

Supporting Information (SI Appendix)

MicroRNA-224 promotes tumor progression in non-small cell lung cancer

Ri Cui[#], Wei Meng[#], Hui-Lung Sun, Taewan Kim, Zhenqing Ye, Matteo Fassan, Young-Jun Jeon, Bin Li, Caterina Vicentini, Yong Peng, Tae Jin Lee, Zhenghua Luo, Lan Liu, Dongyuan Xu, Esmerina Tili, Victor X. Jin, Justin Middleton, Arnab Chakravarti, Tim Lautenschlaeger*, Carlo M. Croce*

SI Materials and Methods

SI References

SI Figures

Fig. S1. MiR-224 expression in NSCLC and its clinical implications.

Fig. S2. MiR-224 is involved in lung cancer cell proliferation and migration.

Fig. S3. *TNFAIP1* and *SMAD4* expressions in NSCLC.

Fig. S4. TNFAIP1 and SMAD4 are targets of miR-224.

Fig. S5. MiR-224 promoter methylation modulates miR-224 expression.

Fig. S6. ERK signaling modulates miR-224 expression.

Fig. S7. Representative pictures of cell migration assay and animal studies using miR-224 manipulated lung cancer cells.

SI Tables

Table S1. Identification of differentially expressed miRNAs between N0 and N1+ groups of patients

Table S2. Primers for 3' UTR mutagenesis and SMAD4 cloning

Table S3. miR-224 expression from TCGA dataset and methylation probes in the promoter

region of miR-224 from TCGA Illumina Infinium Human DNA Methylation 450k beadchip in NSCLC

SI Materials and Methods

Genome-wide miRNA sequencing and data analysis

As described previously (1), small RNAs were purified from fresh frozen tissue samples (6 with N0 stage and 4 with N1+ stage lung ADC) and ligated with 5' and 3' adapter mix using RNA ligase. Ligated products were amplified by PCR using SOLID Small RNA Expression Kit (Applied Biosystems) and applied for Applied Biosystems SOLID next generation high-throughput sequencing system for data acquisition. Expression analysis of miRNA sequencing data was conducted using the R/Bioconductor package EdgeR as described previously (1).

Plasmid construction, cell lines and reagents

The human pre-miRNA expression construct Lenti-miR-224 vector and human miRZip-224 anti-miR-224 miRNA construct were purchased from System biosciences. pLightSwitch empty, TNFAIP1-3'UTR, and SMAD4-3'UTR vectors were ordered from Active Motif. Mutations were generated by using the QuickChange II XL Site-Directed Mutagenesis Kit (Stratagene). To generate SMAD4 expressing constructs, the ORFs were amplified by PCR and cloned into pCDH-CMV-MCS-EF1-Neo vector at NheI and NotI restriction enzyme sites (System biosciences). The primers used in this study were shown in Table S2. shRNA control (SHC001) and shTNFAIP1 (TRCN0000156906) were purchased from Sigma. siRNA against c-Jun and control siRNA were purchased from Dharmacon. The cell lines used in this study were purchased from the American Type Culture Collection (ATCC). Human lung cancer cell lines H1299, H1573, A549 and H460 were maintained in RPMI1640 medium containing 10% FBS

and 100U/ml penicillin-streptomycin. 293T cells were cultured in DMEM medium supplemented with 10% FBS and 100U/ml penicillin-streptomycin. The Phorbol 12-myristate 13-acetate (PMA) and 5-Aza-deoxycytidine were purchased from Cayman chemicals and Sigma-Aldrich, respectively. Antibodies against Histon H3 (D1H2), cleaved PARP1 (Asp214) and SMAD4 were purchased from Cell signaling technology. Anti-vinculin antibody and anti-TNFAIP1 antibody were ordered from Sigma-Aldrich and Abcam, respectively.

Virus infection and transfection

The pre-miR224 expression construct, miRZip-224 anti-miR-224 construct and control vector were packaged with pPACKH1 Lentivector Packaging Plasmid mix (System Biosciences) in a 293T packaging cell line. The Transdux reagent (System Bioscience) was used for virus transduction, and infected cells were selected by fluorescence-activated cell sorting (FACS) analysis (FACSCalibur, BD Bioscience). Transfection of SMAD4 expressing vector and shRNA against TNFAIP1 were carried out with Lipofectamine LTX according to the manufacturer's instruction (Invitrogen) and transfected cells were selected by neomycin or puromycin, respectively. siRNA against C-jun was transfected to cells by using Lipofectamine RNAiMAX according to the manufacturer's instruction (Invitrogen).

Quantitative Real-time PCR

Total RNAs were extracted using TRIzol Reagent (Invitrogen) according to the manufacturer's instruction. Expression of miRNA was quantified by qRT-PCR with TaqMan miRNA Reverse Transcription Kit (Applied Biosystems). Small endogenous nucleolar U6 snRNA was used as control for normalization of miRNA. TaqMan gene expression assays for TNFAIP1 and SMAD4 were purchased from Applied Biosystems to determine their expression. β 2M was used as control for normalization of mRNA expression. cDNAs were synthesized from 3 μ g total RNAs

using SuperScript First-Strand Synthesis System for RT-PCR (Invitrogen). All reactions were conducted in triplicates.

Cell migration, invasion and proliferation assays

In vitro cell migration and invasion assays were conducted using Boyden chambers (BD bioscience) that use 8-mm micropore membranes without Matrigel (for migration assay) or with Matrigel (for invasion assay) as previously described (1). Both assays were carried out according to the manufacturer's instructions. *In vitro* cell growth rates were assessed by Cell Counting Kit8 (Dojindo) as described previously (2). Briefly, two thousand cells in 2% RPMI 1640 medium were added to each 48 well plate in quintuplex, and cultured at 37°C for 4 days. The number of cells was quantified using a Cell Counting Kit8.

Flow cytometry analysis

For cell cycle analysis, the cells were fixed with 70% ethanol at -20°C overnight, washed with PBS and resuspended in PBS containing 50ug/ml propidium iodide, 100ug/ml RNase A, 0.05 % Triton X-100. The cells were incubated at 37°C for 40 min and analyzed by FACS Calibur Flow Cytometer (BD Biosciences).

Western blot analysis

The cells were lysed with RIPA buffer (25mM Tris-HCl (pH 7.6), 150mM NaCl, 1% NP-40, 1% sodium deoxycholate, 0.1% SDS) supplemented with Protease/Phosphatase inhibitor Cocktail (Cell Signaling Technology), and separated on 4-20% Mini-protein TGX Gels (Bio-Rad). The cytosol and nuclear proteins were extracted by using NE-PER Nuclear and Cytoplasmic extraction Reagents (Thermo). After SDS-PAGE, the proteins were electrotransferred to Immun-Blot PVDF membrane (Bio-Rad). Then, Membranes were blocked with 5% BSA in Tris-Buffered Saline with Tween 20 (TBST) buffer and incubated with primary antibody in 3% BSA

in TBST, followed by incubation with appropriate horseradish peroxidase (HRP)-conjugated secondary antibody. Specific proteins were detected using the enhanced chemiluminescence system (GE Healthcare).

In situ hybridization of miR224

In situ hybridization (ISH) was performed as previously described (3). The *miR-224* probe was tagged with 5' digoxigenin and LNA modified (Exiqon). The probe–target complex was detected using an antidigoxigenin-alkaline phosphates conjugate and nitro-blue tetrazolium and 5-bromo-4-chloro-3'-indolyphosphate as the chromogen. Cases were classified according to cytoplasmic *miR-224* intensity as: negative= negative or faint expression in most cells; low expression= low expression in most cells or moderate expression in <50% of the cells; high expression= moderate to strong expression in most cells.

Isolation of cells from normal lung tissues from lung cancer patients

Normal lung tissues were obtained distal to tumor sites when the surgical specimen was large enough to include both tumor and normal lung margin from two lung cancer patients. The cell isolation was processed within 2 hours after surgical resection as previously described (4) with some modification. Briefly, the tissues were minced with scissors until ~1mm² or smaller and digested with 1×collagenase (Stemcell Technologies) in HBSS at 37°C for 4 hours with shaking. Following incubation, using 18 gauge needle to further dissociate undigested pieces of tissues. The tissue slurry was filtered with 10µm filters and overlay onto the 40%-70% percoll gradient. After centrifugation at 2500 rpm for 20min, collect cells from the 40%-70% percoll interface. The cells were maintained with RPMI1640 medium containing 20% FBS and 100U/ml penicillin-streptomycin. We isolated two cell lines from normal lung tissues and designated them as 469NAT and 485 NAT, respectively.

SI References

1. Meng W, *et al.* (2013) MicroRNA-31 predicts the presence of lymph node metastases and survival in patients with lung adenocarcinoma. *Clin Cancer Res* 19(19):5423-5433.
2. Cui R, *et al.* (2009) Osteopontin is involved in the formation of malignant pleural effusion in lung cancer. *Lung Cancer* 63(3):368-374.
3. Fassan M, *et al.* (2011) MicroRNA expression profiling in human Barrett's carcinogenesis. *Int J Cancer* 129(7):1661-1670.
4. Cook JA, *et al.* (1991) Cellular glutathione and thiol measurements from surgically resected human lung tumor and normal lung tissue. *Cancer Res* 51(16):4287-4294.

SI Figures

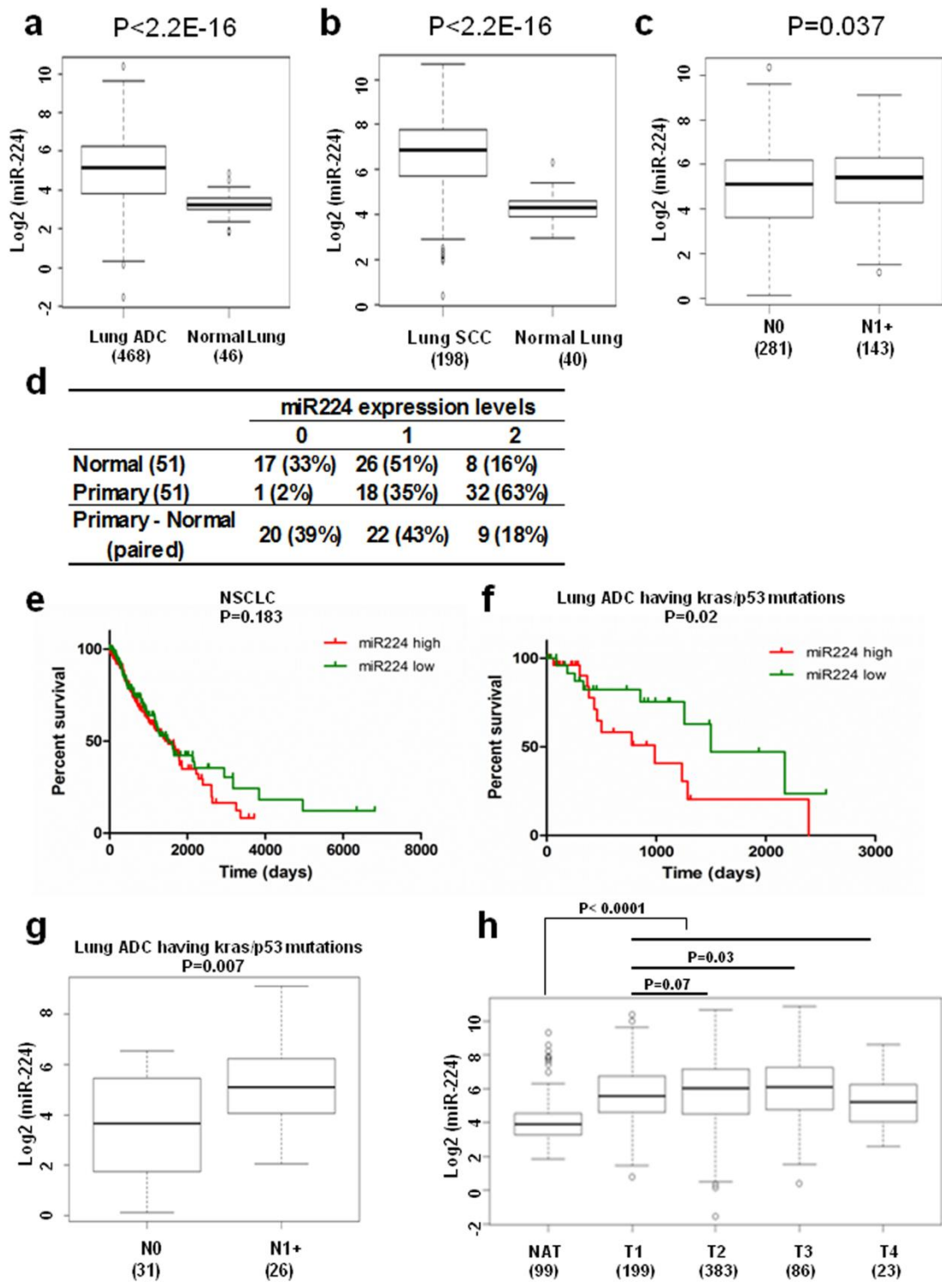


Fig.S1. MiR-224 expression in NSCLC and its clinical implications. The miR-224 expression data for lung ADC (a) and lung SCC (b) were acquired from TCGA miR-seq dataset. Of the evaluable 468 primary lung ADC patients, 46 patients had matched normal adjacent tissues. Of the evaluable 198 primary lung squamous carcinoma patients, 40 patients had matched normal adjacent tissues. (c) miR-224 expression in 281 N0 patients versus 143 N1+ patients from TCGA lung ADC dataset. (d) Summary of tissue micro array data for lung cancer and corresponding normal tissue. Paired analysis represents subtraction of NAT staining score from matched lung cancer tissue scores. (e) Kaplan Meier survival analysis for 642 available TCGA NSCLC patients. The p-value was calculated from Cox proportional hazards model. (f) Cox proportional hazards model on 59 lung ADC patients who have TP53 and/or KRAS mutations (mutation data were downloaded from www.cbioportal.org, Lung Adenocarcinoma TCGA Provisional). (g) miR-224 expression in 31 N0 patients versus 26 N1+ patients from lung ADC patients who have TP53 and/or KRAS mutations. (h) miR-224 expression according to T stage in the TCGA dataset.

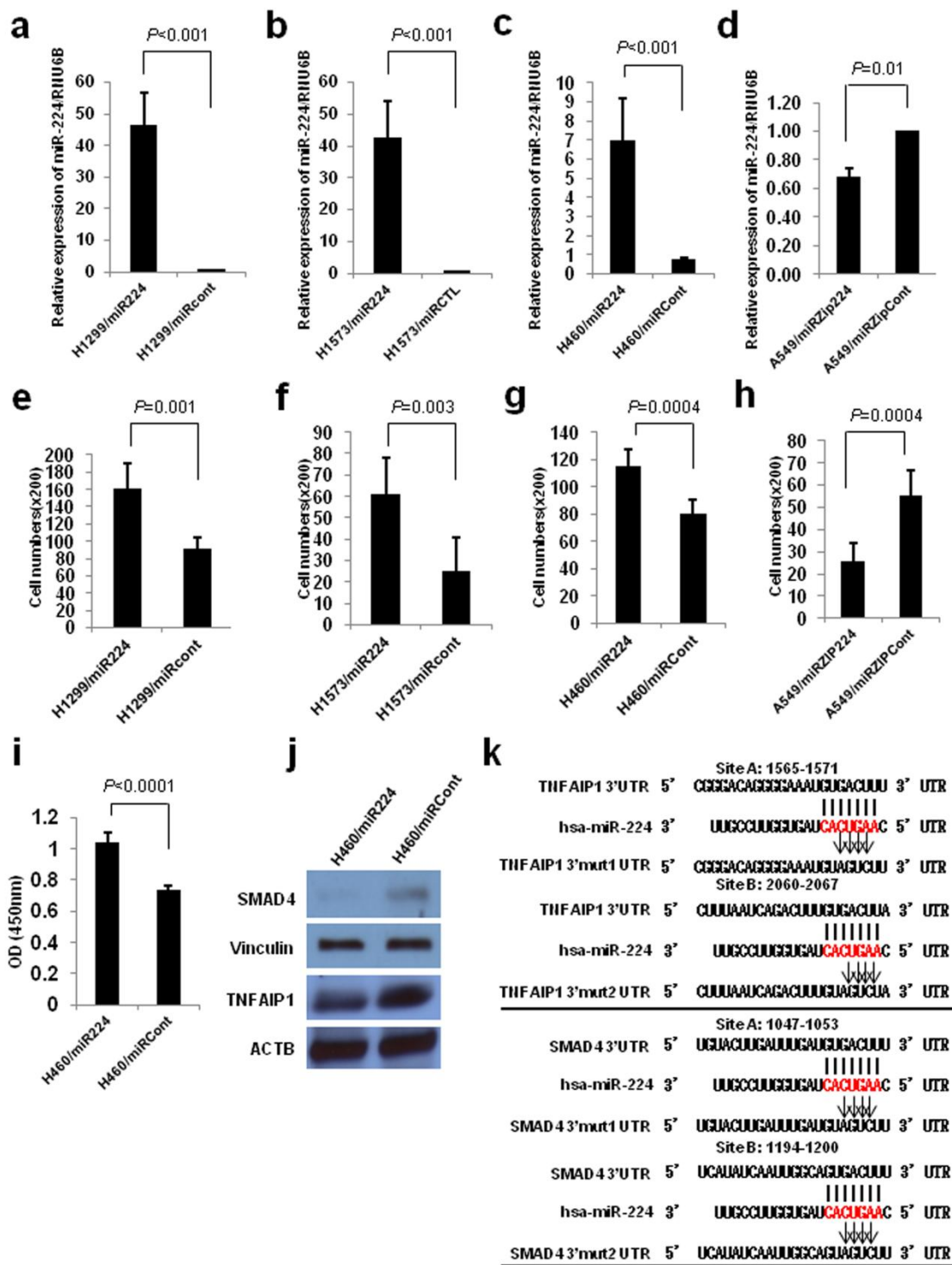
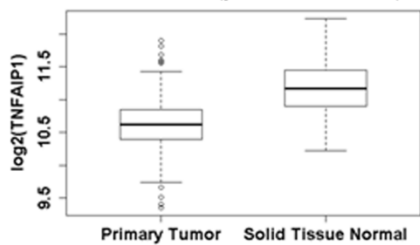
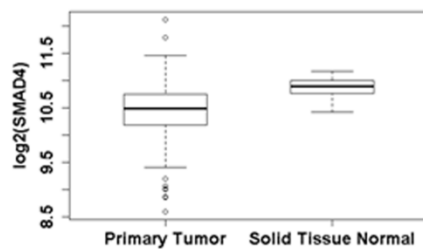


Fig. S2. MiR-224 is involved in lung cancer cell proliferation and migration. (a-c) Expressions of *miR-224* following infection with Lenti-miR vector containing miR-224 precursor into H1299 (a), H1573 (b), and H460 (c) were confirmed by TaqMan real-time PCR. (d) Expression of *miR-224* following infection with miRZip-224 anti miR-224 construct was confirmed by qRT-PCR. (e-g) Migrated cell counts after overexpressing miR-224 in H1299 (e), H1573 (f), and H460 (g) cells. (h) Migrated cell counts after knockdown miR-224 in A549 cells. (i) Cell proliferation assay for miR-224 overexpressing H460 cells. (j) Western blot analyses to measure SMAD4 and TNFAIP1 protein levels in H460 cells infected with lentivirus expressing premiR-224 or control vector. (k) Schematic diagram showing matched sequences between the seed sequence of *miR-224* and 3'UTRs of TNFAIP1 and SMAD4. The arrows indicate the mutagenesis nucleotides.

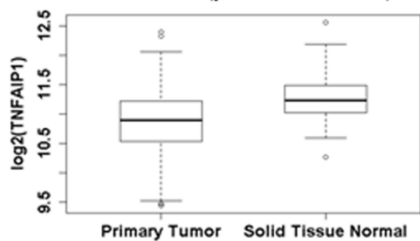
a Lung ADC
TNFAIP1 (p-value=4.92E-15)



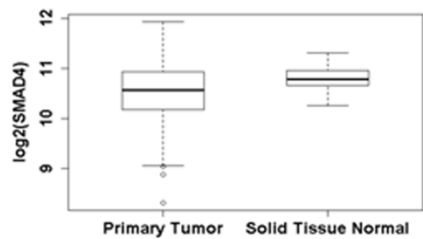
b Lung ADC
SMAD4 (p-value=2.2E-16)



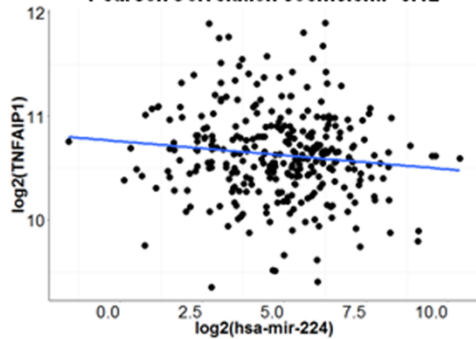
c Lung SCC
TNFAIP1 (p-value=1.79E-07)



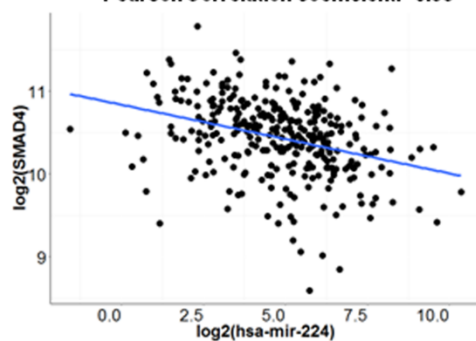
d Lung SCC
SMAD4 (p-value=6.20E-09)



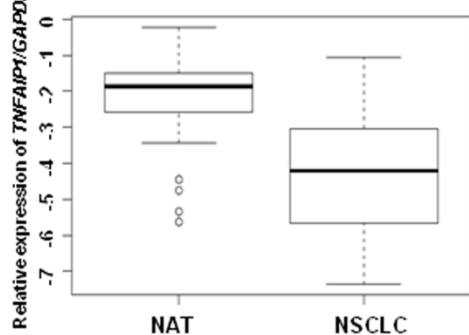
e Lung ADC (P value =0.038)
Pearson Correlation coefficient: -0.12



f Lung ADC (P value= 3.23E-09)
Pearson Correlation coefficient: -0.33



g P= 5.23E-14



h P= 2.67E-06

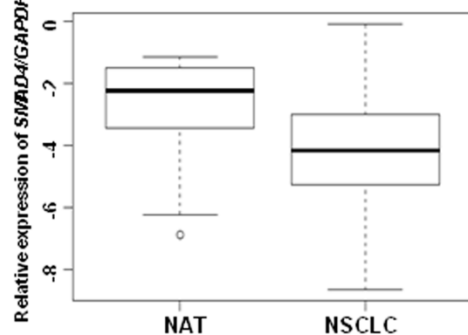


Fig. S3. *TNFAIP1* and *SMAD4* expressions in NSCLC. The *TNFAIP1* and *SMAD4* expression data were obtained from TCGA RNA-seq data containing a total of 334 lung ADC patients with 57 matched normal adjacent tissue (NAT) samples, and 349 lung SCC patients with 51 matched NAT samples. In lung ADC, *TNFAIP1* (a) and *SMAD4* (b) expression were dramatically reduced in primary tumor tissues. In lung SCC, a similar reduction of *TNFAIP1* (c) and *SMAD4* (d) expression levels were observed. (e and f) *SMAD4* and *TNFAIP1* expression from TCGA RNA-seq data and *miR-224* expression from miR-seq data were examined for correlation between *miR-224* expression and its target genes in lung ADC dataset (n = 306). (g and h) *TNFAIP1* and *SMAD4* expressions in 63 NSCLC patients with 48 matched NATs from OSU cohort.

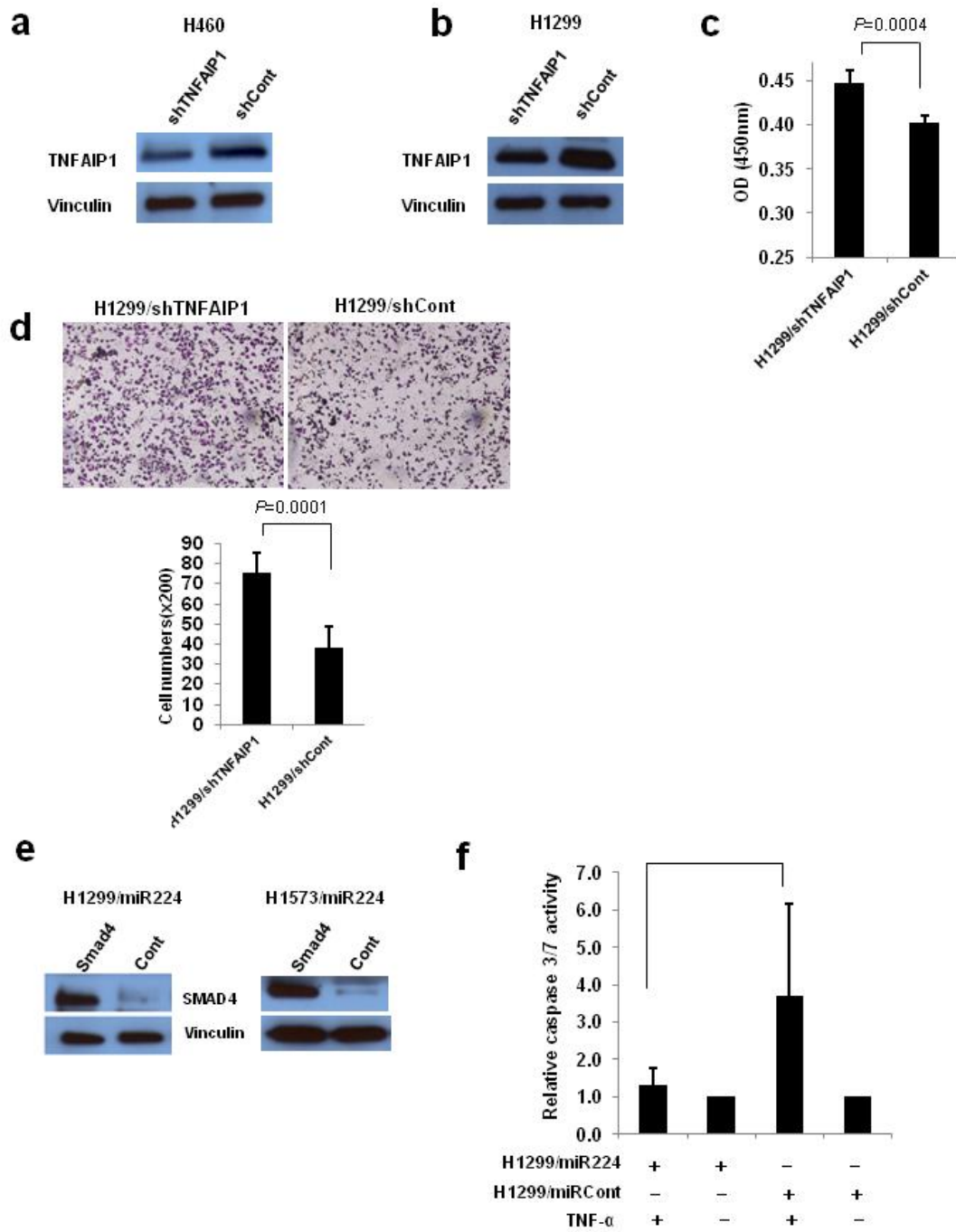
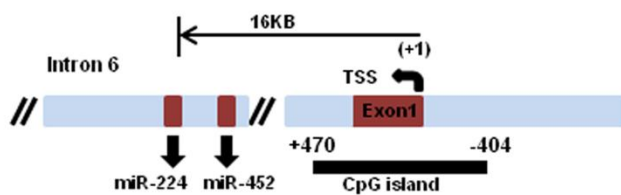
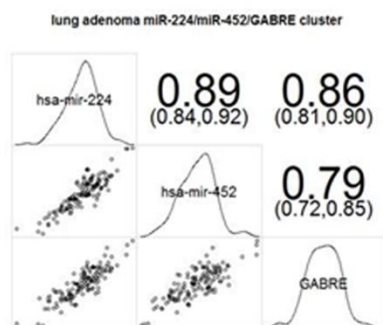


Fig. S4. TNFAIP1 and SMAD4 are targets of miR-224. (a and b) Western blot analyses of TNFAIP1 knockdown H460 (a) and H1299 (b) lung cancer cells. (c) Cell proliferation assay for TNFAIP1 knockdown H1299 lung cancer cells. The cell growth rates were measured by cell counting kit 8. (d) Cell migration assay for TNFAIP1 knockdown H1299 cells using transwell membranes. The average counts were derived from six random microscopic fields. (e) Western blot analysis of miR-224 overexpressing cells transfected with SMAD4 (do not contain 3'UTR) or empty vector. (f) Caspase 3/7 assay after TNF- α treatment in miR-224 overexpressing H1299 cells.

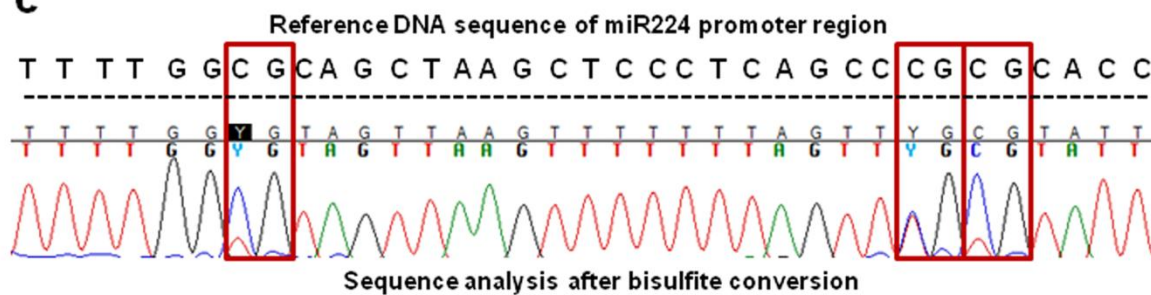
a GABRE genomic region



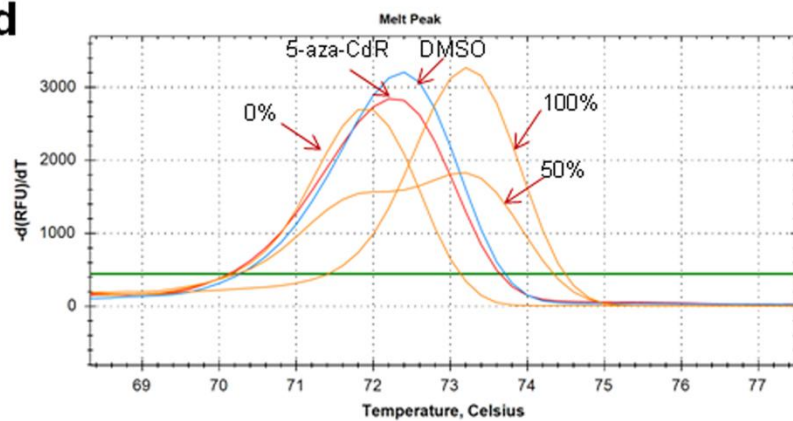
b



c



d



e

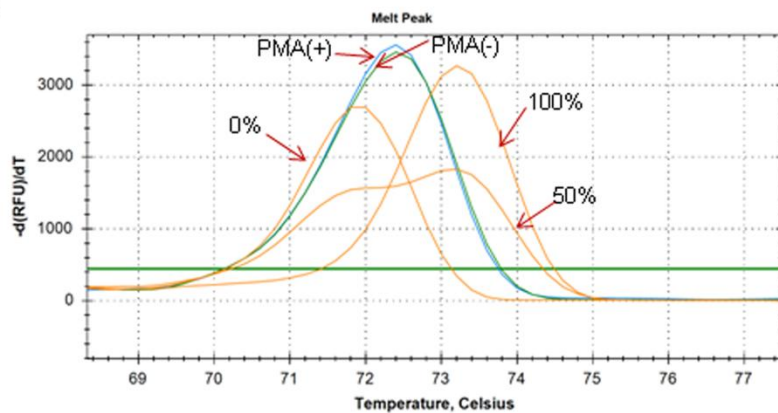


Fig. S5. MiR-224 promoter methylation modulates miR-224 expression. (a) Schematic structure of *miR224~miR452~GABRE* genomic locus (not to scale). *miR-224* and *miR452* are located in intron 6 of *GABRE*. TSS indicated putative transcription start site. Black box shows CpG island which expand exon1 of *GABRE*. (b) The normalized expression values of *miR-224*, *miR-452* and *GABRE* were extracted from TCGA dataset for Pearson correlation analysis between β -values of each other. (c) Bisulfate modified genomic DNA was amplified by *miR224* promoter specific primers and subject to sequencing analysis. (d) MS-HRM analyses of 485NAT cells with and without 5-aza-CdR treatment. (e) MS-HRM analyses of 485NAT cells with and without PMA treatment.

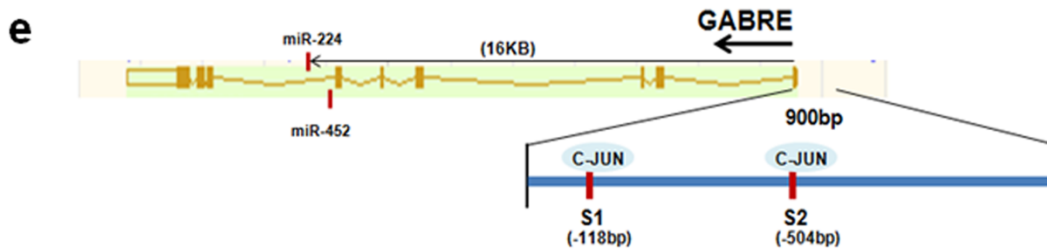
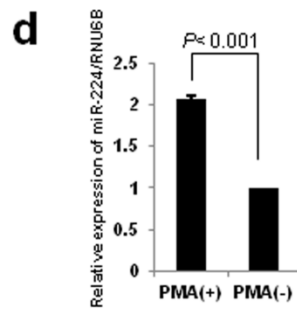
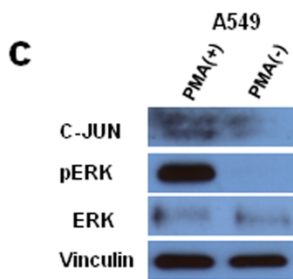
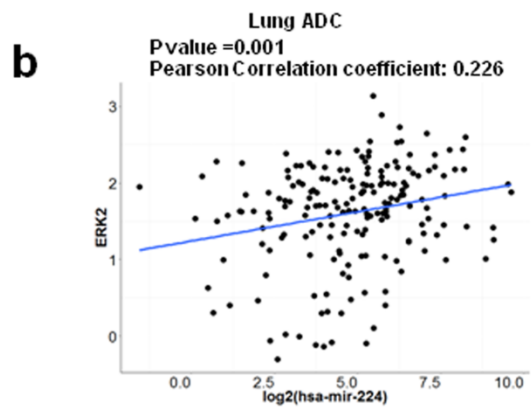
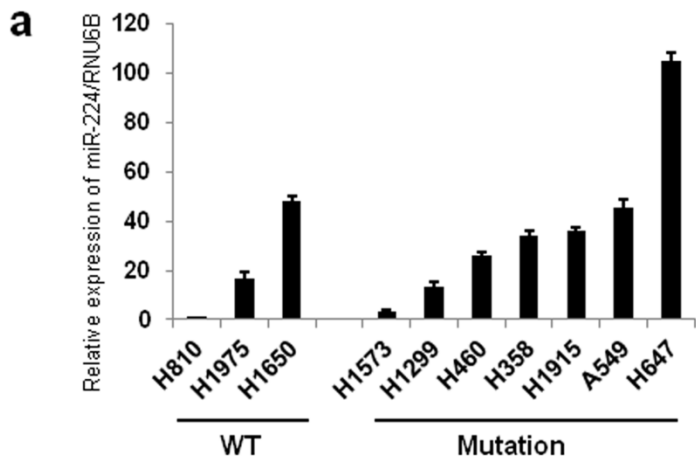


Fig. S6. ERK signaling modulates miR-224 expression. (a) miR-224 expression in different lung cancer cell lines. The cell lines were divided two sub-groups, RAS wild type and RAS mutation groups. The RAS gene status was collected from http://cancer.sanger.ac.uk/cancergenome/projects/cell_lines/. (b) Total ERK2 expression from TCGA RPPA data and *miR-224* expression from miR-seq data were examined for correlation between miR-224 expression and ERK2 proteins in lung ADC dataset (n = 188). Western blot analysis (c) and qRT-PCR (d) for A549 cells with and without PMA treatment to measure miR-224 expression. (e) Schematic diagram of *miR224~miR452~GABRE* genomic locus. *MiR-224* is located in intron6 of *GABRE*. Two C-jun binding sites (red line) were located in -118 bp and -504 bp upstream of *GABRE* transcription start site.

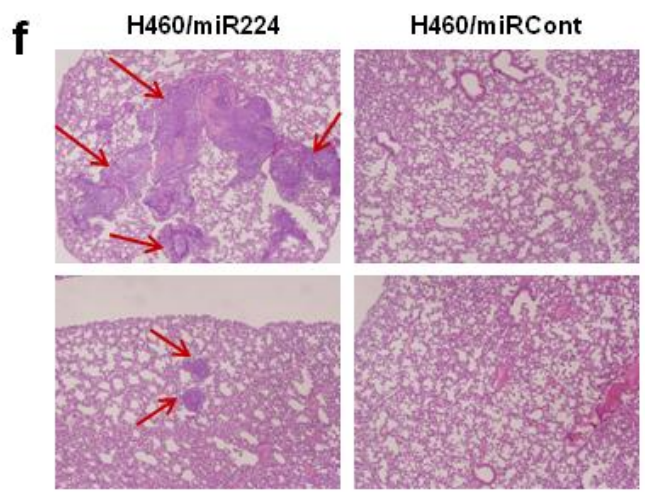
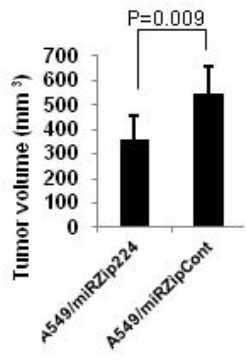
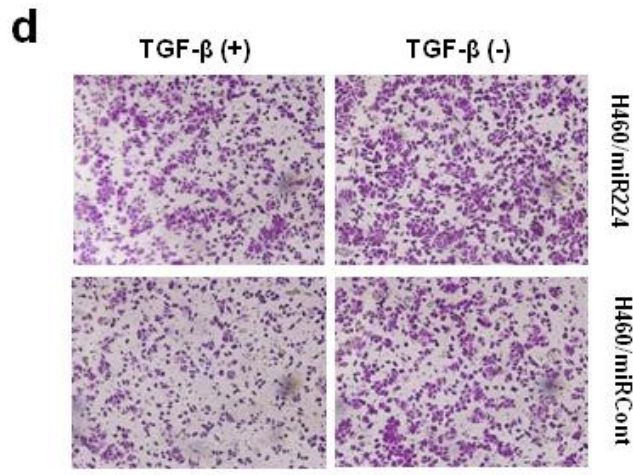
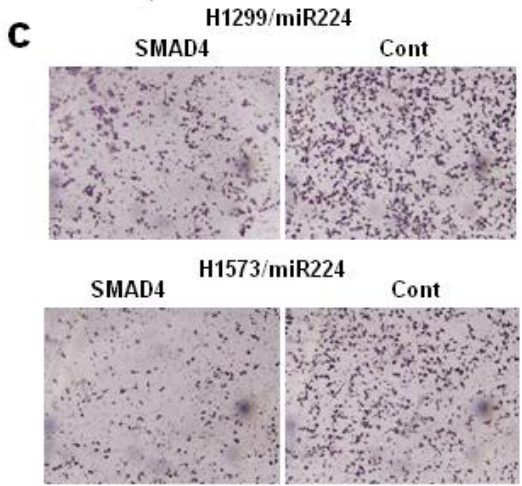
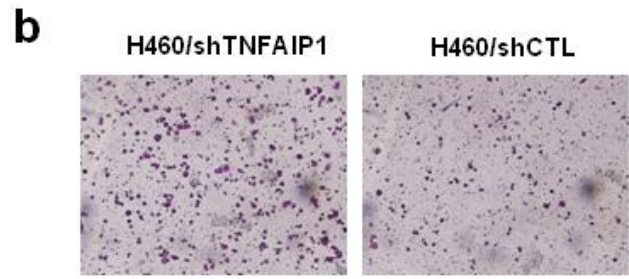
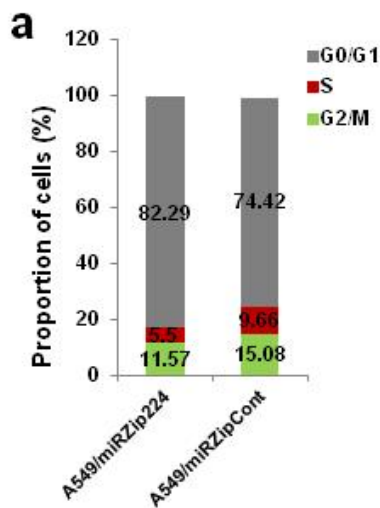


Fig. S7. Representative pictures of cell migration assay and animal studies using miR-224 knockdown lung cancer cells. (a) Proportion of cells in each cell-cycle phase in miR-224 knockdown A549 cells were determined by flow cytometry analysis. (b) Representative pictures of cell migration assay for TNFAIP1 knockdown H460 cells ($\times 100$ magnification). (c) Representative pictures of cell migration assay for miR-224 overexpressing cells transfected with SMAD4 or empty vector. ($\times 100$ magnification). (d) Representative pictures of cell migration assay for H460/miR224 and H460/miRCont cells treated with or without TGF- β . ($\times 100$ magnification). (e) Tumor growth in nude mice s.c. injected into flanks with A549/miRZip224 or A549/miRZipCont after day 45. Data are presented as mean \pm SD (n=6 per group). (f) Representative pictures of lung metastasis after s.c. injection of H460/miR-224 or H460/miRCont cells. Red arrow indicates tumor mass in lung in H460/miR-224 injected mice.

SI Tables

miR ID	Fold change (N1+ vs N0)	P value
hsa-miR-371b-5p	0.01	1.37E-05
hsa-miR-371a-3p	0.01	1.37E-05
hsa-miR-373-3p	0.04	2.35E-03
hsa-miR-34b-5p	0.06	4.99E-03
hsa-miR-3975	47.65	1.82E-05
hsa-miR-1269a	28.69	2.15E-04
hsa-miR-31-3p	27.17	1.65E-04
hsa-miR-224-5p	21.07	3.46E-04
hsa-miR-147b	19.55	4.58E-04
hsa-miR-31-5p	17.07	7.38E-04
hsa-miR-3977	13.49	1.72E-03
hsa-miR-1269b	12.60	2.07E-03
mir-novel-59	12.41	2.79E-03
hsa-miR-363-3p	11.84	2.72E-03
hsa-miR-452-3p	8.70	7.86E-03
hsa-miR-224-3p	8.44	8.31E-03

Table S1. Identification of differentially expressed miRNAs between N0 and N1+ groups of patients.

Primers for 3' UTR mutagenesis	
TNFAIP1_3UTR_mut1_s	catgacggcgggacaggggaaatgtagtcttctaataggcattttatgtagtc
TNFAIP1_3UTR_mut1_as	gactaacataaaatgcctaattagaagactacatttcccctgtcccgcgcatg
TNFAIP1_3UTR_mut2_s	acatgatcttagcttctttaatcagactttgtagtctaaaagtttgggggttttctttgaaagttcc
TNFAIP1_3UTR_mut2_as	ggaaactttcaaagaaaacccccaaacttttagactacaaagtctgattaaagaagctaagatcatgt
SMAD4_3UTR_mut1_s	gtttggatattttgtacttgatttgatgtagtctttttgggtataatgtttaaatcatgtatg
SMAD4_3UTR_mut1_ns	catacatgatttaaacattataccaaaaagactacatcaaatcaagtacaaaaatccaaac
SMAD4_3UTR_mut2_s	ctttgccatcaatgatcatatcaattggcagtagtcttgtatagagaatttaagtagaaaagttgcag
SMAD4_3UTR_mut2_as	ctgcaacttttctacttaaatctctatacaagactactgccaattgatatgatcattgatggcaaag
primers for SMAD4 cloning	
SMAD4_nhe_F	CTAGCTAGCACCATGGACAATATGTCTATTACG
SMAD4_not_R	ATAGTTTAGCGCCGCTCAGTCTAAAGGTT

Table S2. Primers for 3' UTR mutagenesis and SMAD4 cloning

Methylation Probes (Chr:position)	Log2 (has-miR-224)	
	Pearson correlation coefficients	P-value
cg27049053 (chr X: 151142517)	-0.36	1.29E-08
cg12204574 (chr X: 151142773)	-0.19	2.68E-03
cg04929599 (chr X: 151142863)	-0.15	2.09E-02
cg08783090 (chr X: 151143126)	-0.24	1.31E-04
cg18748981 (chr X: 151143213)	-0.06	3.56E-01
cg25528646 (chr X: 151143303)	-0.13	4.43E-02
cg07053880 (chr X: 151143516)	-0.17	8.95E-03
cg01480550 (chr X: 151143527)	-0.17	8.52E-03
cg10276549 (chr X: 151143686)	-0.23	3.34E-04

Table S3. miR-224 expression from TCGA dataset and methylation probes in the promoter region of miR-224 from TCGA Illumina Infinium Human DNA Methylation 450k beadchip in NSCLC.