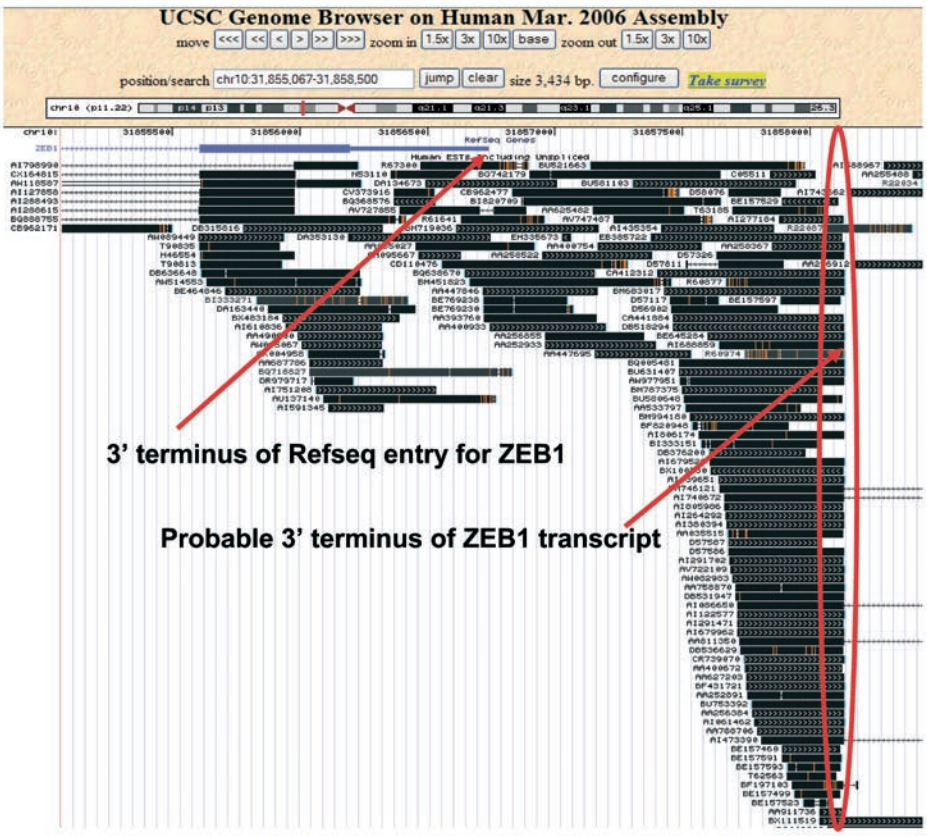

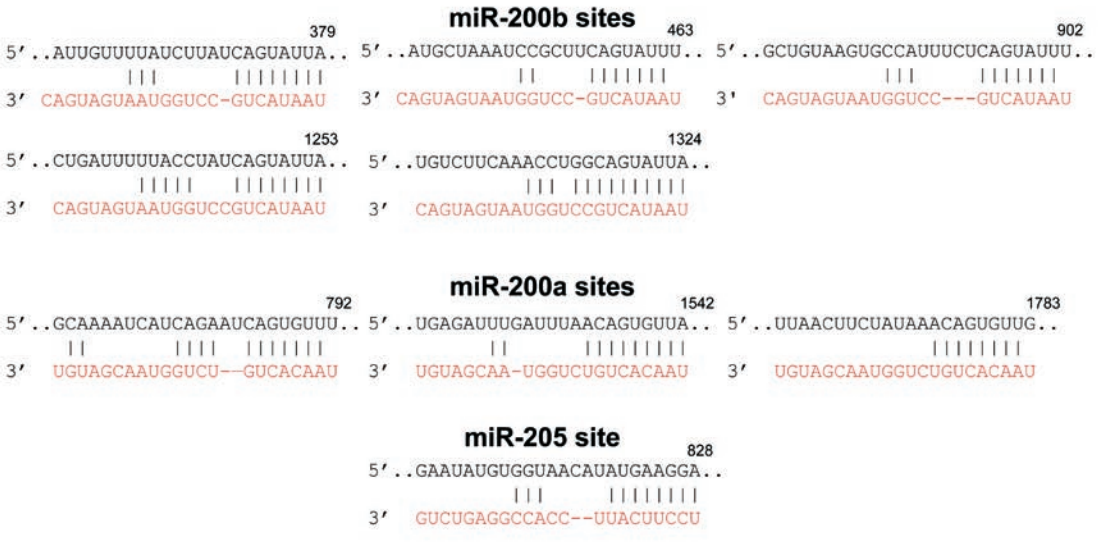

Figure S1 ZEB1 full length mRNA. **(a)** Analysis of the ZEB1 mRNA using the UCSC genome browser (http://genome.ucsc.edu) revealed truncation of the annotated Refseq sequence (NM_030751). The probable terminus of the *ZEB1* 3'UTR is indicated by multiple expressed sequence tags (ESTs) ending at the same position (~1.2 kb downstream of the Refseq terminus). (**b**) Predicted interactions of miR-200a, miR-200b and miR-205 with their binding sites in the *ZEB1* 3'UTR. MicroRNA sequences are shown in red.

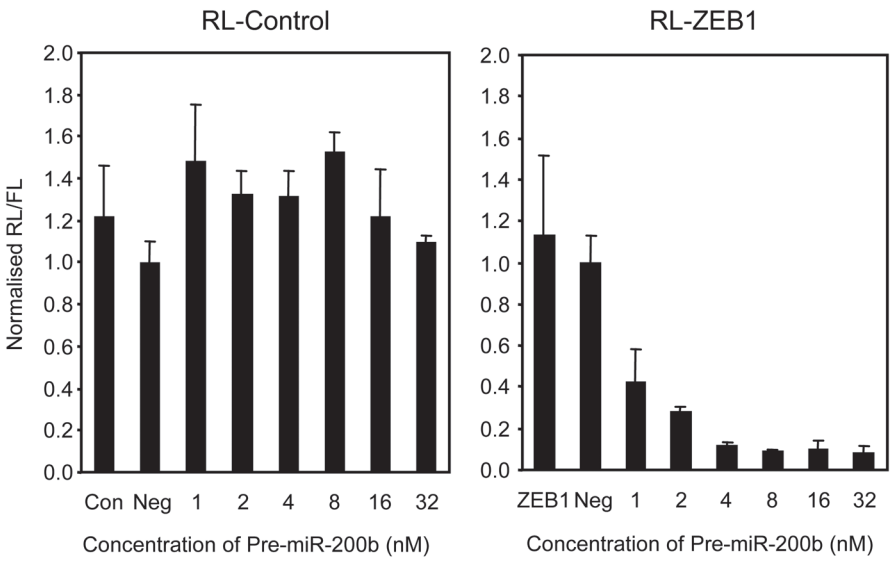

Figure S2 Determination of the minimum effective concentration of Pre-miR for inhibiting reporter expression. A range of concentrations of miR-200b Pre-miR was cotransfected with RL-Control or RL-ZEB1. Comparisons are made with samples without cotransfected microRNA (Con or ZEB1) or cotransfected with a negative control Pre-miR (Neg). The pGL3 plasmid was cotransfected to normalise for transfection efficiency and the ratio of Renilla to Firefly luciferase activity is shown from a triplicate transfection ± s.e.m. (n=3).

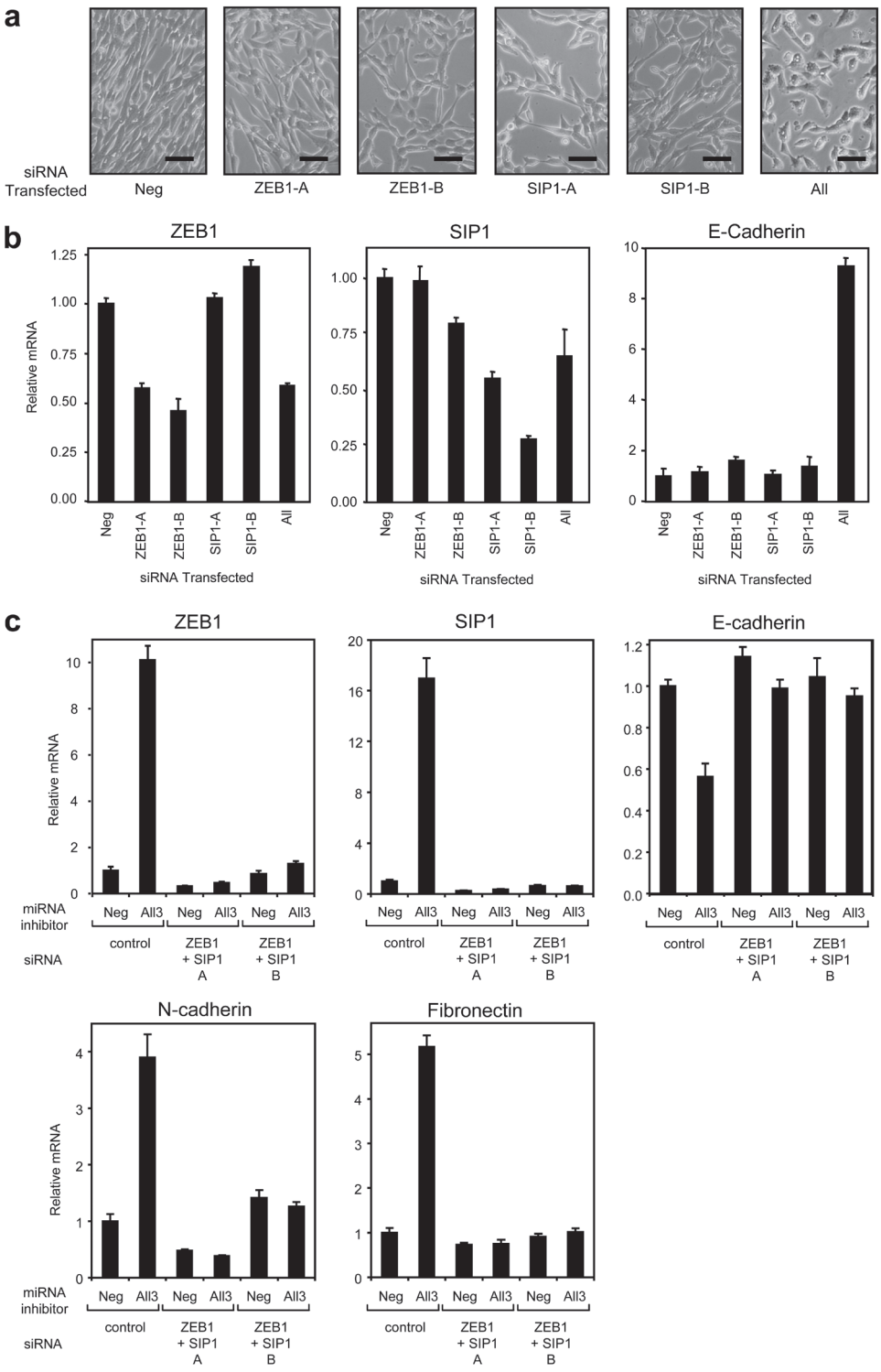

Figure S3 ZEB1 and SIP1 siRNA knockdown and their effect on MDCK-Pez cells, and EMT in MDCK cells mediated by microRNA inhibition. **(a)** Phase contrast microscopy of MDCK-Pez cells transfected with individual ZEB1 and SIP1 siRNAs (A and B) either separately or in combination (All) for 3 days. Comparisons were made with a non-targeting siRNA (Neg). Scale bars represent 50 μm. **(b)** Quantitation by real time PCR of ZEB1, SIP1 and E-cadherin after ZEB1 and or SIP1 knockdown. Transfections were carried out with 20 nM of each individual siRNA or 5 nM each of the four siRNAs (All). The data are taken from a representative experiment of three transfection experiments with PCR performed in triplicate and are shown ± s.e.m. (n=3). **(c)** Quantitation by real time PCR of EMT markers after concurrent miR-200a, miR-200b, and miR-205 inhibition (All3) or treatment with a negative control inhibitor (Neg), with or without ZEB1 and SIP1 knockdown. Successive transient transfections were performed at 3-4 day intervals for a total of nineteen days. The data are taken from two transfection experiments with qPCR performed in triplicate and are shown ± s.e.m. (n=6).

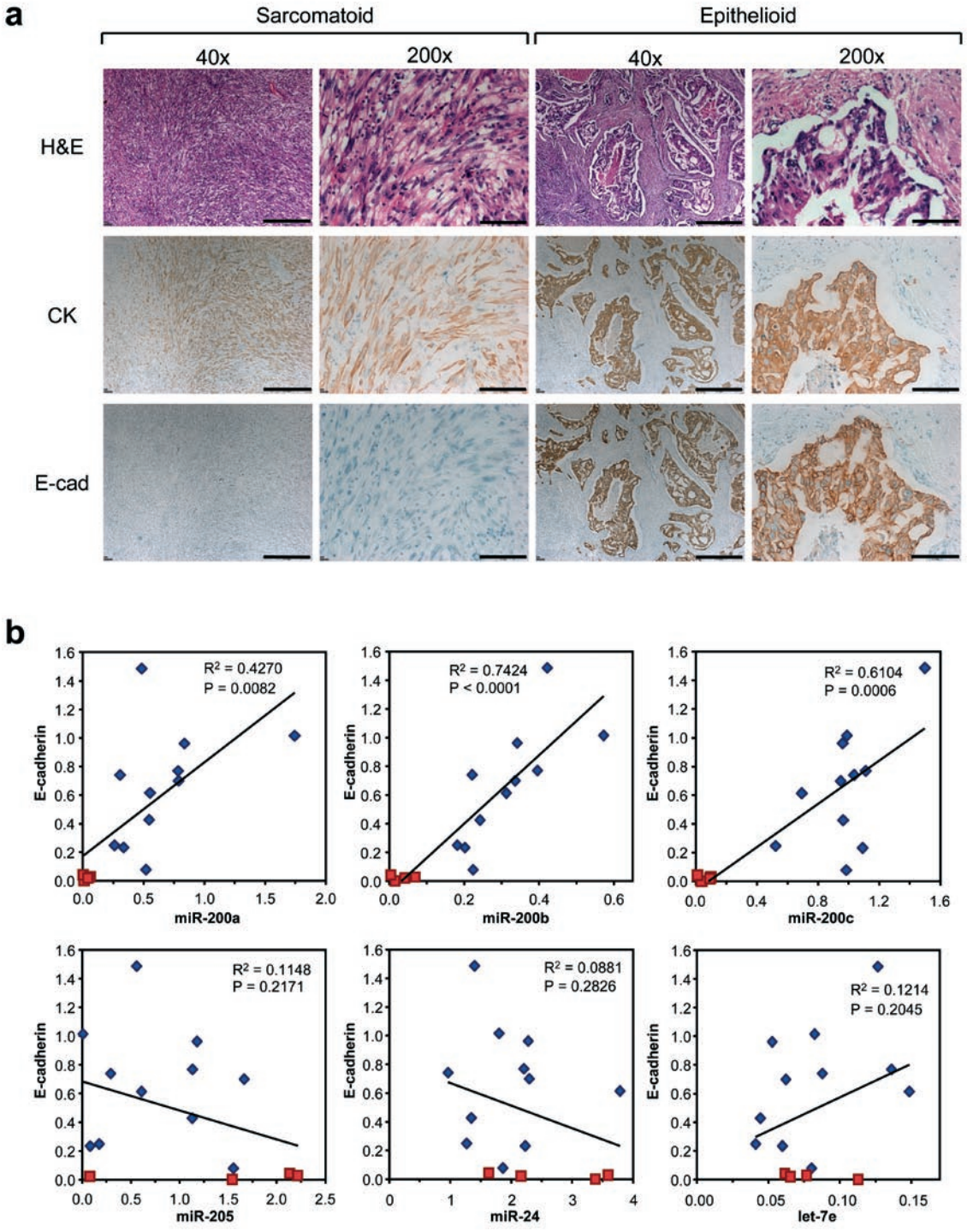

Figure S4 E-cadherin and microRNA levels in metaplastic and ductal tumours. **(a)** Representative photographs (taken at 40x and 200x) of the mesenchymal and the epithelial areas of a metaplastic breast cancer, from a histologic section stained with Haematoxylin and Eosin (H&E), with corresponding results of immunohistochemistry using cytokeratin (CK) and E-cadherin. In this specimen, discrete mesenchymal (sarcomatoid) and epithelial (epithelioid) regions were observed. The epithelioid region expressed E-cadherin and retained epithelial features, whereas the sarcomatoid region displayed a spindle shaped morphology and lacked E-cadherin. Despite the morphologic heterogeneity of metaplastic carcinoma, cytokeratin is often expressed in both the epithelioid and sarcomatoid regions, as in the example above, implying they are derived from epithelial origin. In addition, the patterns of metastasis of metaplastic tumours are similar to those of the more common types of breast cancer, so that for diagnostic and therapeutic purposes this tumour is regarded as a form of breast carcinoma, rather than a sarcoma. **(b)** Relationship between E-cadherin and microRNA levels in ductal and metaplastic tumours. Ductal tumours are coloured in blue and metaplastic in red. Pearson correlation coefficients are shown.

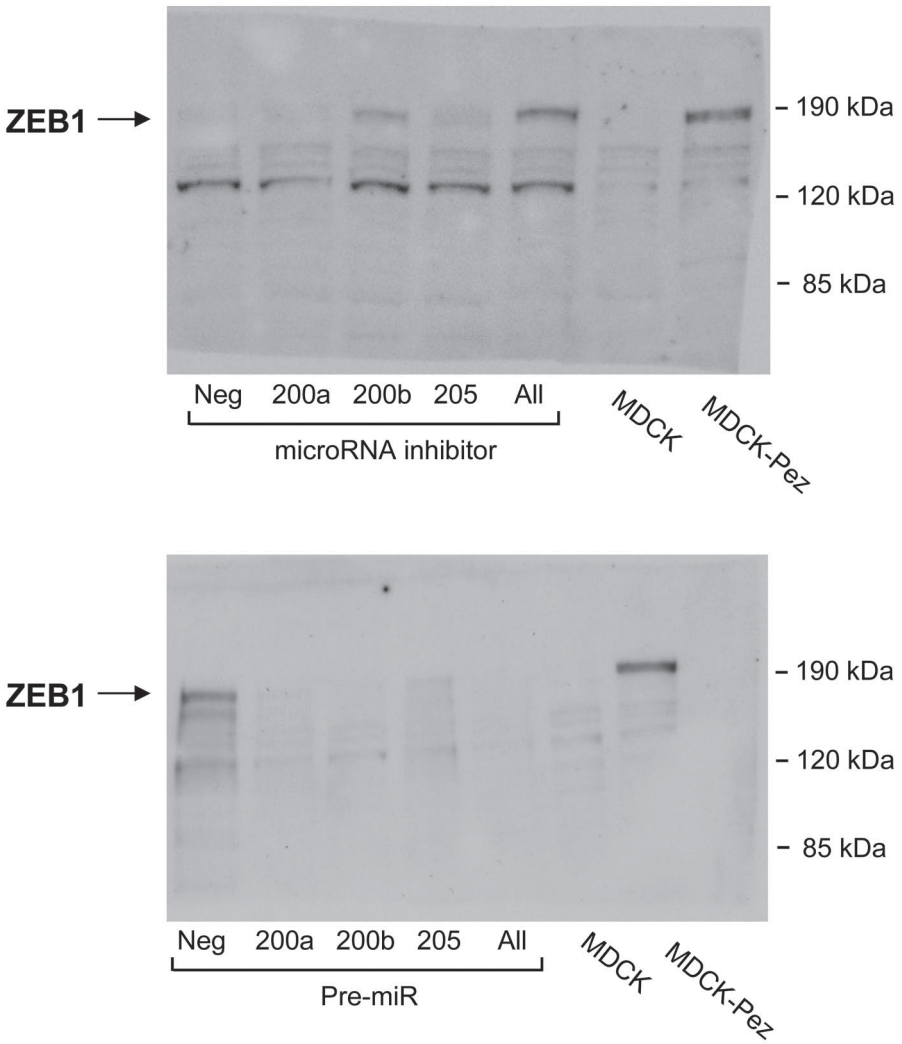

Figure S5 Full scan images of immunoblots shown in Figure 4. Membranes were cut according to the pre-stained molecular weight markers and probed with a ZEB1-specific antibody.

Table S1 Oligonucleotides used in real-time PCR, cloning, and knockdown studies. Primers are labelled by gene name, species (where appropriate) and direction (5' = for, 3' = rev). miRNA knockdown RNAs are 2'-O-methyl modified. Primers used for mutating the miR-200b seed regions in the ZEB1 and SIP1 3'UTRs are numbered according to their position from 5' to 3' with mutated bases marked in bold.

Table S1 Oligonucleotides used in real-time PCR, cloning, and knockdown studies

Primer Name	Sequence (5' to 3')
Real time PCR	
ZEB1 hum-dog for	TTCAAACCCATAGTGGTTGCT
ZEB1 hum-dog rev	TGGGAGATACCAAACCAACTG
SIP1 hum for	CAAGAGGCGCAAACAAGC
SIP1 hum rev	GGTTGGCAATACCGTCATCC
SIP1 dog for	CGGTCCAGAAGAAATGAAGG
SIP1 dog rev	TCCTCAAAGTCTGATGTGCAA
E-cadherin hum for	CCCACCACGTACAAGGGTC
E-cadherin hum rev	CTGGGGTATTGGGGGCATC
E-cadherin dog for	AAGCGGCCTCTACAACTTCA
E-cadherin dog rev	AACTGGGAAATGTGAGCACC
N-cadherin dog for	CAACTTGCCAGAAAACTCCAGG
N-cadherin dog rev	ATGAAACCGGGCTATCAGCTC
Fibronectin dog for	GCAACTCTGTGGACCAAGG
Fibronectin dog rev	CACTGGCACGAGAGCTTAAA
GAPDH hum for	ACCCAGAAGACTGTGGATGG
GAPDH dog for	CATCACTGCCACCCAGAAG
GAPDH hum-dog rev	CAGTGAGCTTCCCGTTCAG
ZEB1/SIP1 3'UTR cloning	
ZEB1 3'UTR for	CAACTAGTCAAAATAAATCCGGGTGTGC
ZEB1 3'UTR rev	TTACTAGTACAGCAGTTCAGGCTTGTTGA
SIP1 3'UTR for	ATACTAGTGGAGTTGGAGCTGGGTATTG
SIP1 3'UTR rev	ACACTAGTTGGAATCAGGATCAGTTGAGAA
ZEB1/SIP1 3'UTR mutagenesis	
ZEB1 miR-200b site 1 for	GCATCTGGCATTGTTTTATCTTATCAG**G**A**C**TATCACTCTTATGTTGGTTTATTCTTA
ZEB1 miR-200b site 1 rev	CGTAGACCGTAACAAAATAGAATAGTC**C**T**G**ATAGTGAGAATACAACCAAATAAGAAT
ZEB1 miR-200b site 2 for	ATTGGTAAACATATGCTAAATCCGCTTCAG**G**A**C**TTTATTATGTTTTTTAAAATGTGAGAACTT
ZEB1 miR-200b site 2 rev	TAACCATTTGTATACGATTTAGGCGAAGTC**C**T**G**AAATAATACAAAAAATTTTACACTCTTGAA
ZEB1 miR-200b site 3 for	TGTAAGTGCCATTTCTCAG**G**A**C**TTTCAAGGCTCTAACCCGC
ZEB1 miR-200b site 3 rev	ACATTCACGGTAAAGAGTC**C**T**G**AAAGTTCCGAGATTGGGCG
ZEB1 miR-200b site 4 for	ATTAACAACATTAGCTGATTTTTACCTATCAG**G**A**C**TATTTTATTTCTTTTAGTTTATAGATCTGTGC
ZEB1 miR-200b site 4 rev	TAATTGTTGTAATCGACTAAAAATGGATAGTC**C**T**G**ATAAAATAAAGAAAATCAAATATCTAGACACG
ZEB1 miR-200b site 5 for	CTGTATGTCTTCAAACCTGGCAG**G**A**C**TAATACCCTTCTTACTGACATAT
ZEB1 miR-200b site 5 rev	GACATACAGAAGTTTGGACCGTC**C**T**G**ATTATGGGAAGAATGACTGTATA
SIP1 miR-200b site 1 for	GGTGCCCGCACTACCATACATCAG**G**A**C**TTTTATTATTATTATTGTTATTCCT
SIP1 miR-200b site 1 rev	CCACGGGCGTGATGGTATGTAGTC**C**T**G**AAAATAATAATAATAACAATAAGGA
SIP1 miR-200b site 2 for	GCTCGCACTACAATGCATCAG**G**A**C**TATGATTCCTCTGTACTTTCC
SIP1 miR-200b site 2 rev	CGAGCGTGATGTTACGTAGTC**C**T**G**ATACTAAGGAGACATGAAAGG
SIP1 miR-200b site 3 for	CTTTGAAGCACCCATGTCAG**C**A**G**TAGAAGAATAGGCAGCAGTT
SIP1 miR-200b site 3 rev	GAAACTTCGTGGGTACAGTC**G**T**C**ATCTTCTTATCCGTCGTCAA
SIP1 miR-200b site 4 for	CTGTACTTTTTTGTTCATTAATTTTGTCAG**C**A**G**TACACCAAACTGTTTTTGCAACAAAAAAAT
SIP1 miR-200b site 4 rev	GACATGAAAAAACAAGTAATTAAAACAGTC**G**T**C**ATGTGGTTTGACAAAAACGTTGTTTTTTTA
SIP1 miR-200b site 5 for	TATTTCCTAATTTTATTTATTTCATACTGTAGTGTACAG**C**A**G**TATAGTTCTTCAATATATAGATATATTTTAGTAAAAAAG
SIP1 miR-200b site 5 rev	ATAAAGGATTAAAATAAATAAAGTATGACATCACATGTC**G**T**C**ATATCAAGAAGTTATATATCTATATAAAATCATTTTTTC
SIP1 miR-200b site 6 for	ATGACAAAAATCTTTCCTGAATTGTCTTTGTAAAAG**G**A**C**TATTGAATTTTCAATTTGTAATTTCTTTTGAAAATG
SIP1 miR-200b site 6 rev	TACTGTTTTTAGAAAGGACTTAACAGAAACATTTTC**C**T**G**ATAACTTAAAAGTTAAACATTAAAGAAAACTTTTAC
miR-200 family cluster cloning	
miR-200b-200a-429 for	AGAATTCCCACTCCGACCTAGTCCTC
miR-200b-200a-429 rev	AACCTAGGCTCCGGGTATCTGTGACTGTGAC
miR-200c-141 for	AGAATTCAGGGCTCACCAGGAAGTGT
miR-200c-141 rev	GTCCTAGGAGGGGTGAAGGTCAGAGGTT
miRNA knockdown	
2'-O-methyl miR-200a	UUGAACAUCGUUACCAGACAGUGUUAGAGUC
2'-O-methyl miR-200b	CGCCGUCAUCAUUACCAGGCAGUAUUAGAGA
2'-O-methyl miR-205	AUAGUCAGACUCCGGUGGAAUGAAGGACGAU
2'-O-methyl control	CAUCACGUACGCGGAAUACUUCGAAAUGACC
ZEB1 and SIP1 knockdown	
ZEB1-A siRNA	GCCAACAGUUGGUUUGGUATT
ZEB1-B siRNA	GCAUCCAAAGAACAAGAAATT
SIP1-A siRNA	CCUCUUGUCAUCUGUACUUTT
SIP1-B siRNA	GCAUGUAUGCAUGUGACUUTT