Supporting Information

Pro-metastatic GPCR Cd97 is a Direct Target of Tumor Suppressor microRNA-126

Ying Y. Lu, Michael J. Sweredoski, David Huss, Rusty Lansford, Sonja Hess, David A. Tirrell*

Lentivectors and plasmids. SparQ™ cumate switch inducible expression constructs, pCDH-EF1-CymR-T2A-Puro (QM200PA-2), pCDH-CuO-MCS-EF1-GFP dual promoter lentivector (QM511B-1), and customized construct pCDH-CuO-miR126-EF1-GFP were purchased from System Biosciences (SBI). pMIR-REPORT™ miRNA expression reporter (AM5795) and pcDNA™3.1(+) vector (V790-20) were from Invitrogen. *cd97* 3'-UTR was amplified from human genomic DNA using the following primers: AAT TAA CTA GTA GGC GCA TGG TTC TGG ACG GCC CAG (forward) and ATT TTT CAG TGT TGA CAC TTA AAA TTA AAC ACA TGC ATA CAG AAG AAA GCT TTA TA (reverse).

Lentiviral transduction. MDA-CymR cells stably expressing pCDH-EF1-CymR-T2A-Puro were generated by transducing MDA-MB-231 breast cancer cells (ATCC) with pCDH-EF1-CymR-T2A-Puro virus (SBI, QM200VA-2) followed by a 14-day selection with 10 μg/mL puromycin. To produce lentivirus containing inducible miR-126 constructs, HEK 293FT cells (Invitrogen) at 80% confluency were transiently transfected with either pCDH-CuO-miR126-EF1-GFP or pCDH-CuO-MCS-EF1-GFP using Lipofectamine 2000 (Invitrogen) along with the ViraPower™ Lentiviral Packaging Mix (Invitrogen, K4975-00), which consists of an optimized mixture of the three packaging plasmids, pLP1, pLP2, and pLP/VSVG. These plasmids supply the helper functions as well as structural and replication proteins in trans required to produce a recombinant lentivirus. The cell culture medium was collected 24, 48 and 72 h post transfection, filtered through a 0.8 μm filter, concentrated on a 30,000 MW cutoff spin column (Millipore), and stored at -80 °C. The MDA-CymR cells were infected with a 1:10 dilution of the concentrated lentiviral stocks before further passaging and subdivision by

FACS analysis based on the intensity of GFP fluorescence in order to obtain populations with similar numbers of transgene integrations. Cells were subdivided into populations with low, medium, and high fluorescence (Supporting Information Figure S1); cells with low fluorescence were used for further studies.

Fluorescence activated cell sorting. Cells were detached and pelleted by centrifugation (200 *g*, 3 min) and washed once with 1 mL Hank's balanced salt solution (0.2% BSA, 10 mM HEPES), and once with a cushion of FBS added to the bottom of the tube. Cells were then resuspended in 400 µL Hank's buffer before filtering through a 50 µm Nytex nylon mesh screen (Sefar). Cells were stored on ice until flow cytometry analysis on a FACSAria flow cytometer (BD Biosciences Immunocytometry Systems) at the Flow Cytometry Facility of the Beckman Institute at Caltech.

RT-PCR. Cumate solution (SBI) was added at concentrations of 0, 30, 150, 300, or 600 µg/mL to cells carrying inducible miR-126 constructs. At 2, 8, 24, and 72 h post induction, cells were washed with PBS, trypsinized, and harvested. Total RNA from these cells was isolated with mirVana miRNA isolation kits (Invitrogen, AM1560). Reverse transcriptase (RT)-PCR was performed according to manufacturer's protocols (TaqMan® microRNA Assays, Invitrogen, 4427975). Briefly, RNA concentrations were measured and 10 ng of RNA was used as template for reverse transcription using TaqMan® microRNA reverse transcription kit (Invitrogen, 4366596). Products from the RT reaction served as templates for quantitative PCR (qPCR) using a Model 7300 real-time PCR system (Applied Biosystems). Reagents used for qPCR included TaqMan® Universal PCR master mix without AmpErase® UNG (Invitrogen, 4324018) and primers from TaqMan® microRNA assays (Invitrogen, has-miR-126, ID000450; U6 snRNA

control, ID001973). Relative miR-126 expression (Supporting Information Figure S2) was calculated using the ΔΔCt method as previously described.[1]

Cell culture and amino acid incorporation. Cells were maintained in a 37 °C, 5% CO_2 humidified incubator chamber in customized Dulbecco's modified Eagle's medium (DMEM) (Invitrogen) supplemented with 10% dialyzed fetal bovine serum (Invitrogen), 50 U/mL penicillin, 50 μg/mL streptomycin, 0.2 mM methionine, 0.398 mM arginine, and 0.798 mM lysine or [$^{13}C_6$, $^{15}N_2$]L-lysine (Cambridge Isotopes). [2] One day prior to cumate addition, cells were passaged and seeded at a density of 3×10^6 in 150 cm^2 tissue culture flasks. Cumate was added to cells at a final concentration of 300 μg/mL. At 24 h post induction, cells were washed with PBS and fresh medium containing 1 mM Aha was added.

Enrichment of newly synthesized proteins. After incubation with Aha for 4 h, cells were washed, detached, and lysed using cold 4% SDS/PBS solution containing protease inhibitors (Roche). Lysates were clarified by centrifugation at 14,000 g for 5 min. Supernatant was collected and diluted to a final SDS concentration of 0.1%. Total protein concentration in each lysate sample was quantified using a BCA protein assay according to the manufacturer's protocol (Pierce). Lysates from light and heavy cultures were mixed at equal protein concentration. Affinity probes were attached using published conditions. [3] Excess probes were removed by precipitating proteins with cold acetone at -20 °C.

To enrich for newly synthesized proteins, pelleted protein was dissolved in 0.1% SDS/PBS solution and allowed to bind Streptavidin Plus UltraLink resin (Pierce) for 1 h

at RT. Samples were washed thoroughly with 1% SDS/PBS and 0.1% SDS/PBS and bound proteins were cleaved by incubation with 5% formic acid for 4 h. Newly synthesized proteins were then eluted with 0.1% SDS/PBS and concentrated for further preparation for MS.

Sample preparation for MS. Enriched proteins were separated by SDS-PAGE and each lane of the gel was cut into eight equal slices. Gel slices containing proteins were subjected to reduction with 7 mM DTT solution for 30 min at 50 °C before alkylation with 37 mM chloroacetamide for 30 min in the dark. In-gel digestion was carried out in 3 ng/μL LysC (Promega) solution at 37 °C overnight. Peptides were extracted from gel slices and desalted on C_{18} columns prior to MS analysis.

Mass spectrometry. Mass spectrometry experiments were performed on an EASY-nLC (Thermo Scientific) connected to a hybrid LTQ-Orbitrap Classic with a nanoelectrospray ion source (Thermo Scientific). Binding and separation of the peptides took place on a 15-cm silica analytical column (75 μm ID) packed in-house with reversed phase ReproSil-Pur C_{18AQ} 3 μm resin (Dr Maisch GmbH, Ammerbuch-Entringen, Germany). Samples were run for 60 min on a 2% to 30% acetonitrile gradient in 0.2% formic acid at a flow rate of 350 nL/min. The mass spectrometer was programmed to acquire data in a data-dependent mode, automatically switching between full-scan MS and tandem MS acquisition. Survey full scan MS spectra (from m/z 300 to 1,700) were acquired in the Orbitrap after the accumulation of 500,000 ions, with a resolution of 60,000 at 400 m/z. The ten most intense ions were sequentially isolated, and after accumulation of 5,000 ions, fragmented in the linear ion trap by CID (collisional energy 35% and isolation

width 2 Da). A reject mass list representing the most abundant streptavidin precursor ions was used. Precursor ion charge state screening was enabled and singly charged and unassigned charge states were rejected. The dynamic exclusion list was set for a 90-s maximum retention time and a relative mass window of 10 ppm; early expiration was enabled.

MS data analysis. Thermo raw data files were analyzed by MaxQuant (v 1.2.2.5) and searched against the International Protein Index (IPI) human database (75,710 sequences) supplemented with an in-house contaminant database (259 sequences) including human keratins and proteases. All default options were used except as follows: match between runs was enabled (2 min maximum), variable modifications on methionine of Aha (-4.9863), AlST (+835.4300), AlST-X (+93.0868), Oxidation (+15.9949), and protein N-terminal acetylation (+42.0106), fixed modification on cysteine of carbamidomethyl (+57.0215) and multiplicity of 2 with heavy label Lys8 (+8.014199). LysC digest was specified with up to two missed cleavages. Initial precursor mass tolerance was 7 ppm; however, MaxQuant calculates tighter individual precursor tolerances after recalibration. Fragment ion tolerance was 0.5 Daltons. Peptide, protein, and site false discovery rates were fixed at 1% using the target-decoy approach with a reversed database.[4] The minimum number of peptides for quantification was 1. Further data processing was performed to calculate ratios and standard errors of the ratios using in-house scripts described previously.[5] Briefly, hierarchical models are constructed of the data, where the overall ratio for each protein is the geometric mean of the biological replicates and the biological replicate ratio is the median of all of the peptide ratios in the replicate. The standard error of the overall protein ratio is calculated by estimating the global peptide

ratio standard error using pooled variance (calculated separately for peptide ratios based on requantified isotopic patterns) and using a bootstrapping procedure to resample at each level in the hierarchical model. Raw files are downloadable on Chorus.

Luciferase reporter assays. HEK293 cells (20,000 per well) were seeded in a 96-well plate and grown for 24 h. Cells were transfected with a plasmid cocktail containing 50 ng pRL-CMV (renilla luciferase control), 50 ng firefly luciferase construct outfitted with the 3'-UTR of interest, and either pcDNA™3.1(+)-miR126 or pcDNA™3.1(+)-miRex, using lipofectamine LTX (Invitrogen). At 48 h post transfection, luciferase activity was assayed using Dual-Glo® Luciferase Assay System according to manufacturer's instructions (Promega). Luminescence was measured on a microplate reader (Tecan).

Supplementary Results

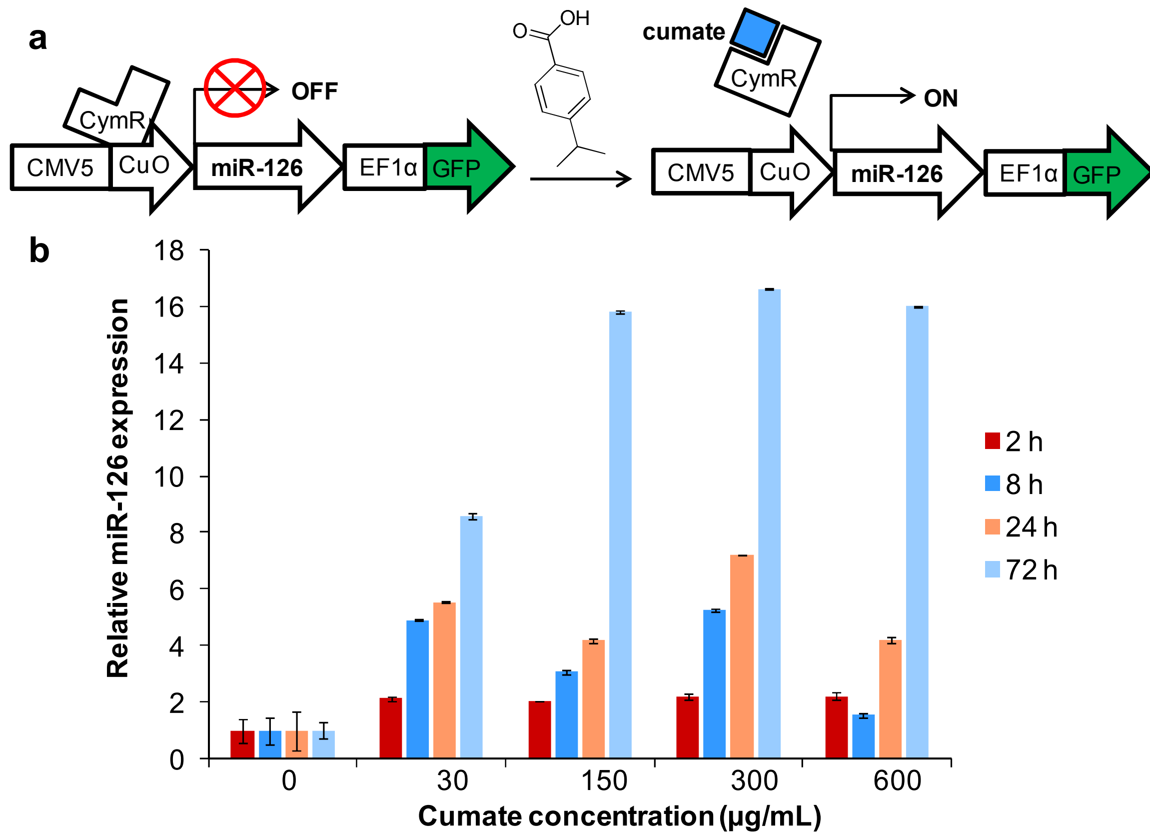

Figure S1. Cumate-inducible expression of miR-126 in MDA-CuO-miR cells. (**a**) Cumate induces a conformational change in the repressor CymR and allows controlled expression of miR-126. (**b**) RT-PCR analysis was performed to measure relative miR-126 expression at various time points after induction with concentrations of cumate ranging from 0 to 600 µg/mL. To achieve both robust and physiologically relevant miR-126 expression, we induced with 300 µg/mL cumate for 24 h for proteomic studies.

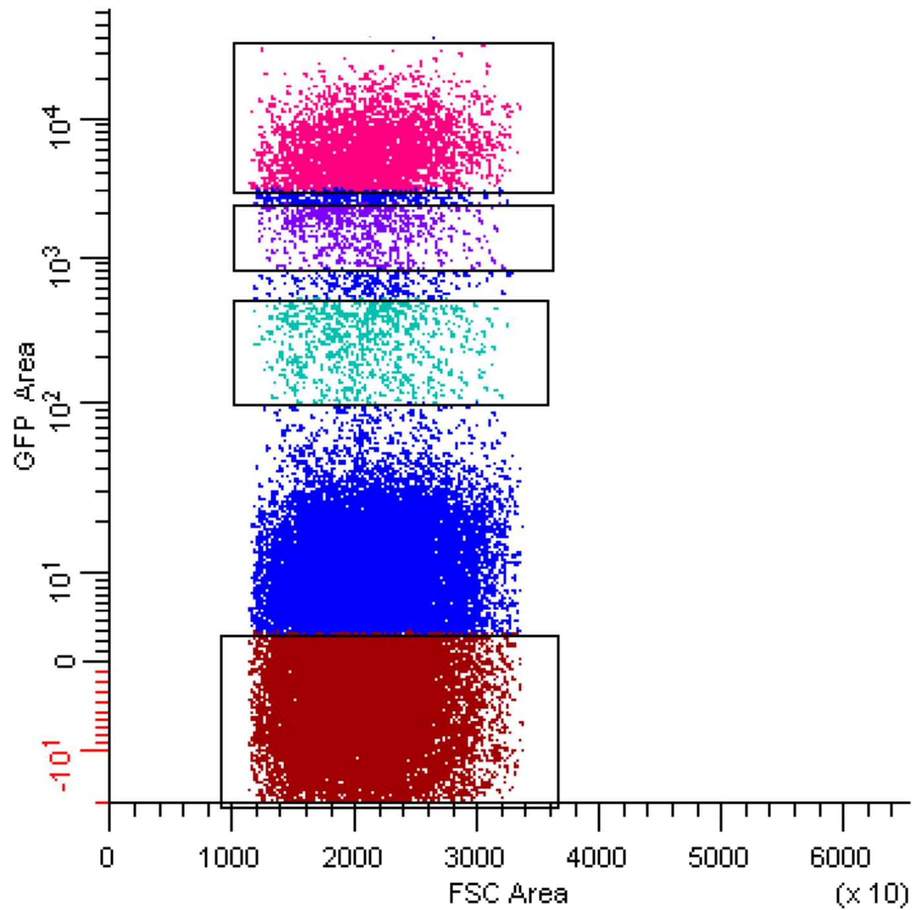

Figure S2. Fluorescence activated cell-sorting (FACS) analysis of MDA-CuO-miR cells. Cells were sorted according to GFP fluorescence to obtain cell populations with similar numbers of transgene integrations. Cell populations with low (green), medium (purple) and high (pink) fluorescence were collected; cells with low fluorescence were used in subsequent studies.

Gene name	Fold change (log_2)	Gene name	Fold change (log_2)
sdpr	-0.68	*etfa*	0.46
kif4	-0.45	*ctsl*	0.47
hsp27	-0.45	*lamb2*	0.47
cnn3	-0.43	*ndufa4*	0.48
cd97	-0.40	*antxr2*	0.49
eif4f	-0.34	*mmp14*	0.52
egfr	0.31	*asns*	0.55
cox6c	0.33	*etfb*	0.56
asph	0.39	*orca*	0.57
nt5	0.39	*ugdh*	0.63
ppil3	0.39	*psa*	0.71
ltb4dh	0.41	*ech1*	0.72
acs3	0.41	*cd49b*	0.82
cars	0.43	*tsc22d1*	0.86
dia4	0.43	*hmox1*	1.13
aars	0.45	*pgdh3*	1.16
msn	0.46		

Table S1. Proteins responsive to miR-126 over-expression were identified by a combined BONCAT-SILAC approach. Listed are proteins quantified by at least 6 independent peptide ratio measurements and exhibiting consistent up- or down-regulation by more than 20%.

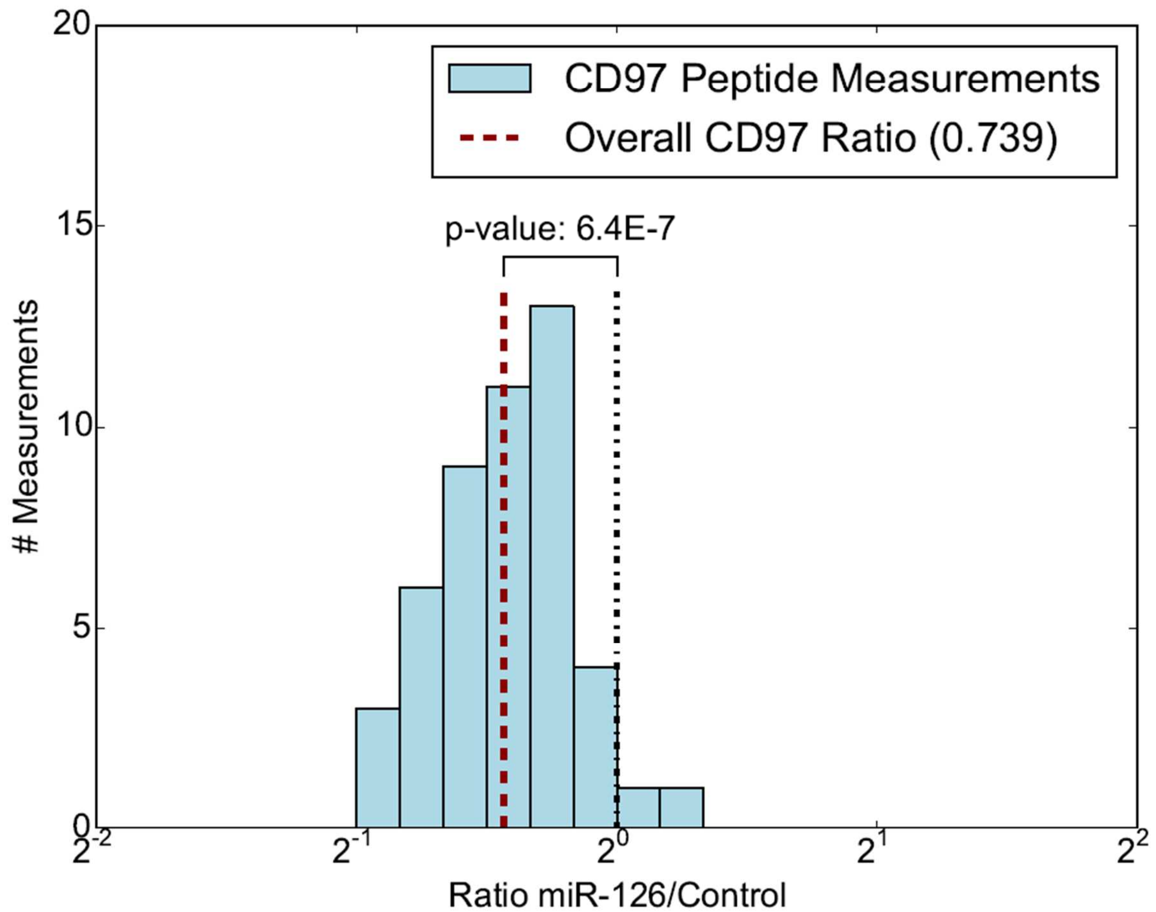

Figure S3. Histogram of independent CD97 peptide measurements. SILAC-labeled proteins from MDA-CuO-miR cells grown in the presence or absence of cumate were mixed at equal protein concentration and enriched using BONCAT prior to MS. Quantitative mass spectrometry, based on 48 independent peptide measurements, indicated down-regulation of CD97 upon cumate induction of miR-126 (*p*-value = 6.4×10^{-7} with the null hypothesis that CD97 is unchanged upon over-expression.

Oligonucleotides used in this study.

cd97 3'-UTR PCR primer fwd:

5'-AAT TA**A CTA GT**A GGC GCA TGG TTC TGG ACG GCC CAG-3' (restriction sites underlined)

cd97 3'-UTR PCR primer rev:

5'-TAT A**AA GCT T**TC TTC TGT ATG CAT GTG TTT AAT TTT AAG TGT CAA CAC TGA AAA AT-3'

Irs1 3'-UTR PCR primer fwd:

5'-AAT TA**A CTA GT**T GAC CTC AGC AAA TCC TCT TC-3' (restriction sites underlined)

Irs1 3'-UTR PCR primer rev:

5'-TAT A**AA GCT T**AC TCT CTC CAC CCA ACG TGA-3'

cd97-ex PCR1 fwd:

5'-GTG TTT GTG GAC GAA GTA CCG AAA GGT C-3'

cd97-ex PCR1 rev:

5'-***GCA GGC CTC TCT CAG GCA GT***G GCC CCA AGG-3' (overlapping region in bold italics)

cd97-ex PCR2 fwd:

5'-ACT GCC TG***A GAG AGG CCT GCC CTG CCT G***-3'

cd97-ex PCR2 rev:

5'-GCA AGG CGA TTA AGT TGG GTA ACG CC-3'

cd97-ex PCR3 fwd:

5'-AAT TA**A CTA GT**A GGC GCA TGG TTC TGG ACG GCC CAG-3'

cd97-ex PCR3 rev:

5’-TAT A**AA GCT T**TC TTC TGT ATG CAT GTG TTT AAT TTT AAG TGT CAA CAC TGA AAA AT-3’

cd97-M3-126:

5’-TGG GGC CAC TGC CTG AGG CTC ***TGC*** GTA CAG AGG CCT GCC CTG CCT-3’ (mutations in bold italics)

cd97-M3-126*:

5’-TCC CTC CAC CCT CCC TCC CTG ATC CCG T***CA T***CC ACC AGG AGG GAG TGG CAG CTA TAG TCT-3’

cd97-M3-R1:

5’-GGA GCC ACT GGT CCT GCT GCT GGC TGC C***AG A***CT GCT CCA CCT TGT GAC CCA GGG TGG GGA-3’

cd97-M3-R2:

5’-TGG GGC TCA GCT TCC CTC TTA AGC TAA G***TG A***GA TGT CAG AGG CCC CAT GGC GAG GCC CCT-3’

miR-126 with flanking region sequence:

5’-

GCTAGCGAATTCGCCCTTGTGGACATTGCCGTGTGGCTGTTAGGCATGGTGGG
GGGCACTGGAATCTGGGCGGAAGGCGGTGGGGACTCCCTCTCCAGGGAGGG
AGGATGGGGAGGGAGGATAGGTGGGTTCCCGAGAACTGGGGGCAGGTTGCC
CGGAGCCTCATATCAGCCAAGAAGGCAGAAGTGCCCCGTCCCGGGGTCCTGT
CTGCATCCAGCGCAGCATTCTGGAAGACGCCACGCCTC***CGCTGGCGACGGGA***
CATTATTACTTTTGGTACGCGCTGTGACACTTCAAACTCGTACCGTGAGTAATA

ATGCGCCGTCCACGGCACCGCATCGAAAACGCCGCTGAGACCTCAGCCTTGAC

CTCCCTCAGCGTGGCCGGGACCCTGAGCCTCTGCGCAGAGCCACCCGCCCCG

ACGTACTTAGGCGGCATAGCCCTGAGACCTCTGGCCAGCGCCAGGCAGGCAG

CGGGGGCGGCAGAGGCCTGGGCCTGAGTCTTCTGGCTCTGCCTCGCGGCCGC

-3' (pre-miR-126 sequence in bold italics)

(1) Livak, K. J., and Schmittgen, T. D. (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method, *Methods 25*, 402-408.

(2) Ong, S. E., and Mann, M. (2006) A practical recipe for stable isotope labeling by amino acids in cell culture (SILAC), *Nat Protoc 1*, 2650-2660.

(3) Hong, V., Presolski, S. I., Ma, C., and Finn, M. G. (2009) Analysis and optimization of copper-catalyzed azide-alkyne cycloaddition for bioconjugation, *Angew Chem Int Ed Engl 48*, 9879-9883.

(4) Cox, J., and Mann, M. (2008) MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification, *Nat Biotechnol 26*, 1367-1372.

(5) Pierce, N. W., Lee, J. E., Liu, X., Sweredoski, M. J., Graham, R. L., Larimore, E. A., Rome, M., Zheng, N., Clurman, B. E., Hess, S., Shan, S. O., and Deshaies, R.

J. (2013) Cand1 promotes assembly of new SCF complexes through dynamic exchange of F box proteins, *Cell 153*, 206-215.