

## Supplementary Files

**Supplementary Table 1.** The numbers of circRNAs predicted before and after the filtering criteria described in Material and Methods. The final percentage of predicted circRNAs that pass the filtering procedure is reported.

	DCIS1	DCIS2	DCIS3	DCIS4	DCIS6	IDC1	IDC2	IDC3	IDC4	IDC6
<b>Number of circRNAs pre filtering procedure</b>	243	188	68	88	916	84	85	94	57	115
<b>Number of circRNAs after filtering procedure</b>	81	61	41	25	360	57	17	27	37	70
<b>Percentage</b>	33.33	32.45	60.29	28.41	39.30	67.86	20.00	28.72	64.91	60.87

**Supplementary Table 2.** Spearman correlation rho values were listed for each of the EMT genes studied in the cohort. VIM and CDH1 were not targeted by miRNAs investigated, but were included because are important EMT markers.

Genes	miRTarBase	Spearman rho	2 tails p-values
<b>ZEB1</b>	miR-200c - miR200b - miR-429	-0.3142857	0.5441
<b>ZEB2</b>	miR-200c - miR200b - miR-429	-0.08571429	0.8717
<b>BMI</b>	miR-200c - miR200b	<b>0.8857143</b>	<b>0.01885</b>
<b>FN1</b>	miR-200c - miR200b	<b>-0.8857143</b>	<b>0.01885</b>
<b>VIM</b>	-	-0.4285714	0.3965
<b>CDH1</b>	-	0.7714286	0.0724

**Supplementary Table 3.** The table listed the diseases enriched for the hsa-circ-001803 using the circ2traits database.

ID circRNA	Disease	<i>p-values</i>
hsa_circ_001803	lung cancer	1.16E-10
hsa_circ_001803	non-small cell lung cancer	3.01E-06
<b>hsa_circ_001803</b>	<b>breast cancer</b>	<b>0.001762331</b>
hsa_circ_001803	head and neck squamous cell carcinoma	0.001762331
hsa_circ_001803	ovarian cancer	0.001762331
hsa_circ_001803	adenoma	0.029454035
hsa_circ_001803	bladder cancer	0.029454035
hsa_circ_001803	Burkitt lymphoma	0.029454035
hsa_circ_001803	cholangiocarcinoma	0.029454035
hsa_circ_001803	colorectal cancer	0.029454035
hsa_circ_001803	coronary artery disease	0.029454035
hsa_circ_001803	gastric cancer	0.029454035
hsa_circ_001803	glioma	0.029454035
hsa_circ_001803	hamartoma	0.029454035
hsa_circ_001803	laryngeal carcinoma	0.029454035
hsa_circ_001803	lipoma	0.029454035
hsa_circ_001803	malignant melanoma	0.029454035
hsa_circ_001803	myoma	0.029454035
hsa_circ_001803	prostate cancer	0.029454035
hsa_circ_001803	sarcoma	0.029454035
hsa_circ_001803	squamous carcinoma	0.029454035
hsa_circ_001803	tongue squamous cell carcinoma	0.029454035
hsa_circ_001803	uterine leiomyoma	0.029454035

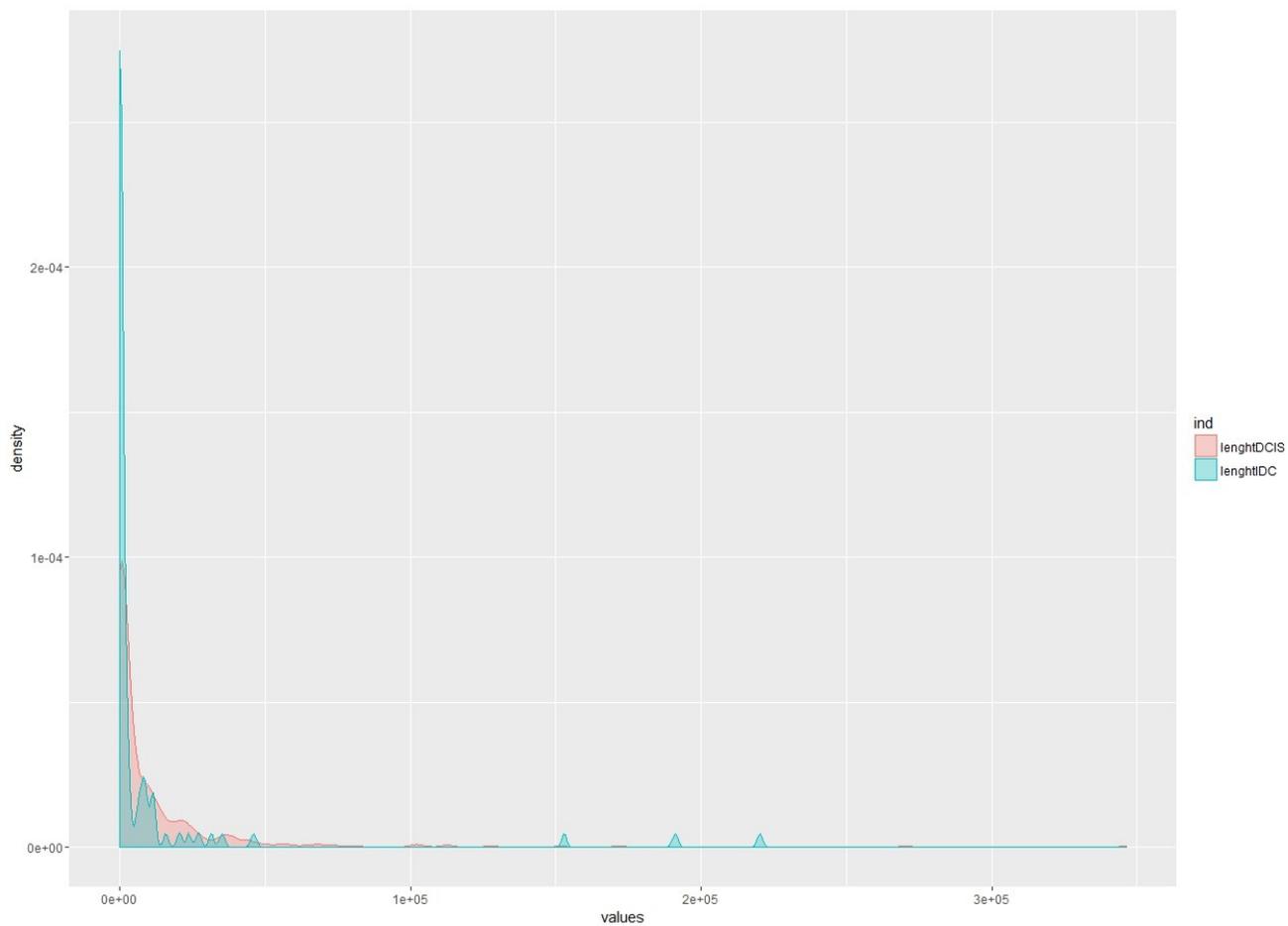
## Supplementary Figures

**Supplementary Figure 1.** The plot showed the density presence of predicted circRNAs (*y* axes) related to their length (*x* axes) comparing the DCIS (rose) and IDC (blue) samples.

**Supplementary Figure 2.** The heatmap showed the log expression values of 18 predicted circRNAs (*y* axis) . The name of each circRNAs was reported including the name of the genes the circRNAs derived from. All the samples were listed (*x*-axis). Log<sub>2</sub> SRBM values were reported. Black label showed no detectable levels.

**Supplementary Figure 3.** The alignment scores of each miRNAs against hsa-circ-0001358 were reported. The alignments were obtained using the Starbase human Pan cancer tool.

**Supplementary Figure 1.**





### Supplementary Figure 3.

miR-200c-3p miRNA 3'-aggtagtAATGGGCCGCATAAt-5'  
                  || : | |||||  
ncRNA 5'-aggctgtTTTGTAGCCAGTATTc-3'  
*alignScore = 144*

miR-429-3p miRNA 3'-tgcCAAAATGGTCTGCATAAt-5'  
                  |||||: :|| |||||  
ncRNA 5'-gctGTTTTG-TAGCCAGTATTc-3'  
*alignScore = 165*

miR-200b-3p miRNA 3'-agtagtAATGGTCCGCATAAt-5'  
                  || :|| |||||  
ncRNA 5'-ggctgtTTTGTAGCCAGTATTc-3'  
*alignScore = 155*

miR-376a-3p miRNA 3'-tgCACCTAAAAG--GAGATACTa-5'  
                  ||| || | : |||||  
ncRNA 5'-agGTGTATGTATGGATCTATGAc-3'  
*alignScore = 146*

miR-376b-3p miRNA 3'-ttgTACCTAAAAG--GAGATACTa-5'  
                  :|| || | : |||||  
ncRNA 5'-gagGTGTATGTATGGATCTATGAc-3'  
*alignScore = 142*