Supplementary Methods

Antibodies

The following antibodies were used for immunoblotting: rabbit polyclonal periostin (Abcam, ab 14041, Cambridge, MA) that recognizes both mouse and human species and β-actin (Sigma-Aldrich, St. Louis, MO) was used as a loading control.

Western blotting

To examine protein expression in cell lysates and secreted periostin expression in mouse sera, Western blots were performed as described previously[1]. Ponceau S Staining was performed as loading control for mouse sera samples.

ELISA

Mouse serum was isolated from pooled whole blood using a syringe, incubated on ice for 30 minutes and centrifuged at 14 000 rpm for 20 minutes. Mouse periostin levels in sera were measured using the mouse periostin DuoSet ELISA Development System (R&D Systems, Minneapolis, MN) according to manufacturer's instructions.

Immunohistochemistry

Rabbit polyclonal periostin (Abcam, ab 14041) antibody was used for immunohistochemistry and immunohistochemistry was performed using with the Vector Elite kit (Vector Laboratories) using manufacturer's protocol, its detailed procedures are as previously described[1]

Optical probe synthesis

An optical imaging probe for detecting periostin was generated by means of covalent modification of the periostin polyclonal antibody (Santa Cruz Biotechnology, Santa Cruz, CA),

that recognizes human and mouse species, with the cyanine 5.5 (Cy5.5, GE Healthcare, Little Chalfont, UK) near infrared dye as described previously[2]. The synthesized probe was used for all imaging studies after purification on a Sephadex G50 column (Amersham, Piscataway, NJ).

Multispectral *in vivo* imaging

TE-11 and TT cells were grown in RPMI + 10% FBS + 1% Penicillin/Streptomycin until they were ~90% confluent. Cells were trypsinized, washed three times with PBS and collected by centrifugation (5min, 1200rpm, RT). Cells were resuspended with Matrigel (BD Biosciences, Bedford, MA) in 1:1 ratio and kept on ice until injection.

Six- to 8-week-old male nude (nu/nu) mice (2 groups, n=5 each) were obtained from Charles River Laboratories International Inc., (Wilmington, MA). The tumors were established by subcutaneous injection of 20μL ($1x10^6$ cells) of the cell suspension. Imaging was performed 2-3 weeks post cell injection when tumors were approximately =150mm^3. Multispectral imaging was performed after bloodpool clearance of the imaging probe (30 μg/animal) using an in vivo multispectral whole animal optical imaging system (Carestream, Rochester, NY). Multiple excitation filters (710-750nm) and an 810 nm emission filter were used to acquire a series of wavelength specific images; this group of images then underwent spectral deconvolution processing to separate the signal generated from the activated NIR probe from autofluorescence and other non-target emission. All animal studies were approved by the respective IACUC at the University of Pennsylvania and Massachusetts General Hospital.

Multispectral upper GI endoscopy

L2-Cre;p120ctn$^{LoxP/LoxP}$ mice with mild or severe dysplasia and a group of control *L2-Cre;p120ctn*$^{LoxP/+}$mice were used for the experiments (3 groups, n=3 each group). All animals

were fed non-fluorescent diet for one week before imaging. 30μg of the imaging probe was injected into each mouse through the tail vein and endoscopy was performed after blood pool clearance of the probe. Anesthesia was induced with isofluorane and the animals were maintained at 37 °C. The animals were then intubated using a 22-guage plastic catheter that was connected to a small animal respirator (Harvard apparatus, Holliston, MA) and the anesthesia was maintained by 1.5% isoflurane[3].

The upper endoscopy was performed using a custom-made multispectral endoscopy system (one channel each for full spectrum white light imaging and near infrared imaging of the periostin reporter); the design and construction of the endoscope system has previously been described[4]. The imaging catheter was introduced through the oropharynx while air was insufflated through the endoscope working channel. Simultaneous WL and NIR images from esophagus were acquired with slow advancement of the fiberscope and recorded for analysis.

Image analysis

Analysis and deconvolution for surface weighted multispectral imaging of subcutaneous tumors was performed using Carestream Molecular Imaging and Multispectral Software (Carestream, Rochester, NY). Quantitative endoscopic image analysis of the NIR signal was performed as previously described[5] using Image J (National Institutes of Health, Bethesda, MD). Briefly, mean signal intensity in regions-of-interest (ROI) of constant size comprising 10 x 10 units of width and height was measured within the lesion, normal surrounding tissue and used to calculate the target to background ratio (TBR). The data are shown as mean±SEM. A two-tailed Student *t* test or ANOVA for independent samples was used for comparison. $P<0.05$ was considered significant.

Microarray analysis

Gene expression data from 3D cultures were generated by using Affymetrix human Exon array and data are available in NCBI GEO databases. Gene Expression Omnibus (GEO) database accession number: GSE21293. In addition to the gene expression data from 3D culture, gene expression data from 53 ESCC patients (Shanxi cohort) were used for examining clinical relevance of the tumor invasion signature. Gene expression and clinical data are available from GEO database (GSE23400). BRB-ArrayTools were primarily used for statistical analysis[6]. We identified genes that were differentially expressed among the two classes using a random-variance t-test. The random-variance t-test is an improvement over the standard separate t-test as it permits sharing information among genes about within-class variation without assuming that all genes have the same variance[7]. Genes were considered statistically significant if their p value <0.001. A stringent significance threshold was used to limit the number of false positive findings. We also performed a global test of whether the expression profiles differed between the classes by permuting the labels of which arrays corresponded to which classes. For each permutation, the p values were re-computed and the number of genes significant at the 0.001 level was noted. The proportion of the permutations that gave at least as many significant genes as with the actual data was the significance level of the global test. A p-value < 0.05 indicated statistical significance, and all statistical tests were two-tailed. We carried out all of the statistical analyses in the R language environment (http://www.r-project.org). Heatmaps were generated by using Cluster and TreeView[8].

Stratification of ESCC patients according to invasion signatures

Before applying the classification algorithm, gene expression data used as training and test sets were normalized by centralizing the gene expression level across the tissues. To stratify

ESCC patients according to the tumor invasion gene expression signature, we applied classification algorithms based on support vector machine (SVM) as described in previous studies[9, 10]. Briefly, gene expression data in training set were combined to form a classifier and robustness of classifier was estimated by misclassification rate determined during the leave-one-out cross-validation (LOOCV) in the training set[6]. The trained classifier was then applied to gene expression data from the test sets.

Supplementary Figure Legends

Supplementary Figure 1

Fluorescent images (red color) of nude mice bearing TE-11 xenograft tumors (yellow arrow) and intravenously injected with (A) fluorescent periostin (POSTN) antibody conjugated to optical probe, (B) fluorescent POSTN antibody conjugated optical probe and non-labeled antibody and (C) saline (control). (D) Quantification of fluorescent signal detected by POSTN optical probe in TE-11 xenograft tumors (n=5 per group). Bar graph represents target-to-background ratios (TBR) +/-SEM.

Supplementary Figure 2

Histopathologic analysis of esophagi isolated from mice in the following three cohorts (A) *L2-cre;p120ctn*$^{LoxP/+}$ (control); (B) *L2-cre;p120ctn*$^{LoxP/LoxP}$ with mild dysplasia and (C) *L2-cre;p120ctn*$^{LoxP/LoxP}$ with severe dysplasia. Scale bars are 100 μM.

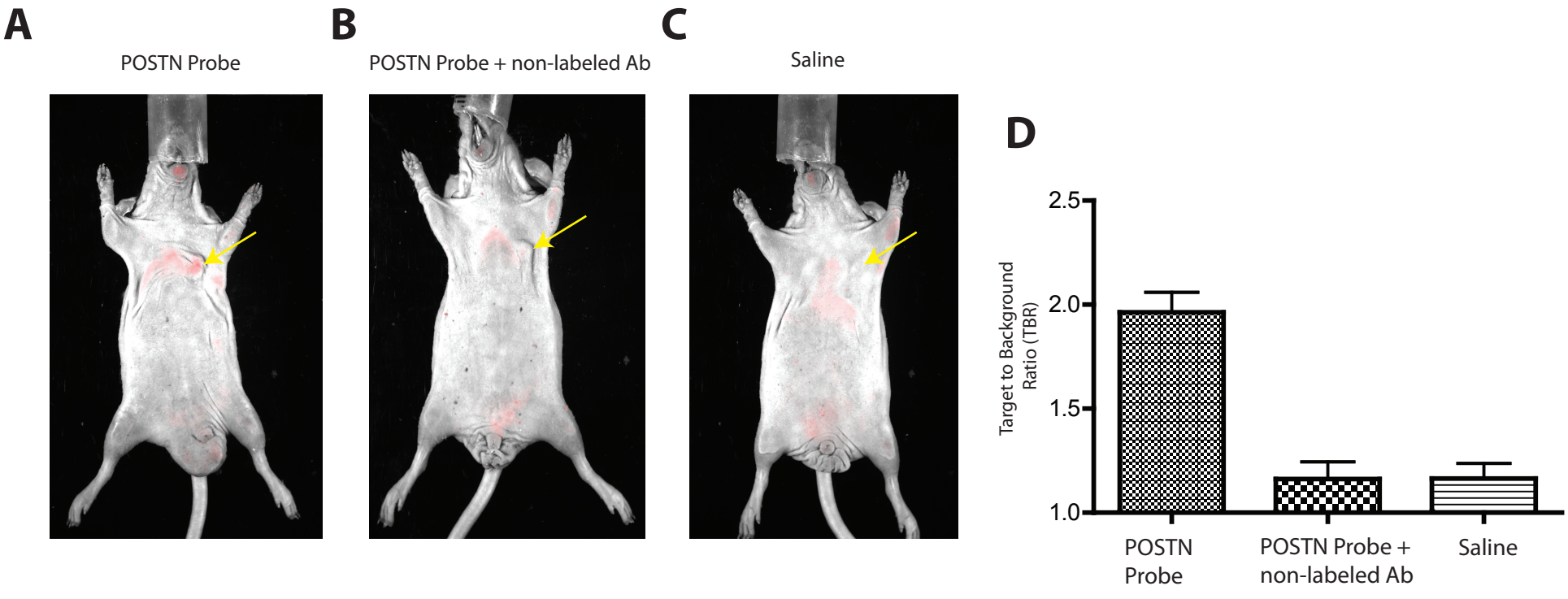
Supplementary Figure 3

Immunohistochemical analysis of periostin expression in esophagi isolated from mice in the following three cohorts (A) *L2-cre;p120ctn*$^{LoxP/+}$ (control); (B) *L2-cre;p120ctn*$^{LoxP/LoxP}$ with mild dysplasia and (C) *L2-cre;p120ctn*$^{LoxP/LoxP}$ with severe dysplasia. Scale bars are 100 μM.

Supplementary References

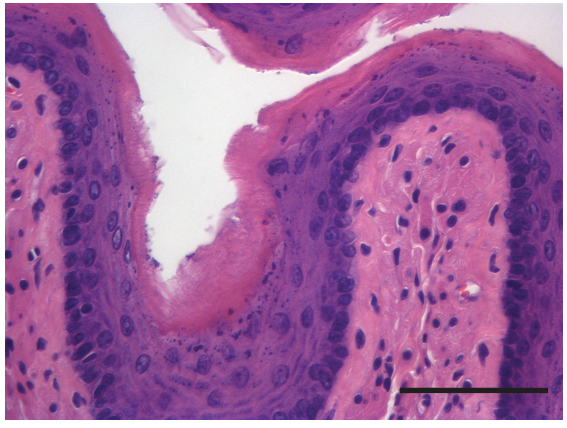
1. Michaylira CZ, et al. Cancer Res 2010;70:5281-92.
2. Gee MS, et al. Radiology 2008;248:925-35.
3. Figueiredo JL, et al. Int J Cancer 2006;118:2672-7.
4. Funovics MA, et al. Mol Imaging 2003;2:350-7.
5. Habibollahi P, et al. Theranostics 2012;2:227-34.
6. Simon R, et al. Cancer Inform 2007;3:11-7.
7. Wright GW, et al. Bioinformatics 2003;19:2448-55.
8. Eisen MB, et al. Proc Natl Acad Sci U S A 1998;95:14863-8.
9. Lee JS, et al. Hepatology 2004;40:667-76.
10. Lee JS, et al. Nat Med 2006;12:410-6.

SUPPLEMENTARY FIGURE 1

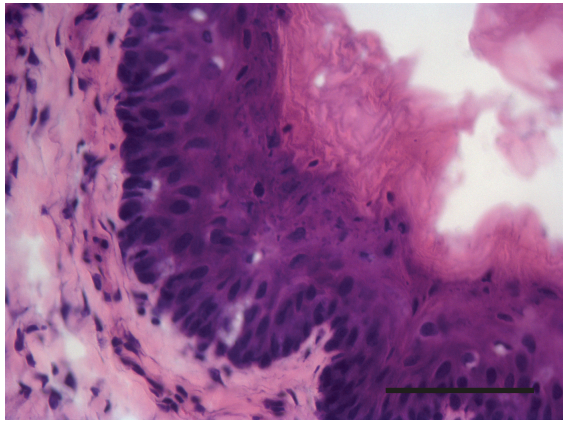


SUPPLEMENTARY FIGURE 2

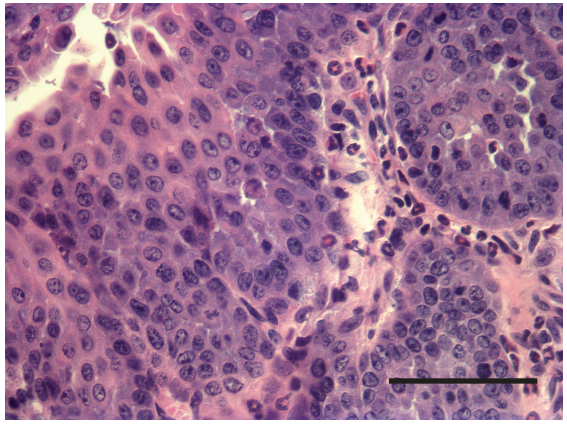
A Control



B Mild dysplasia

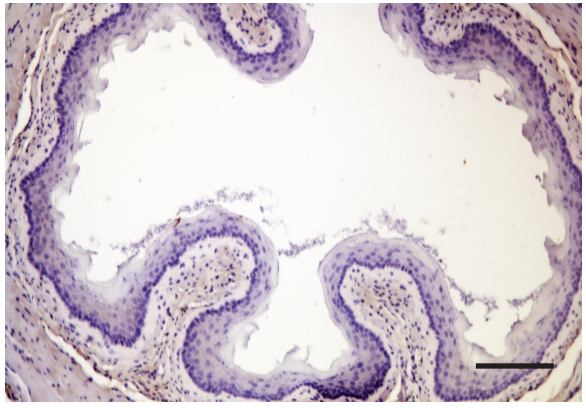


C Severe Dysplasia

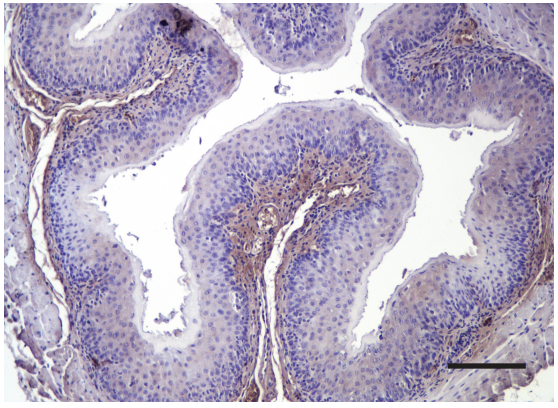


SUPPLEMENTARY FIGURE 3

Control



Mild Dsyplasia



Severe Dsyplasia