011	50 %	Test	CRC	CRC	CRC
CRC	2010	75 %	Test	Liver	GC	GC
CRC	2010	25 %	Test	CRC	CRC	CRC
CRC	2010	100 %	Test	CRC	CRC	CRC
CRC	2006	20 %	Test	CCA	CCA	Pancreas
CRC	2008	100 %	Test	CRC	CRC	CRC
CRC	2003	40 %	Test	Liver	Lung	GC

GC	2008	30 %	Test	CCA	GC	GC
GC	2008	50 %	Test	GC	GC	Squamous
GC	2008	50 %	Test	Liver	GC	GC
GC	2009	40 %	Test	Liver	CRC	CRC
GC	2010	40 %	Test	CCA	CCA	GC
GC	2004	80 %	Test	CCA	CCA	CCA
GC	2008	70 %	Test	Liver	GC	GC
GC	2011	90 %	Test	Liver	Breast	Pancreas
GC	2010	100 %	Test	CCA	GC	GC
GC	2010	20 %	Test	GC	GC	GC
GC	2010	20 %	Test	Liver	GC	GC
GC	2010	40 %	Test	CCA	GC	GC
HCC	2009	15 %	Test	Cirrhosis	Cirrhosis	Cirrhosis
HCC	2010	90 %	Test	HCC	HCC	HCC
HCC	2009	60 %	Test	Cirrhosis	Cirrhosis	Cirrhosis
Liver	2008	None	Test	Liver	Liver	Liver
Liver	2007	None	Test	Liver	Liver	Liver
Liver	2011	None	Test	Cirrhosis	Liver	Cirrhosis
Liver	2008	None	Test	Liver	Liver	Liver
Liver	2008	None	Test	Liver	Liver	Liver
Liver	2009	None	Test	Liver	Liver	Liver
Liver	2010	None	Test	Liver	Liver	Liver
Lung	2008	75 %	Test	Lung	Lung	Lung
Lung	2011	30 %	Test	Liver	CCA	CCA
Breast	2010	40 %	Test	Cirrhosis	Cirrhosis	CCA
Breast	2005	100 %	Test	Pancreas	Pancreas	Breast
Breast	2002	40 %	Test	Breast	CCA	CCA
Breast	2002	40 %	Test	Liver	Lung	Breast
Breast	2009	30 %	Test	HCC	GC	GC
Breast	2002	100 %	Test	Breast	Liver	Breast

Breast	2010	15 %	Test	Liver	Liver	Breast
Pancreas	2010	90 %	Test	Pancreas	Pancreas	Pancreas
Pancreas	2010	80 %	Test	CCA	Liver	Liver
Pancreas	2008	25 %	Test	Cirrhosis	Pancreas	Pancreas
Pancreas	2011	25 %	Test	CCA	CCA	Pancreas
Pancreas	2011	80 %	Test	CCA	Pancreas	Pancreas
Pancreas	2008	90 %	Test	CRC	CRC	CRC
Pancreas	2011	35 %	Test	Liver	GC	GC
Pancreas	2011	25 %	Test	HCC	Bladder	Squamous
Pancreas	2011	95 %	Test	CCA	GC	GC
Pancreas	2008	50 %	Test	CCA	Breast	Breast
Squamous/Esophagus	2006	25 %	Test	Squamous	Squamous	Squamous
Squamous/Lung	2011	40 %	Test	HCC	Squamous	Squamous
Squamous/Anal	2011	25 %	Test	CCA	Bladder	Bladder
Squamous/Esophagus	2009	40 %	Test	CCA	Squamous	Squamous
Squamous/Anal	2009	50 %	Test	HCC	Squamous	Squamous
Squamous/Head and Neck	2011	80 %	Test	Bladder	Bladder	Squamous
Squamous/Head and Neck	2009	75 %	Test	CCA	Squamous	Squamous
Squamous/Head and Neck	2005	50 %	Test	Bladder	Squamous	Squamous
Squamous/Esophagus	2011	60 %	Test	HCC	Squamous	Squamous
Squamous/Esophagus	2010	75 %	Test	GC	GC	GC
Squamous/Cervix	2011	95 %	Test	GC	GC	GC
Squamous/Head and Neck	2008	75 %	Test	Squamous	Squamous	Squamous
Bladder	2007	35 %	Validation	Cirrhosis	CCA	CCA
Bladder	2005	50 %	Validation	Liver	Bladder	Bladder
Bladder	2011	60 %	Validation	Bladder	Bladder	Bladder
Bladder	2011	75 %	Validation	HCC	Squamous	Breast
Bladder	2010	90 %	Validation	CCA	Bladder	Bladder
CCA	2011	75 %	Validation	Cirrhosis	CCA	CCA
CCA	2001	85 %	Validation	Cirrhosis	Cirrhosis	Cirrhosis

CCA	2002	80 %	Validation	CCA	CCA	Cirrhosis
CCA	2009	40 %	Validation	Liver	Liver	Liver
CCA	2011	80 %	Validation	Cirrhosis	CCA	GC
Cirrhosis	2011	None	Validation	Cirrhosis	Cirrhosis	Cirrhosis
Cirrhosis	2007	None	Validation	Cirrhosis	Cirrhosis	Cirrhosis
Cirrhosis	2004	None	Validation	Cirrhosis	Cirrhosis	Cirrhosis
Cirrhosis	2010	None	Validation	CCA	Liver	Liver
Cirrhosis	2010	None	Validation	Cirrhosis	Cirrhosis	Cirrhosis
CRC	2009	25 %	Validation	HCC	Squamous	Squamous
CRC	2008	35 %	Validation	Liver	CRC	CRC
CRC	2007	95 %	Validation	CRC	CRC	CRC
CRC	2005	90 %	Validation	CRC	CRC	CRC
CRC	2005	60 %	Validation	CRC	CRC	CRC
GC	2008	25 %	Validation	Liver	Squamous	GC
GC	2005	50 %	Validation	CCA	CCA	CCA
GC	2007	90 %	Validation	HCC	GC	GC
GC	2004	25 %	Validation	Liver	CCA	Pancreas
GC	2010	35 %	Validation	GC	GC	GC
HCC	2010	50 %	Validation	HCC	HCC	HCC
HCC	2009	75 %	Validation	HCC	HCC	HCC
HCC	2009	70 %	Validation	CCA	Liver	HCC
HCC	2011	90 %	Validation	HCC	HCC	HCC
HCC	2010	25 %	Validation	Cirrhosis	HCC	HCC
Liver	2005	None	Validation	Liver	Liver	Liver
Liver	2010	None	Validation	Liver	Liver	Liver
Liver	2010	None	Validation	Liver	Liver	Liver
Liver	2009	None	Validation	Liver	Liver	Liver
Liver	2005	None	Validation	Liver	Liver	Liver
Lung	2008	80 %	Validation	CCA	Lung	Lung
Lung	2010	50 %	Validation	CCA	Lung	Lung

Lung	2002	70 %	Validation	CCA	Lung	Lung
Lung	2007	50 %	Validation	CCA	CCA	CCA
Lung	2001	90 %	Validation	CCA	CCA	Pancreas
Breast	2007	50 %	Validation	Liver	Breast	Breast
Breast	2011	25 %	Validation	Cirrhosis	Breast	Pancreas
Breast	2003	90 %	Validation	Squamous	Squamous	Breast
Breast	2004	95 %	Validation	CCA	Breast	Breast
Breast	2011	60 %	Validation	Liver	Breast	Breast
Pancreas	2010	40 %	Validation	GC	GC	Pancreas
Pancreas	2010	30 %	Validation	Liver	CCA	Pancreas
Pancreas	2010	90 %	Validation	CCA	CCA	Pancreas
Pancreas	2010	80 %	Validation	GC	GC	Pancreas
Pancreas	2007	75 %	Validation	CCA	Bladder	Pancreas
Squamous/Lung	2004	70 %	Validation	CCA	Squamous	Squamous
Squamous/Cervix	2004	75 %	Validation	Squamous	Squamous	Breast
Squamous/Esophagus	2008	10 %	Validation	Squamous	Squamous	Squamous
Squamous/Esophagus	2009	50 %	Validation	CCA	Cirrhosis	Cirrhosis
Squamous/Esophagus	2011	50 %	Validation	Liver	Squamous	Squamous

Supplementary Table S2

A complete list of selected microRNAs for the PRIM classifier, CCM classifier and CCM+CB classifier.

MicroRNA	CCM+CB	CCM	PRIM
let.7c	X	X	X
let.7f	X	X	
miR.1		X	
miR.9	X	X	
miR.10a	X	X	X
miR.10b	X	X	X
miR.17		X	X
miR.18a	X	X	
miR.21	X		X
miR.23b	X	X	X
miR.25	X	X	X
miR.27a	X	X	X
miR.27b	X		X
miR.28.5p	X	X	
miR.29b	X	X	
miR.29c	X		X
miR.31	X	X	
miR.33b	X	X	
miR.34a	X	X	X
miR.34c.5p	X	X	
miR.92a	X	X	X
miR.98	X		
miR.99b	X	X	X
miR.105		X	
miR.106a	X		
miR.107	X	X	X
miR.122			X
miR.128	X		
miR.129.3p	X	X	X
miR.129.5p		X	
miR.130b	X	X	X
miR.133a	X	X	X
miR.133b	X		
miR.135a	X		
miR.135b	X	X	X
miR.136	X	X	
miR.138	X	X	X
miR.139.3p	X	X	
miR.139.5p	X		
miR.140.3p	X	X	

miR.141	X	X	
miR.143	X	X	X
miR.146b.5p	X		
miR.147b	X	X	X
miR.148a			X
miR.149	X	X	
miR.152	X	X	
miR.181a	X	X	X
miR.181c	X	X	X
miR.182	X	X	X
miR.187	X	X	X
miR.190		X	X
miR.191	X	X	X
miR.192	X	X	X
miR.193a.3p	X	X	
miR.193a.5p	X		
miR.196b	X	X	X
miR.198	X	X	X
miR.200a	X	X	
miR.200c	X	X	
miR.203	X	X	X
miR.205	X	X	X
miR.210	X	X	X
miR.214	X	X	X
miR.216a	X		
miR.217			X
miR.221	X	X	X
miR.223	X	X	X
miR.224	X	X	X
miR.296.3p	X	X	
miR.301b	X		
miR.302a	X		
miR.302b	X	X	X
miR.302c	X		
miR.323.3p	X	X	X
miR.324.3p	X	X	
miR.324.5p	X	X	
miR.331.5p	X	X	
miR.338.3p	X	X	X
let.7b	X	X	
miR.342.3p	X		
miR.342.5p	X	X	
miR.362.3p		X	
miR.363	X		
miR.365	X	X	
miR.370	X		

miR.375	X	X	X
miR.376a		X	
miR.379	X	X	X
miR.381	X	X	
miR.411	X	X	
miR.423.5p	X	X	
miR.425	X		
miR.429	X		
miR.433		X	
miR.449a	X		
miR.449b	X	X	
miR.453		X	
miR.483.5p	X	X	
miR.484		X	X
miR.494	X		
miR.501.3p	X	X	X
miR.502.3p	X	X	
miR.505			X
miR.508.3p		X	X
miR.510		X	
miR.512.3p		X	
miR.515.5p	X	X	
miR.516b	X	X	
miR.517a	X	X	
miR.520a.5p		X	
miR.526b	X	X	X
miR.532.3p	X	X	X
miR.532.5p		X	
miR.542.5p	X		
miR.545	X	X	
miR.548c.3p	X		
miR.548c.5p	X		
miR.548d.3p		X	X
miR.551b		X	
miR.582.3p	X		
miR.589		X	
miR.615.5p	X	X	X
miR.616	X		
miR.625	X	X	
miR.627	X	X	
miR.628.5p	X		
miR.642	X	X	
miR.653	X		
miR.654.3p	X		
miR.655	X	X	X
miR.671.3p	X	X	

miR.758		X	
miR.873		X	
miR.875.3p	X		
miR.885.5p			X
miR.891a	X	X	
miR.346		X	
miR.492	X	X	
miR.509.3.5p	X		
miR.511		X	
miR.517b	X		
miR.520f	X		