Supplemental methods

Quantitative DNA methylation analysis

DNA methylation analysis was carried out using the Epityper system from Sequenom (San Diego, CA). The EpiTYPER assay is a tool for the detection and quantitative analysis of DNA methylation using base-specific cleavage of bisulfite-treated DNA and Matrix-Assisted Laser Desorption/Ionization Time-of-Flight Mass Spectrometry (MALDI-TOF MS). Specific PCR primers for bisulfite-converted DNA were designed using the EpiDesigner software (www.epidesigner.com), for the entire CpG island of the genes of interest. T7-promoter tags are added to the reverse primer to obtain a product that can be *in vitro* transcribed, and a 10-mere tag is added to the forward primer to balance the PCR conditions. For primer sequences, target chromosomal sequence, and Epityper specific tags, (see Fig S11). One µg of tumor DNA was subjected to bisulfite treatment using the EZ-96 DNA methylation Kit, which results in the conversion of unmethylated cytosines into uracil, following the manufacturer's instructions (Zymo Research, Orange, CA). PCR reactions were carried out in duplicate, for each of the 2 selected primer pairs, for a total of 4 replicates per sample. For each replicate, 1 µl of bisulfite-treated DNA was used as template for a 5 µl PCR reaction in a 384-well microtiter PCR plate, using 0.2 units of Kapa2G Fast HotStart DNA polymerase (Kapa Biosystems, Cape Town, South Africa), 200 µM dNTPs, and 400 nM of each primer. Cycling conditions were: 94 ^{0}C for 15 minutes, 45 cycles of 94 ^{0}C for 20 seconds, 56 ^{0}C for 30 seconds, 72 ^{0}C for 1 minute, and 1 final cycle at 72 ^{0}C for 3 minutes. Unincorporated dNTPs were deactivated using 0.3 U of shrimp alkaline phosphatase (SAP) in 2 µl, at 37 ^{0}C for 20 minutes, followed by heat inactivation at 85 ^{0}C for 5 minutes. Two µl of SAP-treated

reaction were transferred into a fresh 384-well PCR plate, and *in vitro* transcription and T cleavage were carried out in a single 5 µl reaction mix, using the MassCleave kit (Sequenom) containing 1 X T7 polymerase buffer, 3 mM DTT, 0.24 µl of T Cleavage mix, 22 units of T7 RNA and DNA polymerase, and 0.09 mg/ml of RNAseA. The reaction was incubated at 37 ^{0}C for 3 h. After the addition of a cation exchange resin to remove residual salt from the reactions, 10 µl of Epityper reaction product were loaded onto a 384-element SpectroCHIP II array (Sequenom). SpectroCHIPs were analyzed using a Bruker Biflex III matrix-assisted laser desorption/ionization–time of flight (MALDI-TOF) mass spectrometer (SpectroREADER, Sequenom). Results were analyzed using the Epityper Analyzer software, and manually inspected for spectra quality and peak quantification.

Migration, invasion and transendothelial assay

A scratch-wound–closure assay was used to compare migration of SN12C and LM cells. Monolayers were wounded by two perpendicular linear scratches across each well with a 10-µL pipette tip, to produce 300-µm-wide strips. Wounds were photographed at the indicated times.

For invasion assay cells (1×10^5) were counted using Trypan blue reagent (Sigma-Aldrich, St. Louis, MO), normalized for the number of viable cells, and placed in SFM on Transwell inserts coated with 2 µg Matrigel. Following incubation in 24-well plates containing complete medium for 6–8 hours, the noninvasive cells were removed with a cotton swab. The cells that have migrated through the membrane and stuck to the lower surface of the membrane were fixed with methanol and scored under a microscope.

For transendothelial assay, HUVECs were seeded into collagen-coated trans-well inserts (1 μm pore size, BD Falcon) at 100,000 cells per well, and allowed to grow to confluence for 4 days. Tumour cells were pulsed with 5 μM cell tracker green (Invitrogen, Carlsbad, CA) for 30 min before being conditioned overnight in 0.2% FBS ECM media without growth factors. The next day, 50,000 tumour cells were seeded into trans-well inserts with or without a confluent endothelial monolayer, and the wells were fixed in 4% paraformaldehyde after 10 h. Cells on the apical side of each insert were scraped off and the trans-well membrane mounted onto slides. Migration to the basolateral side of the membrane was visualized with a microscope.

Bioinformatics

Affymetrix CEL files were normalized using the RMA preprocessing algorithm and significantly differentially expressed genes were identified using the Significance Analysis of Microarrays (SAM) algorithm [24]. SAM analysis was performed on the normalized RMA data with a two-class unpaired t-test option and 500 iterations. In accordance with Minimum Information About a Microarray Experiment guidelines, raw data files for all cell lines and clinical specimens have been deposited in the Gene Expression Omnibus public repository (GEO accession no. GSE 23631). Genomic copy number data and their integration with gene expression data were performed using the Partek® software, version 6.5 (Partek Inc., St. Louis, Missouri). The Partek segmentation algorithm was used to identify significant copy number change regions following quantile normalization of the raw data. The smoothing parameters for this segmentation algorithm were set to 'minimum genomic markers' as 7, 'the segmentation p-value threshold' as 0.001, and the signal-to-noise at 0.3. The reporting parameters of

a region called to be aberrant were set as 'expected range' at 0.3 with a P-value of 0.01. Regions of significant gain or loss were filtered based on the category attribute, here LM versus the parental. Aberrant regions were annotated to the corresponding genes and correlated with the expression data obtained from SAM analysis. Data from the Agilent miRNA profiles were quantile-normalized using the Partek Genome Suite and the significantly differentiated miRNAs were identified by SAM analysis. The GLAD (Gain and Loss Analysis of DNA) algorithm aims at identifying the chromosomal regions with identical DNA copy number, which is delimited by breakpoints. More details about the GLAD algorithm can be found in Hupé P, et al. Analysis of array CGH data: from signal ratio to gain and loss of DNA regions. Bioinformatics (2004) 20:3413–3422. The GLAD algorithm segments the genomic profile, defining regions of homogeneous DNA copy number. This function allows the detection of breakpoints in genomic profiles obtained by array CGH technology, for each of these regions; it provides a smoothing value and a status (gain, normal or loss).

Analyisis of mRNA expression

Four-hundred nanograms of total purified RNA was subjected to a reverse transcriptase reaction according to the manufacturer's protocol (invitrogen). cDNA corresponding to approximately 4 ng of starting RNA was used in three replicates for quantitative PCR. Indicated Taqman gene expression assays (Applied Biosystems) and the Taqman universal PCR master mix (Applied Biosystems) were used to quantify expression. Quantitative expression data were acquired and analyzed using an ABI Prism 7900HT Sequence Detection System (Applied Biosystems).