Supplementary figure 1: Determine the gene mutations and expression which could contribute to the erlotinib resistance

(A)

EGFR exon 19

| A | A | A | C | A | T | C |

AAGGAATTAAGAGAAGCAACATC **wt**
A[AGGAATTAAGAGAAGC]AACATC **delE746-A750**

EGFR exon 20

| T | C | A | C | G | C | A |

T C A C G C A **wt**
T C A T G C A **T790M**

KRAS exon 2

| C | T | G | G | T | G | G |

C T G G T G G **wt**
C T G A T G G **G12D**

KRAS exon 3

| A | A | G | A | G | G | A |

A A G A G G A **wt**
A A C A G G A **A61Q**

(B)

H1993   H1650   H1650 M3

P-MET

c-MET

Tubulin
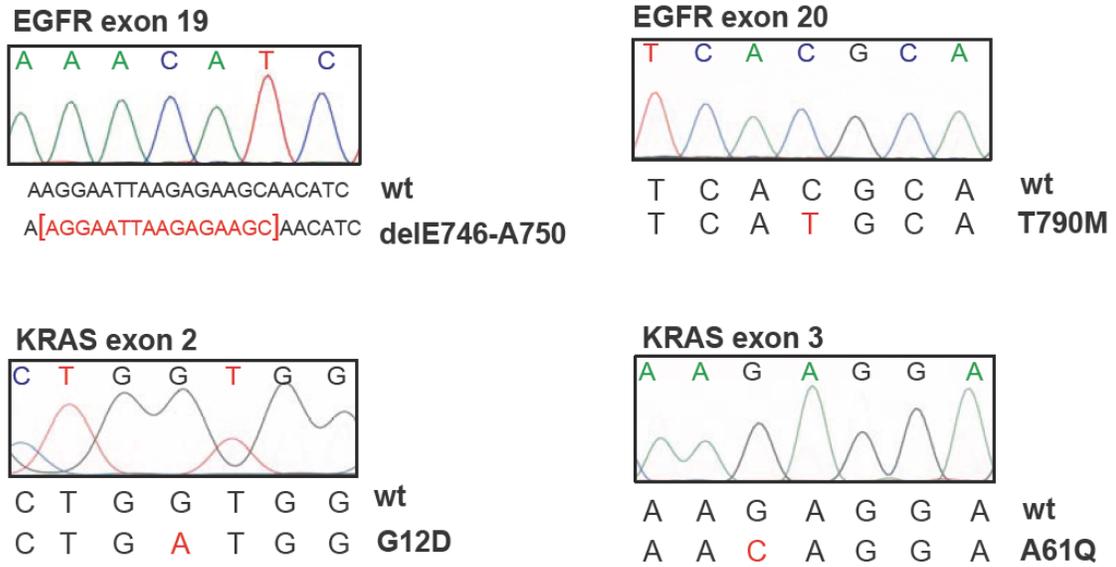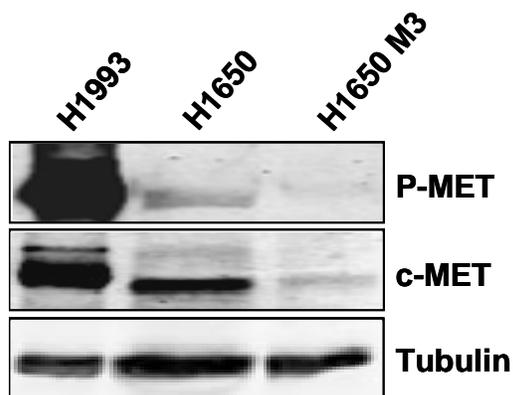
(A) Mutational analysis of EGFR exon 19, EGFR exon 20 and KRAS exon2 and 3 in the H1650 M3 cells. **(B)** Immunoblot analysis of cell extracts from H1650 (parental-sensitive), H1650 M3 (resistant) and H1993 cells. The latter harbors an amplification of c-MET and as such is used as a control for c-MET over-expression and c-MET activation.

Supplementary Figure 2: Summary of gene expression profile data from H1650 and H1650-M3 cells. (A) Genes that showed at least 2.5 fold in expression in the H1650-M3 compared to the H1650 were clustered in the above functional categories. A total of 215 genes were analyzed. The array-based gene expression profiles were validated by **(B)** RT-PCR analysis and **(C)** real-time RT-PCR analysis of selected genes. The relative expression of genes by real time PCR was calculated relative to Actin (n=4).

(A)   Genes up-regulated during the epithelial-mesenchymal transition

Gene symbol	Description	Probe number	Fold difference
COL3A1	collagen, type III, alpha 1 (Ehlers-Danlos syndrome type IV, autosomal dominant)	215076_s_at	7.806079
		211161_s_at	7.329219
		201852_x_at	6.245043
CTGF	connective tissue growth factor	209101_at	6.650338
COL5A2	collagen, type V, alpha 2	221730_at	5.895145
		212488_at	
TGFB2	transforming growth factor, beta 2	209909_s_at	4.289272
TGFB1I1	transforming growth factor beta 1 induced transcript 1	209651_at	4.170635
MMP2	matrix metallopeptidase 2 (gelatinase A, 72kDa gelatinase, 72kDa type IV collagenase)	201069_at	3.720662
COL12A1	collagen, type XII, alpha 1	225664_at	3.585455
		231766_s_at	2.697956
LHX2	LIM homeobox 2	206140_at	3.496055
		211219_s_at	2.803768
DAB2	disabled homolog 2, mitogen-responsive phosphoprotein (Drosophila)	201278_at	3.48811
		210757_x_at	2.910185
		201280_s_at	3.464301
SPARC	secreted protein, acidic, cysteine-rich (osteonectin)	200665_s_at	3.060882
DKK3	dickkopf homolog 3 (Xenopus laevis)	202196_s_at	4.236412
		214247_s_at	5.105802
		230508_at	2.947388
TWIST1	twist homolog 1 (acrocephalosyndactyly 3; Saethre-Chotzen syndrome) (Drosophila)	213943_at	2.875102

Genes involved in tissue remodelling and repair

Gene symbol	Description	Probe number	Fold difference
COL3A1	collagen, type III, alpha 1 (Ehlers-Danlos syndrome type IV, autosomal dominant)	215076_s_at	7.806079
		211161_s_at	7.329219
		201852_x_at	6.245043
CTGF	connective tissue growth factor	209101_at	6.650338
COL5A2	collagen, type V, alpha 2	221730_at	5.895145
		221729_at	4.785885
		212489_at	3.863292
NID1	nidogen 1	202007_at	5.359996
CYR61	cysteine-rich, angiogenic inducer, 61	210764_s_at	4.440958
FBLN1	fibulin 1	202995_s_at	4.392872
RBMS3	RNA binding motif, single stranded interacting protein	238447_at	3.992004
CCL2	chemokine (C-C motif) ligand 2	216598_s_at	3.936064
ETS1	v-ets erythroblastosis virus E26 oncogene homolog 1 (avian)	224833_at	3.795118
FOXF1	forkhead box F1	205935_at	3.779365
MMP2	matrix metallopeptidase 2 (gelatinase A, 72kDa gelatinase, 72kDa type IV collagenase)	201069_at	3.720662
F2R	coagulation factor II (thrombin) receptor	203989_x_at	3.500018
F2RL2	coagulation factor II (thrombin) receptor-like 2	230147_at	4.283747
LEPREL1	leprecan-like 1	218717_s_at	3.336609
SPARC	secreted protein, acidic, cysteine-rich (osteonectin)	200665_s_at	3.060882
PTX3	pentraxin-related gene, rapidly induced by IL-1 beta	206157_at	2.990426
TIMP2	TIMP metallopeptidase inhibitor 2	224560_at	2.792585
ENTPD4 /// LOXL2	lysyl oxidase-like 2 /// ectonucleoside triphosphate diphosphohydrolase 4	202998_s_at	2.773696

Genes involved in migration and motility

Gene symbol	Description	Probe number	Fold difference
THBS1	thrombospondin 1	201110_s_at	5.807242
		201109_s_at	4.923286
MYL9	myosin, light chain 9, regulatory	201058_s_at	5.561672
NEXN	nexilin (F actin binding protein)	226103_at	5.23588
		1552309_a_at	3.497584
PTGE	prostaglandin E receptor 4 (subtype EP4)	204897_at	5.175662
CYR61	cysteine-rich, angiogenic inducer, 61	210764_s_at	4.440958
		201289_at	3.71511
MMP2	matrix metallopeptidase 2 (gelatinase A, 72kDa gelatinase, 72kDa type IV collagenase)	201069_at	3.720662
DOCK10	dedicator of cytokinesis 10	219279_at	3.142447
FRMD6	FERM domain containing 6	225481_at	3.135242
SPARC	secreted protein, acidic, cysteine-rich (osteonectin)	200665_s_at	3.060882
FERMT2	fermitin family homolog 2 (Drosophila)	209209_s_at	2.966072
		214212_x_at	2.643792
TSPAN5	tetraspanin 5	225387_at	2.921342
DOCK1	dedicator of cytokinesis 1	241708_at	2.833222
ITGA4	integrin, alpha 4 (antigen CD49D, alpha 4 subunit of VLA-4 receptor)	205885_s_at	2.798731
		205884_at	2.798518
ACTN1	actinin, alpha 1	208637_x_at	2.783385
LOXL2	lysyl oxidase-like 2 /// ectonucleoside triphosphate diphosphohydrolase 4	202998_s_at	2.773696

Genes relevant in metastasis

Gene symbol	Description	Probe number	Fold difference
SUSD5	sushi domain containing 5	214954_at	5.84075
THBS1	thrombospondin 1	201110_s_at	5.807242
		201108_s_at	2.707674
		201109_s_at	4.923286
MCAM	melanoma cell adhesion molecule	211340_s_at	4.486635
		209087_x_at	4.224188
		210869_s_at	3.864155
FBLN1	fibulin 1	202995_s_at	4.392872
CCL2	chemokine (C-C motif) ligand 2	216598_s_at	3.936064
S100A2	S100 calcium binding protein A2	204268_at	3.785131
BST2	bone marrow stromal cell antigen 2	201641_at	3.333186
RUNX2	runt-related transcription factor 2	232231_at	3.147917
TWIST1	twist homolog 1 (acrocephalosyndactyly 3; Saethre-Chotzen syndrome) (Drosophila)	213943_at	2.875102
JAM2	junctional adhesion molecule 2	219213_at	2.813103
ITGA4	integrin, alpha 4 (antigen CD49D, alpha 4 subunit of VLA-4 receptor)	205885_s_at	2.798731

TGF-β regulated genes

Gene symbol	Description	Probe number	Fold difference
CTGF	connective tissue growth factor	209101_at	6.650338
COL3A1	collagen, type III, alpha 1 (Ehlers-Danlos syndrome type IV, autosomal dominant)	215076_s_at	7.806079
		211161_s_at	7.329219
		201852_x_at	6.245043
SRPX	sushi-repeat-containing protein, X-linked	204955_at	4.692204
TGFB2	transforming growth factor, beta 2	209909_s_at	4.289272
CCL2	chemokine (C-C motif) ligand 2	216598_s_at	3.936064
ETS1	v-ets erythroblastosis virus E26 oncogene homolog 1 (avian)	224833_at	3.795118
		1555355_a_at	3.276713
MMP2	matrix metallopeptidase 2 (gelatinase A, 72kDa gelatinase, 72kDa type IV collagenase)	201069_at	3.720662
RUNX2	runt-related transcription factor 2	232231_at	3.147917
SPARC	secreted protein, acidic, cysteine-rich (osteonectin)	200665_s_at	3.060882
		212667_at	3.530221
TWIST1	twist homolog 1 (acrocephalosyndactyly 3; Saethre-Chotzen syndrome) (Drosophila)	213943_at	2.875102
LOXL2	lysyl oxidase-like 2 /// ectonucleoside triphosphate diphosphohydrolase 4	202998_s_at	2.773696
CLIC4	chloride intracellular channel 4	201559_s_at	2.740826
		221881_s_at	3.205253
KLF11	Kruppel-like factor 11	218486_at	2.673633

Genes regulated by IL-6/ STAT signalling pathway

Gene symbol	Description	Probe number	Fold difference
BST2	bone marrow stromal cell antigen 2	201641_at	3.333186
NFATC2	Nuclear factor of activated T-cells, cytoplasmic, calcineurin-dependent 2	226991_at	3.065041
PTX3	pentraxin-related gene, rapidly induced by IL-1 beta	206157_at	2.990426
LHX2	LIM homeobox 2	211219_s_at	2.803768
		206140_at	3.496055
IFI27	interferon, alpha-inducible protein 27	202411_at	2.54035
IL-6	interleukin 6 (interferon, beta 2)	205207_at	2.530991
ID2	inhibitor of DNA binding 2, dominant negative helix-loop-helix protein	201565_s_at	2.523807

(B)
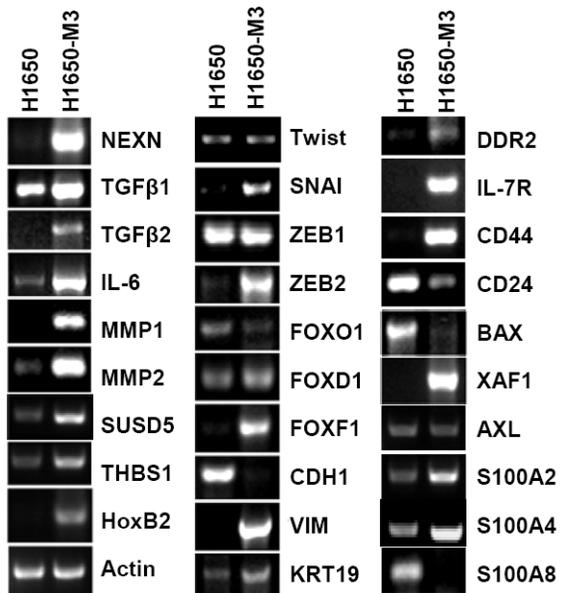
(C)
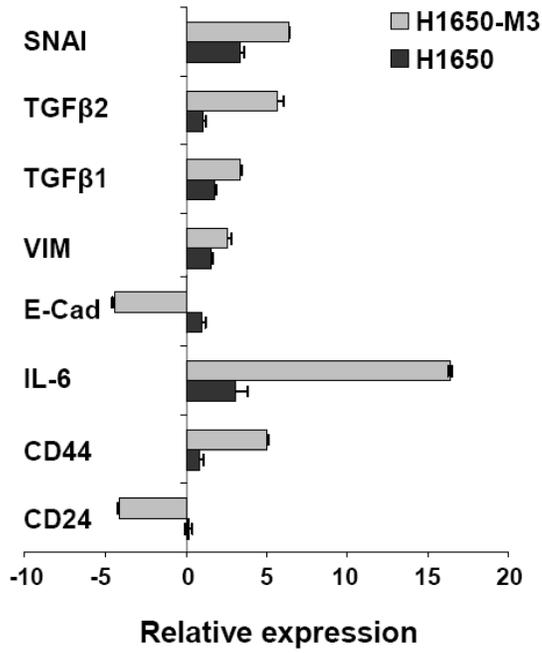
Supplementary Figure 3: TGF-β signaling is upregulated in an EGFR independent manner.

(A)
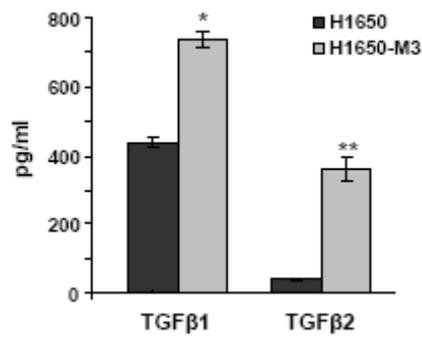
(B)
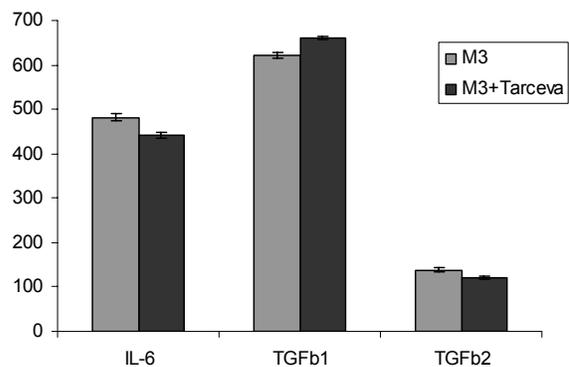
(C)
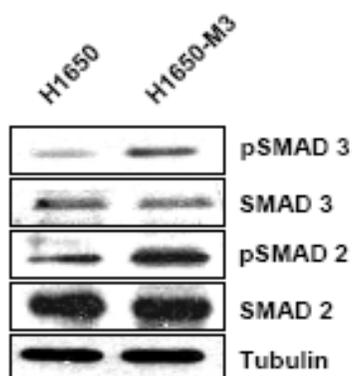
(D)
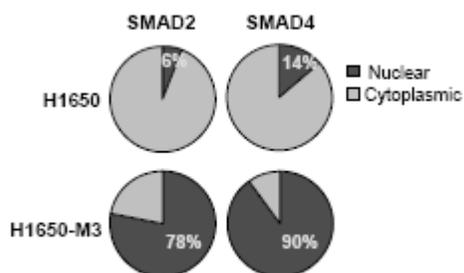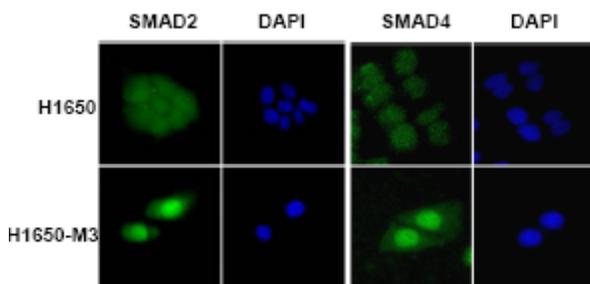
(A) The increased TGF-β1 and TGF-β2 mRNA levels were correlated with augmented TGF-β1, and TGF-β2 secretion. Levels of TGF-β1, and TGF-β2 were measured by ELISA. Each bar represents the mean ± SD (n=4) with a Student's t test, p<0.0001. **(B)** TGF-β1, TGF-β2 and IL-6 levels in the media of the H1650 M3 cells do not change in the presence of erlotinib (10uM). TGF-β levels were assessed by ELISA assay. Each bar represents the average of four data points. **(C)** Augmented bioavailability of TGF-β1, and TGF-β2 resulted in higher activation of TGF-β mediated signalling. Cell extracts were immunoblotted to detect levels of expression and phosphorylation of indicated proteins **(D)** Immunofluorescence staining was employed to determine the nuclear translocation of SMAD2 and SMAD4 (Green). Cells were stained with the indicated antibodies. DAPI staining (Blue) was used for nuclear localization and to ensure an equal image exposure. The chart represents the percentage of cells (n=50) with nuclear versus cytoplasmic distribution of SMAD2 and SMAD4 in erlotinib-sensitive (H1650) and in representative erlotinib-resistant cells (H1650-M3).

Supplementary Figure 4: TGF-β1/2 are required and sufficient to induce EMT, increase motility and invasion

(A) **(B)**
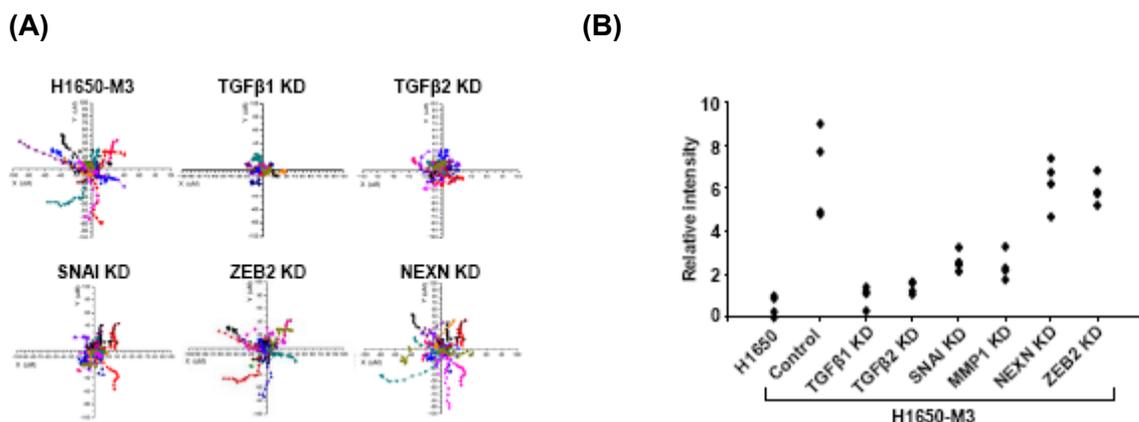
(C)
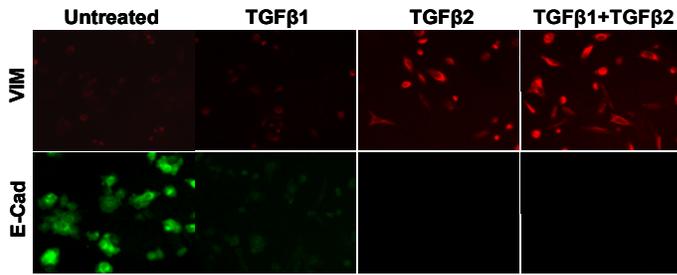
(D)
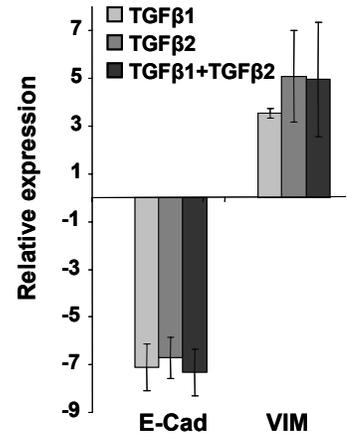
(E)
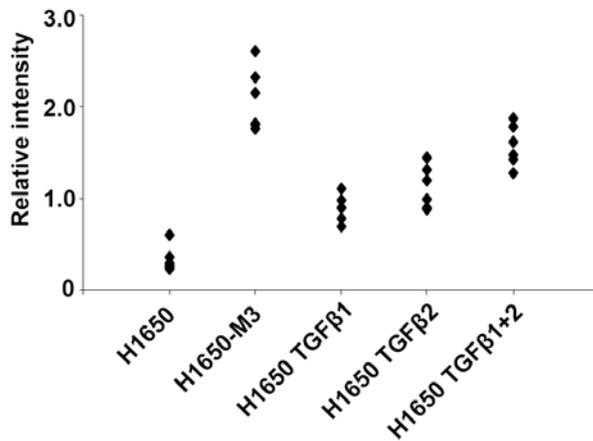
(A) Decreased expression of TGF-β1, TGF-β2, and SNAI resulted in decreases in single cell motility and **(B)** in inhibition of Matrigel invasive capabilities. Of note, no substantial difference in cell motility or invasion was observed when cells were transfected with shRNAs targeting nexilin and ZEB2.The single cell motility was calculated by tracking the movement of single cells (n=12) for a period of 3.3 hours. Each trace represents the movement of a single cell, whereas individual dots designate a frame of 10 minutes**.** The

3

invasion potential of cells was determined in a standard Matrigel invasion assay using a coating of 20 µl Matrigel. **(C)** Erlotinib-sensitive (H1650 cells) cells treated for 8 days with TGF-β1, TGF-β2, or TGF-β1 and TGF-β2 in combination displayed morphological features of mesenchymal cells. Immunostaining with Vimentin and E-cadherin antibodies and **(D)** real-time RT-PCR indicated an increased expression of Vimentin and a diminished expression of E-cadherin compared to untreated cells. The chart represent relative expression level normalized to actin. **(E)** The cells also displayed an increased motility (data not shown) and invasive capabilities in a modified Boyden chamber assay. Filters coated with different voulmes of Matrigel (i.e., 20 µl and 40 µl) were used. Each dot represents an individual replica (n=5).

Supplementary Figure 5: IL-6 is sufficient and required to modify sensitivity of cells to erlotinib treatment.

(A)
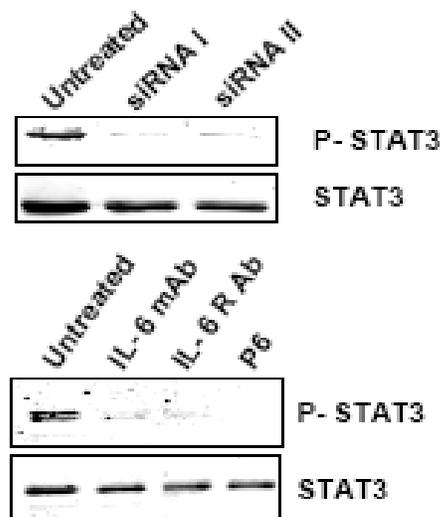
(B)
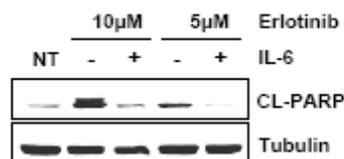
(C)
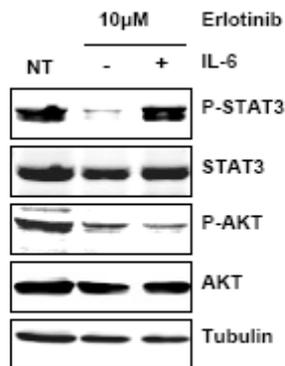
(A) Decreasing IL-6 expression by siRNA interference decreased the phosphorylation of STAT3. H1650-M3 cells were transfected with siRNA targeting IL-6 (siRNA I or II) using Lipofectamine 2000 at a final concentration of 30 nM siRNA. After 6 days, cell extracts were analyzed by immunoblotting for phospho-STAT3, STAT3 and Tubulin. Similarly inhibition of the IL-6 axis decreased levels of phosphorylation of STAT3. H1650-M3 cells were treated with 1 µg/mL IL-6 neutralizing antibody, 10 µg/mL IL-6 receptor antibody or 2 µM pan-JAK inhibitor for 2 days. Cell extracts were analyzed by immunoblotting for phospho-STAT3 and total STAT3. **(B)** The decreased sensitivity observed in cells treated with IL-6 is due to a diminished apoptotic response. Cells were grown in the presence or absence of IL-6 for 24 hours and then treated with erlotinib at the indicated concentrations. Control cells were not treated with erlotinib. Cell extracts obtained after 72 hours were analyzed by immunoblotting for cleaved PARP. **(C)** IL-6 induces STAT-3 phosphorylation even in the presence of erlotinib without affecting the inhibition of other signalling EGFR downstream (i.e. AKT). Cells were grown in the presence or absence of

IL-6 and then treated for 4 hours with erlotinib. Cell extracts were analyzed by immunoblotting with the indicated antibodies.

Supplementary Figure 6: Erlotinib-resistant, mesenchymal cells are already present in erlotinib-naive NSCLC tumors.

(A)
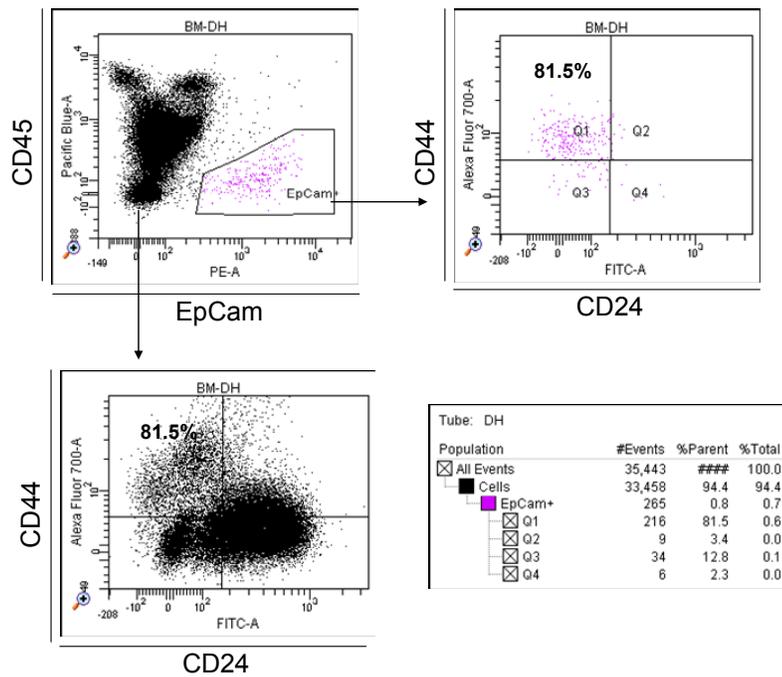
(B)

DTS	Demography (sex/ age)	TNM	Stage	% EpCAM⁺ CD44⁺ CD24⁻ NORMAL	% EpCAM⁺ CD44⁺ CD24⁻ TUMOR	Ratio
13	M/73	T2N1M0	IIB	0.1	2.7	27.0
19	F/80	T4N2M0	IIIB	0.1	0.7	7.0
25	F/74	T4N0M0	IIIB	0.8	2.3	2.9
51	F/61	T1NOM0	IA	0.7	2.4	3.4
54	F/59	T1N0M0	IIA	0.2	2.5	12.5
62	F/65	T1N0M0	IA	0.01	0.1	10.0
69	F/62	T2N0M0	IB	0.02	0.1	5.0
78	F/65	T2N0M0	IB	0.6	3.9	6.5
80	M/77	T4N0M0	IIIB	0.5	4.6	9.2
81	F/80	T1N0M0	IA	0.6	1.4	2.3
Total=20				Median = 0.35	Median = 2.6	Median=11.3

(C)

DTS	Demography (sex/ age)	TNM	Stage	% EpCAM⁺/TOTAL # cells in BM	% EpCAM⁺CD44⁺CD24⁻ EpCAM⁺ in BM
13	M/73	T2N1M0	IIB	0.3	39.6
AR	F/55	T4N0M0	IIIB	0.7	87

(D)

Tumor1

Tumor2
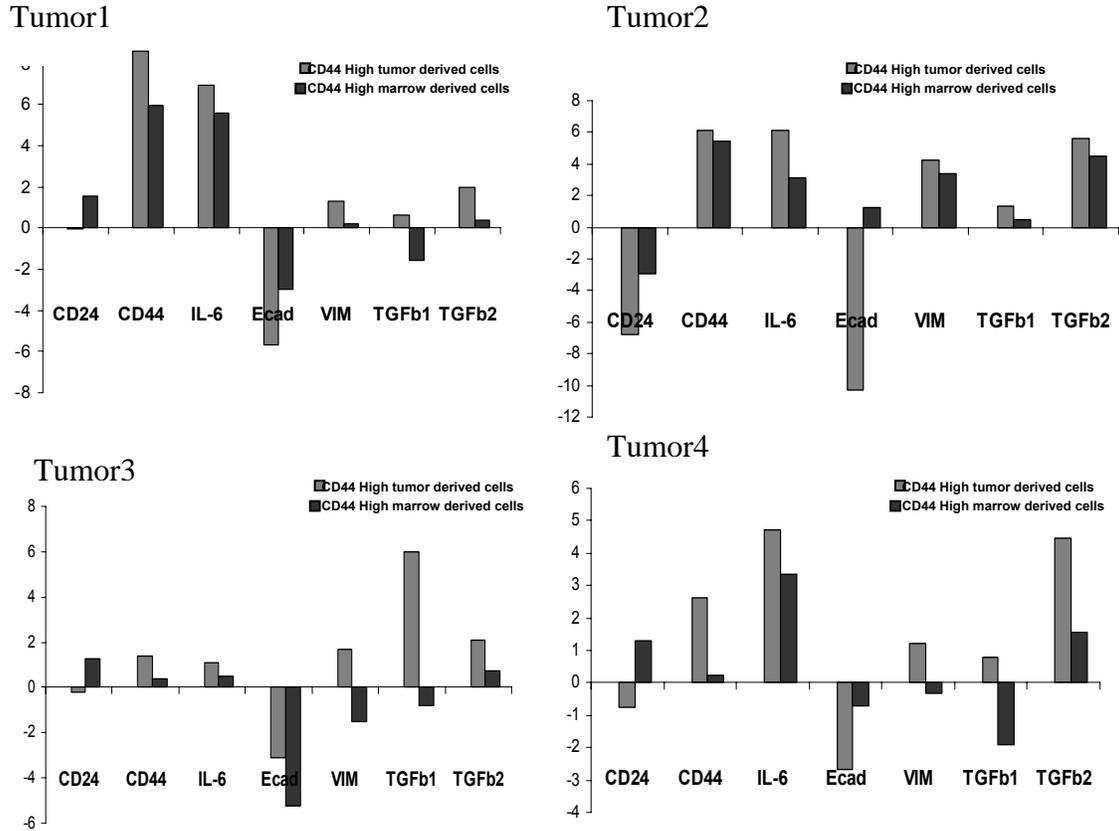
Tumor3

Tumor4

(A) Flow cytometry analysis for CD44 and CD24 in NSCLC tumor cells. Dissociated tumor cells were analyzed for CD45, CD31, EpCAM, CD44, and CD24. The gates shown were utilized to sort CD44^{high}CD24^{low} cells. **(B)** The table summarizes the data collected from 10 samples from lung adenocarcinoma/BAC patients. The percentage of EpCAM^{+}CD44^{high}CD24^{low} cells (i.e., CD45^{-}/CD31^{-}/EpCAM^{+}/CD24^{-}/C44^{+}) was calculated relative to the total number of CD45^{-}CD31^{-}EpCAM^{+} cells in dissociated tumor and normal tissues. **(C)** The table summarizes the data collected from the samples of bone marrow of two adenocarcinoma patients. The percentage of EpCAM^{+}/TOTAL # cells in BM represents the percentage of EpCAM positive cells in the total BM cell preparation, whereas the % EpCAM^{+}CD44^{+}CD24^{-}EpCAM^{+} in BM represents the percentage of EpCAM^{+}CD44^{+}CD24^{-} cells present in the EpCAM^{+} fraction in the bone marrow of

NSCLC patients. **(D)** Real-time PCR analysis comparing gene expression profiles of selected genes in CD44^{high}CD24^{low}CD45$^-$CD31$^-$EpCAM$^+$ cells (i.e., tumor-derived cells) to CD44^{high}CD24^{low}CD45$^+$CD31$^+$EpCAM$^-$ cells indicated that the former cells had a signature similar to the one observed in the erlotinib-derived H1650-M3 cells. The analysis was conducted on cells from four independent tumors.

Supplementary Figure 7: Inflammation induces IL-6 expression and decreased the tumor sensitivity to erlotinib.

(A)
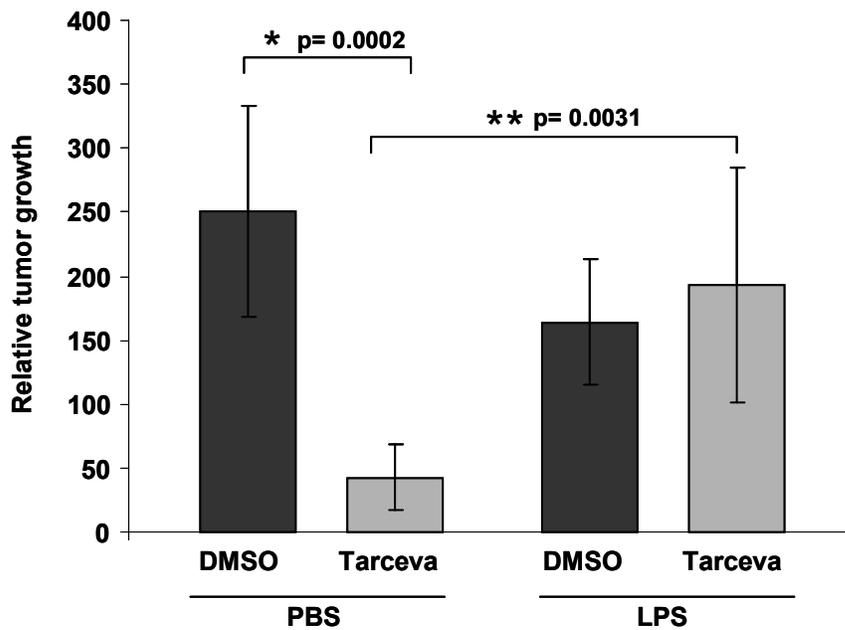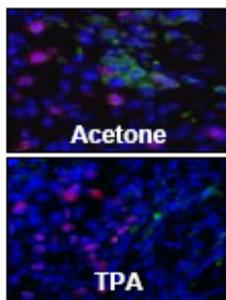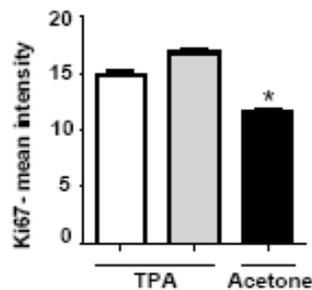
(B)   **(C)**   **(D)**
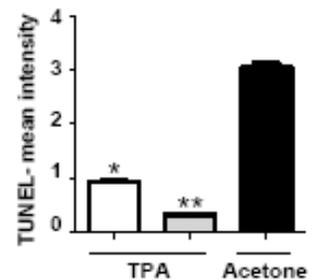
(A) LPS treatment decreases the tumors response to erlotinib treatment. The chart represents growth percentage of tumors. Each column represents the mean volume of 8 different tumors. **(B)** The effects of TPA are due to a decreased of apoptosis rather than of an increased proliferation. The panel represents a representative immunostaining of a paraffin section for Ki67 in purple, TUNEL assay in green and counter-staining with DAPI in blue. (**C**) The chart depicts Ki67 mean intensities of three representative tumors. Protein expression was measured by defining ROIs using automated cell acquisition and quantification software for immunofluorescence (Tissuequest^{TM}). (**D**) The TUNEL mean intensities of three representative tumors are depicted in the chart. The differences are significant as indicated (Student's t test).

Supplementary Figure 8: IL-6 stimulation is not sufficient to induce EMT.
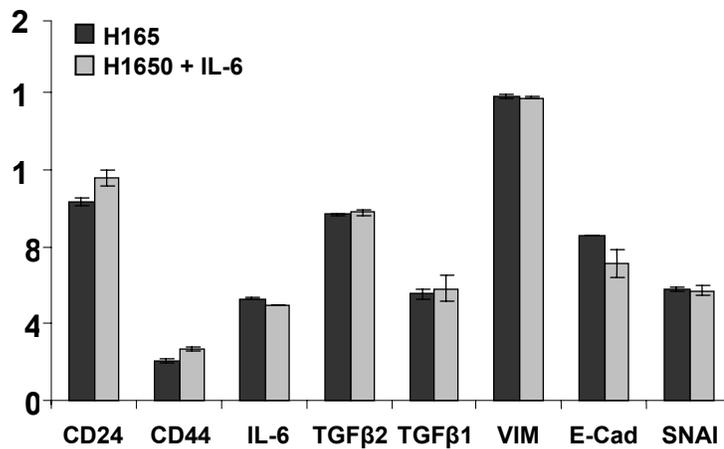
Stimulation of parental erlotinib-sensitive cells (H1650) with IL-6 did not alter expression of CD24, CD44, IL-6, TGF-β2, TGF-β1, Vimentin (VIM), E-cadherin (E-Cad), or Snail (SNAI) as measured by real-time RT-PCR. Relative RNA expression levels were normalized to actin mRNA.

Material and Methods

Cell culture

H1650 (H1650), HCC827, H1975 (NCI-H1975), H1299 (NCI-H1299), HCC4006 were obtained from the American Type Culture Collection (ATCC) repository. The PC9 cell line was a gift from Dr. Jeff Engelman. All the cell lines were cultured in RPMI supplemented with 5% FBS, glutamine, penicillin, and streptomycin. Erlotinib-resistant cell lines were selected from H1650 cells by growing for approximately one month in media containing 20 µM erlotinib and then maintaining cells for 2 additional months in 10 µM erlotinib-containing media. The cells were then switched to regular growth media. The lentiviral packaging cell line HEK293T was cultured in DMEM containing 10% FBS, penicillin, streptomycin, and sodium pyruvate.

Antibodies and reagents

The following antibodies were used in this study: mouse anti-E-cadherin antibody (BD Transduction Laboratories), monoclonal anti-Vimentin antibody (RV202, Santa Cruz Biotechnology), mouse anti-β-Tubulin antibody (2-28-33, Santa Cruz Biotechnology), mouse anti-STAT3 antibody (124H6, Cell Signaling Technology), and rabbit anti-phospho-STAT3 antibody (D3A7, Cell Signaling Technology). The Smad molecules were detected with the Phospho-Smad antibody Sampler Kit from Cell Signaling Technology. Erlotinib hydrochloride was purchased from LGM Pharmaceuticals, Inc. Recombinant interlukin-6 (rhIL6), anti-IL-6 monoclonal antibody (MAB206) and anti-IL-6R antibody (AB-227-NA) were obtained from R&D Systems. Recombinant TGFβ1/2

were purchased from Sigma-Aldrich. Human recombinant EGF was purchased from Millipore. The JAK 1/2 inhibitor tetracyclic pyridone 2-*tert*-butyl-9-fluoro-3,6-dihydro-7*H*-benz[h]-imidaz[4,5-*f*]isoquinoline-7-one, pyridone 6 (P6) was purchased from Calbiochem.

Genomic DNA extraction and mutational analysis

Genomic DNA extraction was performed as described in "Molecular Cloning" by Sambrook and Russell. KRAS exons and EGFR exons were amplified using Pfu-Ultra (Stratagene) using the following primers:

KRAS exon2 (fwd: CTTAAGCGTCGATGGAG; rev: ATCCTCATCTGCTTGGGATG),

KRAS exon3 (fwd: AAATAGTGCTGCTGCGAACA; rev: CTAGGTTTCAATCCCAGCA),

EGFR exon19 (fwd: GCAATATCAGCCTTAGGTGCGGCTC; rev: CATAGAAAGTGAA-CATTTAGGATGTG), EGFR exon 20 (fwd: CCATGAGTACGTATTTTGAAACTC; rev: CATATCCCCATGGCAAACTCTTGC).

PCR products were purified using PrepEase® Gel Extraction Kit (USB) and were sequenced using the following primers:

KRAS exon 2 (fwd: AAGGCCTGCTGAAAATGACTG; rev: CTGTATCAAAGAATGG-TCCTGCAC), KRAS exon 3 (fwd: CCAGACTGTGTTTCTCCCTTCTCAG; rev: AAC-CCACCTATAATGGTGAA), EGFR exon 19 (fwd: CCTTAGGTGCGGCTCCACAGC; rev: CATTTAGGATGTGGAGATGAGC), EGFR exon 20 (fwd: GAAACTCAAGATCG-CATTCATGC; rev: GCAAACTCTTGCTATCCCAGGAG).

Immunoblots

Cells were washed with PBS before collection and lysed directly in RIPA buffer containing 0.2% SDS. Proteins were separated by SDS-PAGE, transferred to nitrocellulose membranes (Bio-Rad), and blotted with antibodies as indicated.

Immunostaining

Cells were cultured in 24-well chamber slides. After washing once with phosphate-buffered saline (PBS), cells were fixed in 4% paraformaldehyde and permeabilized in 0.1% Triton X-100 in PBS for 10 min. Fixed cells were washed three times in PBS and blocked with 3% bovine serum albumin in PBS for 30 min. After washing three times with PBS, cells were incubated with the primary antibody for 1 h at room temperature. Immune complexes were then stained with donkey anti-mouse immunoglobulin G conjugated with fluorescein (FITC) or rhodamine antibodies (Invitrogen). DAPI was used for nuclear staining. Stained cells were mounted with Vectashield Mounting medium (Vector Laboratories) and analyzed under an inverted microscope.

Lentiviral-based and siRNA-based gene knockdown

All the shRNA constructs (lentiviral pGIPz plasmids) were purchased from Openbiosystem. The recombinant construct as well as the packaging plasmids (i.e., BH10, pREV, and pVSVG) were transiently transfected into HEK293T cells using lipofectamine2000 (Invitrogen). Viral supernatants were collected and filtered 48 hours after initial transfection. Infection was performed at low MOI (MOI=0.6) in the presence

of Polybrene. After 5 days, infected cells were sorted based on GFP expression and m
To target the IL-6 gene, commercially prepared Stealth RNAi siRNAs I) HSS105338 and
II) HSS105339 were obtained from Invitrogen. siRNAs were transfected into cells using
Lipofectamine 2000 (Invitrogen) as per the manufacturer's instructions. Two days after
transfection, mRNA levels were determined by RT-PCR and sensitivity to erlotinib was
determined as previously described. Stable knockdown was obtained even at 10 days
following transfection. aintained in 1 ng/ml puromycin for 1 week. mRNA levels were
assessed by RT-PCR.

Migration and invasion assays

Single cell tracks were traced by time lapse photography. Twenty randomly chosen cells
were traced by following the centroid over 500 mins at 10 min intervals. Each cell was
then plotted with time 0 at the origin of the grid.

For wound-closure experiments, cells were plated in 10-cm plates and cultured to
confluence. Cells were scraped with a p20 tip and transferred to pre-warmed fresh
media. The healing of the gap was observed at indicated time points.

The invasion assays were performed in 24-well 6.5-mm diameter inserts (Corning, 8.0-
mm pore size) coated with a indicator layer of growth factor reduced Matrigel (BD
Transduction Laboratories). The cells were plated in the upper well in 0.2% serum and
incubated with 5% FBS and 100 ng/ml fibronectin in the lower chambers. After 24 hours,

cells in the upper chamber were removed with a cotton swab. Cells that had migrated into the lower chamber were fixed in 4% PFA and stained with 0.5% Crystal Violet. Filters were photographed and the total number of cells was quantified using the Odyssey Imaging System. Every experiment was repeated in quadruplicate.

ELISA

The ELISAs were performed with commercially available kits: The kit used to detect the secreted IL-6 was the high-sensitivity kit from Cell Sciences, the TGF-β levels were determine using the Human TGF-β 1/2 Quantikine ELISA Kit (R&D Systems). The tested samples were collected from the cultured or treated cell media and subjected to analysis immediately after a 1:10 dilution.

RNA extraction and quantitative RT-PCR

Total RNA was collected from cultured cells and purified using Trizol reagent following the manufacturer's instructions (Invitrogen). Total RNA (500 ng) was subjected to a reverse transcriptase reaction using the Improm-II Reverse Transcriptase kit (Promega). cDNA corresponding to approximately 1% of the RNA was used in three replicates for quantitative PCR with the SYBR GREEN (Applied Biosystems) labeling or for standard PCR. Indicated Taqman gene expression assays (Applied Biosystems) and the Taqman universal PCR master mix (Applied Biosystems) were used to quantify gene expression. Quantitative expression data were acquired and analyzed using an ABI (Applied Biosystems) Sequence Detection System. A list of oligonucleotides used and PCR conditions are provided in Supplementary Figure 8.

Gene expressing profiling

RNA extraction was performed as described. RNA quality was assessed on an Agilent 2100 Bioanalyzer, RNA 6000 Pico Series II Chips. Samples with RIN scores of 2.0 or greater were processed. Total RNA was amplified by a modified Eberwine Technique. Amplified RNA was then converted to cDNA using a WT Expression kit (Ambion).   Size distribution of aRNA and cDNA was assessed for 3' bias on all samples using Agilent 2100 Bioanalyzer RNA 6000 Nano Series II Chips. The cDNA was then fragmented and terminally labeled with biotin, using the Affymetrix GeneChip WT Terminal Labeling kit. Samples were then prepared for hybridization, hybridized, washed, and scanned according to the manufacturer's instructions on U133 plus 2.0 Affimetrix GeneChips (Affymetrix). Affymetrix Expression Console QC metrics were used to process the image data.

MTT assay

A total of 3,000 cells were plated in 0.2 ml in 96-well flat bottom plates and then exposed to erlotinib as indicated. At the indicated times, 20 µl of 5 mg/ml MTT solution in PBS was added to each well. After 3-4 h, medium was removed and 200 µl of DMSO was added to each well to dissolve the formazan crystals. The absorbance at 540 nm was determined using a plate reader (SpectraMax 190, Molecular Devices, Inc.). Experiments were conducted in octuplicate.

Soft agar assay

Cells were plated in triplicate in 12-well plates at 5000 cells/well in RPMI-1640, 5% FBS and 0.35% agarose. Cells were cultured for 20 days in presence of the indicated treatment. Colonies were visualized by staining with tetrazolium salt The number of colonies were determined using the J-Image software (htpp://rsbweb.hih.gov/ij/) on Z-stack images (Z=1.5 mm) at 5x magnification. Each point represents the average of 3 fields from 3 individual well.

FACS and tumor tissue analysis

H1650 and H1650-M3 and all other NSCLC cell lines were grown in RPMI media supplemented with 5% FBS, sodium pyruvate, penicillin, and streptomycin. Cells were harvested at 70% confluency, dissociated and washed in PBS with 5% FBS. Following this wash, cells were pelleted and resuspended in solution containing PBS (pH 7.2), 0.5% bovine serum albumin, and 2 mM EDTA. Labeling was performed using eFluor 450 conjugated anti-CD24 (eBioscience) and either PE-Cy7 conjugated or Alexa 700 conjugated anti-CD44 (Biolegend). Counting experiments were performed using an LSR II flow cytometer (Becton Dickenson); 100,000 events were collected for each sample and repeated independently in triplicate. Sorting experiments were done with a BD FACS Aria II cell sorter. Gates were set using H1650-M3 as a reference population.

NSCLC tumors were dissociated in collagenase (200 unit/mL) and hyaluronidase (100 unit/mL) and were suspended in DMEM for two hours at 37°C. Cell suspensions were strained using BD Falcon 100 micron Cell Strainers, pelleted and subjected to hypo-osmotic red blood cell lysis. Cell suspensions were then pelleted, washed with PBS and

resuspended in antibody staining buffer previously described and filtered through 40 micron Cell strainers (BD Falcon). Afterwards, samplers were stained for EpCAm, CD24, CD44, CD45 (Abcam) and CD31 (eBioscience) for 30 minutes on ice. Human tissues were obtained from CT Surgery Department, Weill Cornell Medical College and patient consent was obtained according to approved IRB protocols from the institution.

Mouse model system

The in vivo mouse experiments were performed by inoculating 3×10^6 H1650 subcutaneously into each flank of immunodeficient *NuINu* mice on Swiss CD1 background (Charles River Breeding Laboratories). After the tumors reached a volume of approximately 100 mm^3, mice were treated topically every 2 days with TPA (10 µL of 0.25 mg/mL solution in acetone) or with acetone for 5 days. After this treatment, erlotinib (50 mg/kg) or vehicle alone (DMSO) was delivered daily by oral gavage for 12 days. Tumor volume was monitored with a digital caliper and calculated using the following formula: (length) x (width)^2 /2. The neutralizing antibody against IL-6 (2 mg in PBS) was injected intraperitoneally one day before erlotinib treatment began.

Supplementary Figure 9: Oligonucleotides and PCR conditions.

Oligonucleotides for semi-quantitative RT-PCR:

```
MMP1-F:

5'-ATGCACAGCTTTCCTCCA-3'

MMP1-R:

5'-TGAATGTCAGAGGTGTGA-3'

DDR2-F:
```

5'-TCAGTTACACCAATC-3'

DDR2-R:

5'-GCAAGTTCACTACAG-3'

XAF-F:

5'-GAAGATCTATGGACTACAAGGACGACGATGACAAGGAAGGAGACTTCTCG-3'

XAF-R:

5'-GGAATTCCTAGCTGAAATTTCTCAC-3'

CD24-F:

5'-CGGGATCCATGGGCAGAGCAATGGTG-3'

CD24-R:

5'-GGAATTCTTAAGAGTAGAGATGCAG-3'

CD44-F:

5'-ATGGACAAGTTTTGGTGG-3'

CD44-R:

5'-TAGTTATGGTAATTGGTC-3'

TGFB2-F:

5'-CGGGATCCATGCACTACTGTGTGCTG-3'

TGFB2-R:

5'-GGAATTCTTAGCTGCATTTGCAAGA-3'

TGFB1-F:

5'-TGTGCGGCAGTGGTTGAG-3'

TGFB1-R:

5'-TCAGCTGCACTTGCAGGA-3'

BAX-F:

5'-ATGGACGGGTCCGGGGAG-3'

BAX-R:

5'-GGAATTCTCAGACACGTAAGGA-3'

SNAIL-F:

5'-CGGGATCCATGCCGCGCTCTTTCCTC-3'

SNAIL-R:

5'-GGAATTCTCAGCGGGGACATCCTGA-3'

ZEB1-F:

5'-CAGATGAAGCAGGATGTA-3'

ZEB1-R:

5'-CTGATTTATGTGATGTCA-3'

SIP1-F:

5'-CGGGATCCATGCGCCGAGCGGAACTG-3'

SIP1-R:

5'-GGAATTCTCAAGATGGCTCATCAGC-3'

CDH1-F:

5'-AGCACGTGAAGAACAGCA-3'

CDH1-R:

5'-CCTCTAAGGTGGTCACTT-3'

VIM-F:

5'-CGGGATCCATGTCCACCAGGTCCGTG-3'

VIM-R:

5'-GGAATTCTTATTCAAGGTCATCGTG-3'

MMP2-F:

5'-ATGGAGAGGCAGACATCA-3'

MMP2-R:

5'-TGCCATCCTTCTCAAAGT-3'

TWIST-F:

5'-CGGGATCCATGATGCAGGACGTGTCC-3'

TWIST-R:

5'-GGAATTCCTAGTGGGACGCGGACAT-3'

IR7R-F:

5'-ATGACAATTCTAGGTACA-3'

IR7R-R:

5'-TAACTATAGTGGTTAGGT-3'

KRT19-F:

5'-ATGACTTCCTACAGCTAT-3'

KRT19-R:

5'-TCAGAGGACCTTGGAGGC-3'

FOXF1-F:

5'-ATGTCTTCGGCGCCCGAG-3'

FOXF1-R:

5'-TCACATCACGCAAGGCTT-3'

FOXO1-F:

5'-AGAGGGTGGCAAGAGCGG-3'

FOXO1-R:

5'-TCTGCTGCATCATGGTGC-3'

FOXD1-F:

5'-ATGACCCTGAGCACTGAG-3'

FOXD1-R:

5'-TTAACAATTGGAAATCCT-3'

IL6-F:

5'-CGGGATCCATGGACTACAAGGACGACGATGACAAGAACTCCTTCTCCACA-3'

IL6-R:

5'-GGAATTCCTACATTTGCCGAAGAGC-3'

S100A8-F:

5'-ATGTTGACCGAGCTGGAG-3'

S100A8-R:

5'-CTGCCACGCCCATCTTTA-3'

SUSD5-F:

5'-ACGTCCTAGTTACTCCT-3'

SUSD5-R:

5'-AGTTGACTGGTACTTCGT-3'

21

THBS1-F:

5'-TGGAGCGGAAAGACCACT-3'

THBS1-R:

5'-CTGGCCAGGTCTCTGGT-3'

S100A2-F:

5'-ATGTGCAGTTCTCTGGA-3'

S100A2-R:

5'-TCAGGGTCGGTCTGGGCA-3'

S100A4-F:

5'-ATGGCGTGCCCTCTGGAGA-3'

S100A4-R:

5'-TTCCTGGGCTGCTTATCTG-3'

AXL-F:

5'-CGAGGTACATTGGCTTC-3'

AXL-R:

5'-TGGAACATGCAGGCTGC-3'

ACTIN-F:

5'-GCATGGGTCAGAAGGATTC-3'

ACTIN-R:

5'-CATCTCTTGCTCGAAGTCC-3'

HOXB2-F:

5'-CCTTTAGCCGTTCGCTTAGAG-3'

HOXB2-R:

5'-AAGTTGAGGTCGGGAAGGAAA-3'

OLIGODT:

5'- TTTTTTTTTTTTTTTTTTTT-3'

Oligonucleotides for Real-time quantitative RT-PCR:

IL-6-F:

5'-AACCTGAACCTTCCAAAGATGG-3'

IL-6-R:

5'-TCTGGCTTGTTCCTCACTACT-3'

TGFB1-F:

5'-CAACAATTCCTGGCGATACCT-3'

TGFB1-R:

5'-GCTAAGGCGAAAGCCCTCAAT-3'

TGFB2-F:

5'-CCCCGGAGGTGATTTCCATC-3'

TGFB2-R:

5'-CAGACAGTTTCGGAGGGGA-3'

ECAD-F:

5'-CGAGAGCTACACGTTCACGG-3'

ECAD-R:

5'-GGCCTTTTGACTGTAATCACACC-3'

VIM-F:

5'-AGAACTTTGCCGTTGAAGCTG-3'

VIM-R:

5'-CCAGAGGGAGTGAATCCAGATTA-3'

SNAIL-F:

5'-AATCGGAAGCCTAACTACAGCG-3'

SNAIL-R:

5'-GTCCCAGATGAGCATTGGCA-3'

SIP1-F:

5'-GCTCCGAAGCTGGCAAGAA-3'

SIP1-R:

5'-GGGACTTGTCACTATGCAGGTT-3'

CD44-F:

5'-GTGCTACTTCAGACAACCACAA-3'

CD44-F:

5'-GTACTACTAGGAGTTGCCTGGAT-3'

CD24-F:

5'-CTCCTACCCACGCAGATTTATTC-3'

CD24-R:

5'-AGAGTGAGACCACGAAGAGAC-3'

ACTIN-F:

5'-CATGTACGTTGCTATCCAGGC-3'

ACTIN-F:

5'-CTCCTTAATGTCACGCACGAT-3'