SUPPLEMENTAL INFORMATION

EXTENDED EXPERIMENTAL PROCEDURES

Animal Studies

For orthotopic lung tumorigenesis assays, 1.0x 10^4 cells were resuspended in a 1:2 mixture of PBS and growth factor-reduced Matrigel (BD Biosciences) and injected into the lung through the left rib cage of mice (Nguyen et al., 2009). Tumor growth rate in the lung/pleura was measured as a function of lung photon flux in live animals.

Orthotopic mammary fat pad injections in athymic mice were performed as previously described (Minn, 2005). Briefly, cells were resuspended in 1:1 PBS:Matrigel to a concentration of 2.5 x 10^6 cells/ml. Mice were anaesthetized by controlled isoflurane (Henry Schein, Isothesia) administration through a nose cone in a sterile hood. Depth of anesthesia was confirmed by absence of reflex to toe-pinching. The skin covering the fourth nipple was disinfected with three alternate scrubs of 70% Ethanol and Povidone-Iodine. An incision was made between the fourth and fifth nipple of the mouse to expose the mammary fat pad, and 100uL of the cell suspension injected using a 28G insulin syringe. Veterinary glue (3M, Vetbond) and 9mm stainless steel wound clips (Reflex Wound Clip System) were used to close the incision site. Buprenorphine was injected sub-cutaneously to control post-procedural pain in these mice. Wound clips were removed 10 days post-injection. Tumors were monitored weekly by measurement using a calibrated digital caliper (VWR, Cat. 62739-531). For surgical implantation of patient derived tissue (HCI-001, HCI-002 and HCI-008) similar procedure was followed with the following exceptions. Fresh patient derived xenograft tumors were harvested and chopped into small pieces, dissecting out fat and necrotic mass. A small tumor tissue piece was then implanted into the mammary fat pad. Tumors were monitored weekly by measurement using a calibrated digital caliper. Two months post implantation grossly normal organs were harvested and sectioned to detect disseminated tumor cells. For transplantation assays, cells were re-suspended media with matrigel and injected subcutaneously. Palpable tumors were scored.

1.0 x 10^5 cells resuspended in 50µl HBSS were used to perform intrasplenic injections as previously described (Winslow et al., 2011). Briefly, mice were anaesthetized and an incision was made between the left abdominal and thoracic regions and spleen was exposed. Cells were injected into the spleen, allowed to pervade for 10 minutes, followed by splenectomy. Intrasplenically injected mice were analyzed 3-4 weeks post injection.

In vivo depletion of NK cells was achieved by intraperitoneal injection of either 100 µl of Anti asialo-GM1 (Wako Chemicals Cat. 98610001) or 300ul of PK136 antibody (Supernatant from hybridoma cell line; gift from Sun lab) every five days from the start of the NK cell depletion regimen. NK cell depletion was monitored by labeling NK cells from peripheral blood with NK1.1 (BD Biosciences) and analyzed by flow cytometry (BD Fortressa, BD Biosciences).

Cell Lines and Recombinant Proteins

HCC1954, H2030-BrM, T-nonMet (368T1) and T-Met (482T1) were cultured in RPMI 1640 media supplemented with 10% Fetal Bovine Serum (FBS), 2 mM L-Glutamine (Glu), 100 IU/ml Penicillin/Streptomycin (P/S) and 1 µg/ml amphotericin B (ampoB). MDA231-BrM2, MDA23-LM2, and 4T07 were cultured in DME media supplemented with 10% FBS, 2 mM Glu, 100 IU/ml P/S and 1 µg/ml ampoB. H2087 (ATCC) was cultured in RPMI 1640 media supplemented with 10% FBS, 2 mM Glu, 100 IU/ml P/S, 1 µg/ml ampoB, 0.5 mM Sodium Pyruvate, 10 mM HEPES, 50 nM Hydrocortisone, 25 nM Sodium Selenite, 20ug/ml Insulin, 10 ug/ml Transferrin (SITE), 0.5% Bovine Serum Albumin (BSA), and 1ng/ml recombinant human EGF. 293T was cultured in DME media supplemented with 10% FBS. Culture conditions above containing 10% FBS were defined as Mitogen-Rich Media (MRM), whereas culturing in media containing 2% (H2087) or 0.2% (HCC1954) FBS was defined as Mitogen-Low Media (MLM). For *in vitro* experiments, recombinant WNT3A and DKK1 were purchased from R&D systems, and recombinant IL-2 was purchased from Peprotech.

Immunofluorescence, Confocal Microscopy and Image Analysis

Harvested organs were fixed in 4% paraformaldehyde overnight at 4° C and washed in 1X PBS. Organs were cryo-protected by sequential immersion in 15% and 30% sucrose. Cryo-protected organs were mounted using OCT (Sakura) onto a sliding microtome outfitted with a platform freezing unit (Thermo Scientific, Cat. Microm KS-34 and Microm HM-450). 80µm sections were cut and sequentially stored in anti-freezing solution (30% ethylene glycol, 30% glycerol in PBS) at -20° C. Free-floating frozen sections representative of the entire organ of interest were permeabilized with three consecutive washes in PBS supplemented with 0.25% Triton X-100 (PBS-T). If required, antigen unmasking was done by incubation of sections in citrate buffer (Vector Labs, H-3300) and steamed for 30 minutes. Sections were cooled to room temperature and rinsed twice in cold 1X PBS, followed by incubation in blocking buffer (10% normal goat serum (Life Technologies) supplemented with 2% BSA (Gemini Bio-Products) and 0.25% Triton

X-100) for 1 hour. Blocked sections were incubated in primary antibodies diluted in blocking buffer overnight at 4° C. Sections were thoroughly washed in PBS-T for at least six times before incubation in secondary antibodies (1:500, Life Technologies, respective Alexa-Fluor conjugated antibodies raised in goat) for 2 hours. Sections were washed three times in PBS-T, followed by three washes in PBS. Sections were counter-stained with the nuclear dye DAPI or Hoechst 33342 (1:1000) for 5 minutes, followed by two washes in PBS. Sections were transferred onto slides and allowed to air-dry until translucent, followed by mounting using ProLong diamond antifade mountant (Life Technologies, Cat. P36970).

Primary antibodies used: GFP (Aves Labs, Cat. GFP-1020), Collagen IV (Millipore, Cat. AB756P, antigen unmasking required), DyLight 594 labeled GS-I isolectin B4 (Vector Labs, Cat. DL-1207), Human Vimentin (Vector labs, VP-V684), CD31 (Dianova, DIA-310), Ki-67 (Vector labs, VP-K451), CD45 (abcam, ab10559), F4/80 (abcam, ab6640), Ly6B.2 (Bio-Rad, MCA771GA) Sox2 (abcam, ab97959, ab59776), Sox9 (abcam, ab185230), β-Catenin (CST, 8814) and DKK1 (abcam, ab61034). Respective AlexaFlour secondary antibodies from Life Technologies were used for staining. Images were acquired with an SP5 confocal microscope (Leica Microsystems) and micrographs analyzed with ImageJ, Imaris and Metamorph software packages.

Gene-Expression Analysis

Whole RNA was isolated from cells using RNAeasy Mini Kit (QIAGEN) or PrepEase RNA spin kit (USB). Transcriptor First Strand cDNA synthesis kit (Roche) was used to generate cDNA. Relative gene expression was determined using Taqman assays (Life Technologies). Assays used for human genes are: Sox2 (Hs01053049-M1), Sox9 (Hs01001343_g1), DKK1 (Hs00183740_M1), Axin2 (Hs00610344_M1), CD155 (Hs00197846_M1), ULBP1 (Hs04194603_S1), ULBP2 (Hs00607609_mH), MICA (Hs04233590_s1), Fas (Hs00163653_m1), TRAILR1 (Hs00269492_m1), TRAILR2 (Hs00366278_m1), ICAM1 (Hs00164932_M1), RAET1E/ULBP4 (Hs01026643_M1) and RAET1G/ULBP5 (Hs01584111_M1). Relative gene expression was normalized to the housekeeping gene β2M (Hs99999907_m1) or GAPDH (Hs99999905_m1). Quantitative PCR was performed on ViiA 7 Real-Time PCR System (Life Technologies) and analyzed using companion software.

Immunoblotting

Cells in described culture conditions were washed with 1X PBS and lysed using RIPA cell lysis buffer (Cell Signaling Technology) supplemented with protease (Roche, *cOmplete, mini, EDTA-free protease inhibitor tablets*, Cat. 11836170001) and phosphatase (Thermo Scientific, *Halt Phosphatase Inhibitor Cocktail*, Cat. 78427, 1:1000) inhibitors. Total protein concentration was determined using BCA Protein Assay Kit (Pierce). Proteins were separated in NuPAGE Novex 4-12% Bis-Tris gels using 1X MOPS SDS running buffer, and transferred to nitrocellulose membranes. Membranes were immunoblotted with primary antibodies against Sox2, Sox9, p38, p-p38, ERK, p-ERK, (Cell Signaling Technology) and beta-tubulin (Sigma-Aldrich). Proteins of interest were detected using near-infrared (NIR) secondary antibodies captured on an Odyssey CLx infrared imaging system (LI-COR Biosciences). Human DKK1 Quantikine ELISA KIT (R&D systems) was used to determine DKK1 protein levels in the supernatant of the described 3-day cultures.

Knockdown and Overexpression Constructs

Stable DKK1 knockdown in H2087 cells was generated using two independent shRNAs in pGIPZ lentival vector (Open Biosystems) targeting the following sequences: GCCATAGCAAGATTGCTTA (shRNA1) and GGGAGTATATGAATGTGAA (shRNA2)(Vanharanta et al., 2014). H2087 Sox2 (TRCN0000085748-52) knockdown cells were generated using two independent shRNAs in pTRIPZ lentival vector (Open Biosystems). Sox9 knockdown in HCC1954 cells were generated using 2 independent shRNA hairpins cloned into the mIR-E SGEN vector targeting the following mRNA sequences: (shSox9-1: ATCGTGTGATCAGTGTGCTAAA, shSox9-2: ATCCTGTTGTATTAACATTTAA). pCS2-hDKK1-flag, a gift from Sergei Sokol (Addgene plasmid # 15494) was subcloned into the pLVX-puro lentiviral expression vector using EcoRI and XbaI restriction enzyme sites.

Cell Surface Protein Expression Analysis

Adherent cells were detached using 0.25% Trypsin to generate single cell suspensions, and resupended in FACS buffer (1X PBS, 0.25mM EDTA, 2% FBS). Cells were incubated with fluorochrome-conjugated monoclonal antibodies for 15-30min, and washed three times in FACS buffer. Cell surface expression of proteins was analyzed by flow cytometry on a BD Fortressa (BD Biosciences). APC conjugated secondary antibodies were used to analyze expression of cell-surface proteins CD155 (R&D, MAB25301) and ULBP2/5/6 (R&D, MAB1298).

Cell Proliferation and Viability Assays

eFluor670 (eBioscience, Cat. 65-0840-90) dye was used to track proliferation for cell under described culture conditions according to the manufacturer's recommendation. The amount of dye retention was analyzed 6 days post labeling by flow cytometry using a BD FACSCalibur (BD Biosciences) fitted with a 660/20 band pass filter to detect the APC channel. Cell cycle state analysis was performed using APC BrdU Flow Kit (Cat. 557892). Cancer cells cultured for 3 days in either MRM or MLM conditions, were pulsed with BrDU for 15 minutes and then fixed with PFA. Cell cycle analysis was performed using FACSCalibur flow cytometer (BD Biosciences). Cell lines were labeled with EdU for 24 hours; pulsed every 6 hours before injecting into mice. Cells were also plated *in vitro* to monitor proliferation and dye retention. Harvested organs were sectioned and stained using Click-iT® plus EdU Alexa Fluor 594 Imaging Kit (Cat. C10639).

Cancer cells were cultured under described conditions in 96-well plates (Corning Costar Cat.3610). Cell proliferation and cell death were detected by CellTiter-Glo® (Promega) and Caspase-Glo® (Promega) assays respectively at indicated time points. Cell viability in cells infected with 7TP WNT responsive construct was measured using SYTO® 60 Red Fluorescent Nucleic Acid Stain (S11342) according to manufacture's protocol. Cells were grown in 3D matrigel cultures as described before (Shibue et al., 2012).

Oncosphere Assays

Cells were grown in oncosphere media (RPMI 1640 supplemented with 20ng/mL of recombinant human EGF (Life Technologies, PHG0311) and bFGF (StemCell, 02634) in 24-well ultra-low attachment plates (Corning, 24-well plate) at a density of 1,000 cells per well. Spheres were allowed to grow for 10-21 days and quantified on an EVOS imaging system (Life Technologies) Oncospheres were embedded in paraffin, for sectioning and immunostaining.

NK Cytotoxicity Assays

Splenocyte suspensions were prepared by mechanical dissociation. NK cells were purified from these suspensions by magnetic depletion of non-NK cells using magnetic assisted cell sorting (MACS) using the NK Cell Isolation Kit (Miltenyi Biotec). NK cell media consisted of RPMI 1640 containing 10% FBS, β-2ME, nonessential amino acids, 10mM HEPES, 0.5 mM sodium pyruvate, 2 mM L-glutamine, and 10 IU/ml penicillin/streptomycin. NK cells were expanded in NK cell media containing 1000 U/ml recombinant IL-2 overnight. Target cells labeled with

eflour670 dye were incubated with or without NK (effector to target ratio of 1:5) cells for 3 hrs at 37° C and cytotoxicity was assessed by flow cytometry.

ChIP

For Sox2-ChIP assay, we used the ChIP Assay Kit (Millipore) according to the manufacturer's instructions. Cells were sonicated in 20s pulses for a total of 3 minutes. Anti-human Sox2 antibody (abcam, ab15830) was used. DNA was extracted using the QIAquick PCR Purification Kit (Qiagen) followed by qPCR using the previously described primers (Park et al., 2012).

Chromatin immunoprecipitation was performed as described in (Basnet et al., 2014). In all the experiments, cells were fixed with 1% formaldehyde. Antibodies used were RNA polymerase Abcam (ab817) and H3K27ac Active Motif (39133).

RNA-Seq, ChIP-Seq and Downstream Bioinformatic Analysis

Total RNA isolated from H2087 or HCC1954 derivatives grown in either MRM or MLM conditions were quality-checked with an Agilent BioAnalyzer 2000. RNA with an integrity number of greater than 9.5 were used for subsequent analyses. Libraries were prepared with either TruSeq RNA Sample Prep Kit v2 (Illumina) or NEBNext Ultra RNA Library Prep Kit (New England Biolabs). Generated libraries were run on a HiSeq2500 system and 25-50 million raw reads were collated per sample. Raw reads were quality-checked and subsequently mapped to the human genome (hg19) using Tophat2 (2.2.4) using default settings(Langmead and Salzberg, 2012). The number of reads uniquely mapped to a specific gene were counted with HTSeq (0.6.1p1) (Anders et al., 2015). Downstream bioinformatic analyses was performed in RStudio (0.98.1103) implementing R (3.2.0). Differential gene expression was analyzed using the DESeq2 (1.8.1) package in R using default settings (Love et al., 2014). Principal component analyses were generated using the *prcomp* function in R and plotted with *ggplot2*. Human and mouse gene homologues are matched using the Mouse Genome Informatics annotation Heatmaps were generated using the *heatmap.2* function in R. Pathway signature scores were calculated using the sum of z-score or gene-set variation analysis methods using previously-reported signatures (Bild et al., 2006; Gatza et al., 2010; Hanzelmann et al., 2013; Nguyen et al., 2009; Zhang et al., 2013).

ChIP-Seq read mapping: 50/50bp paired-end raw FASTQ files were mapped to human genome hg19 by Bowtie2 with default settingS in the aligner(Langmead and Salzberg, 2012). Output

SAM files are converted to BAM format with "*samtools view*", sorted with "*samtools sort*" and indexed with "*samtools index*" from Samtools (Li et al., 2009). Sorted BAM files WERE converted to TDF file format with "*igvtools count -z 5 -w 25 -e 250*" from IGVTools(Robinson et al., 2011), specifying the maximum zoom level at 5bp, window size of 25bp for averaged coverage, and read extension to 250bp(Robinson et al., 2011). TDF files were loaded to Integrative Genomics Viewer 2.3 for coverage visualization. Coverage was normalized to Reads Per Million Reads (RPM) or Counts Per Million TotalCount prior to visualization for cross sample comparisons(Robinson et al., 2011).

ChIP-Seq Peak Calling: H3K27ac peaks were called by "*findPeaks -style histone -F 5 -P 0.00001 -L 5 -LP 0.00001*" function from HOMER (v4.2), specifying the ChIP-Seq data style as broad histone modifications, cut-off for fold enrichment over input or local tag count as 5, cut-off for poisson p-value threshold relative to input or local tag count as 0.00001(Heinz et al., 2010). Sequencing data from input samples were used as background control. To identify parental (H2087 or HCC1954 Parental) or derivatives (H2087-LCC2 or HCC1954-LCC2) specific peaks, the same peak calling method was used whiling using the second experiment as the background control sample (Heinz et al., 2010). Sample specific peaks were randomly picked and validated in the IGV visualization window. Tag distribution matrix was calculated with "*annotatePeaks.pl hg19 -size 6000 -hist 50 -ghist*" function from HOMER (v4.2), specifying calculated region as +/-3000bp from peak center, and window size as 50bp. Heatmap for tag distribution matrix was plotted in R (v3.1.2) with "*heatmap.2*" function from *gplot* package.

Statistics

Data was analyzed and statistics were performed in Prism5 (Graphpad). *P* values shown were calculated using a two-tailed Mann- Whitney test unless other wise noted (* $p<0.05$, ** $p<0.01$, *** $p<0.001$, n.s.= not significant).

REFERENCES

Anders, S., Pyl, P.T., and Huber, W. (2015). HTSeq--a Python framework to work with high-throughput sequencing data. Bioinformatics *31*, 166-169.

Basnet, H., Su, X.B., Tan, Y., Meisenhelder, J., Merkurjev, D., Ohgi, K.A., Hunter, T., Pillus, L., and Rosenfeld, M.G. (2014). Tyrosine phosphorylation of histone H2A by CK2 regulates transcriptional elongation. Nature *516*, 267-271.

Bild, A.H., Yao, G., Chang, J.T., Wang, Q., Potti, A., Chasse, D., Joshi, M.B., Harpole, D., Lancaster, J.M., Berchuck, A.*, et al.* (2006). Oncogenic pathway signatures in human cancers as a guide to targeted therapies. Nature *439*, 353-357.

Gatza, M.L., Lucas, J.E., Barry, W.T., Kim, J.W., Wang, Q., Crawford, M.D., Datto, M.B., Kelley, M., Mathey-Prevot, B., Potti, A.*, et al.* (2010). A pathway-based classification of human breast cancer. Proceedings of the National Academy of Sciences of the United States of America *107*, 6994-6999.

Hanzelmann, S., Castelo, R., and Guinney, J. (2013). GSVA: gene set variation analysis for microarray and RNA-seq data. BMC bioinformatics *14*, 7.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol. Cell *38*, 576-589.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nature Methods *9*, 357-359.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and Genome Project Data Processing, S. (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics *25*, 2078-2079.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. *15*, 550.

Nguyen, D.X., Chiang, A.C., Zhang, X.H., Kim, J.Y., Kris, M.G., Ladanyi, M., Gerald, W.L., and Massague, J. (2009). WNT/TCF signaling through LEF1 and HOXB9 mediates lung adenocarcinoma metastasis. Cell *138*, 51-62.

Park, S.B., Seo, K.W., So, A.Y., Seo, M.S., Yu, K.R., Kang, S.K., and Kang, K.S. (2012). SOX2 has a crucial role in the lineage determination and proliferation of mesenchymal stem cells through Dickkopf-1 and c-MYC. Cell Death and Differ *19*, 534-545.

Robinson, J.T., Thorvaldsdottir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G., and Mesirov, J.P. (2011). Integrative genomics viewer. Nature Biotech *29*, 24-26.
Shibue, T., Brooks, M.W., Inan, M.F., Reinhardt, F., and Weinberg, R.A. (2012). The outgrowth of micrometastases is enabled by the formation of filopodium-like protrusions. Cancer Dis. *2*, 706-721.
Vanharanta, S., Marney, C.B., Shu, W., Valiente, M., Zou, Y., Mele, A., Darnell, R.B., and Massague, J. (2014). Loss of the multifunctional RNA-binding protein RBM47 as a source of selectable metastatic traits in breast cancer. Elife *3*.
Winslow, M.M., Dayton, T.L., Verhaak, R.G., Kim-Kiselak, C., Snyder, E.L., Feldser, D.M., Hubbard, D.D., DuPage, M.J., Whittaker, C.A., Hoersch, S.*, et al.* (2011). Suppression of lung adenocarcinoma progression by Nkx2-1. Nature *473*, 101-104.
Zhang, X.H., Jin, X., Malladi, S., Zou, Y., Wen, Y.H., Brogi, E., Smid, M., Foekens, J.A., and Massague, J. (2013). Selection of bone metastasis seeds by mesenchymal signals in the primary tumor stroma. Cell *154*, 1060-1073.