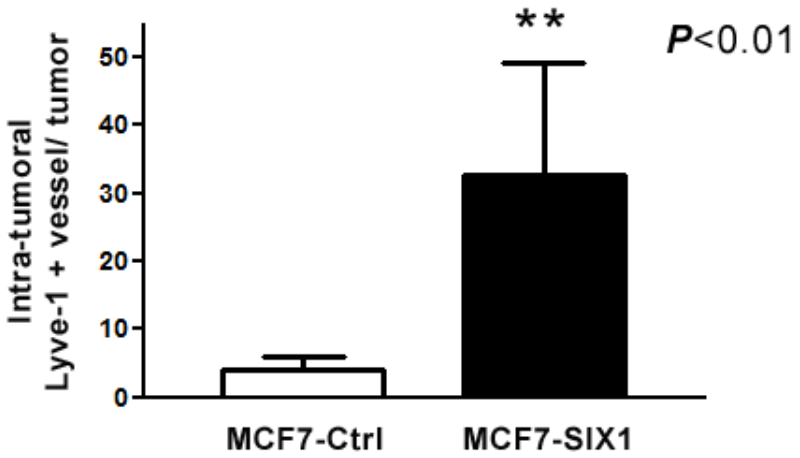
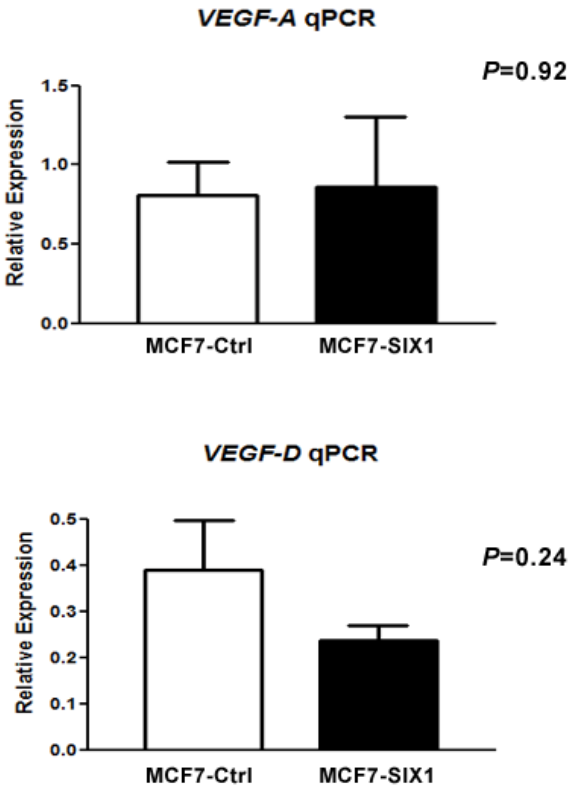
Supplementary Figures

	Nude	NOD/SCID
Location	Prevalence	Prevalence
Lymphatic system	100% (8/8)	100% (15/15)
Axillary lymph node	38% (3/8)	73% (11/15)
Lumbar lymph node	38% (3/8)	33% (5/15)
Other lymph node	25% (2/8)	27% (4/15)
Lymphatic vessel	13% (1/8)	0% (0/17)

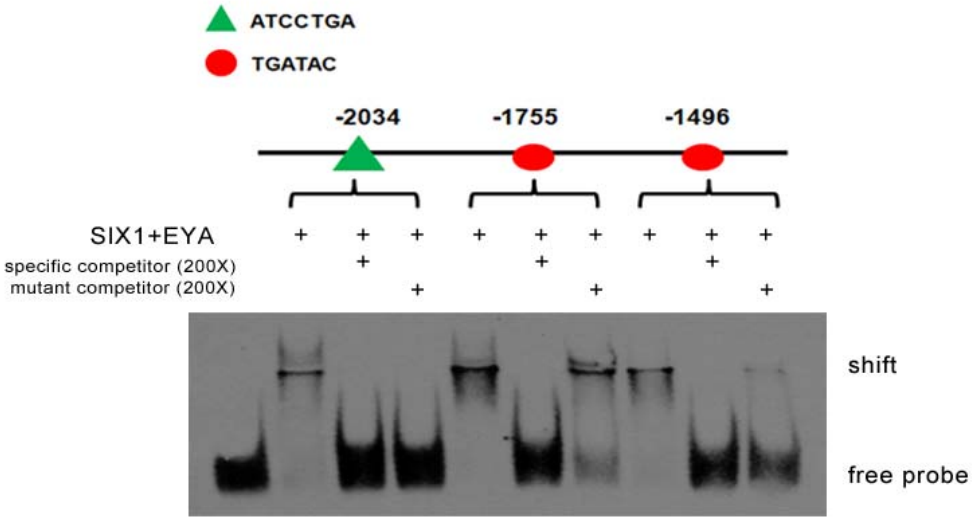
Supplementary Figure 1. Description of the observed lymphatic metastases in two different SIX1-induced MCF7 metastasis models (Nude and NOD/SCID).



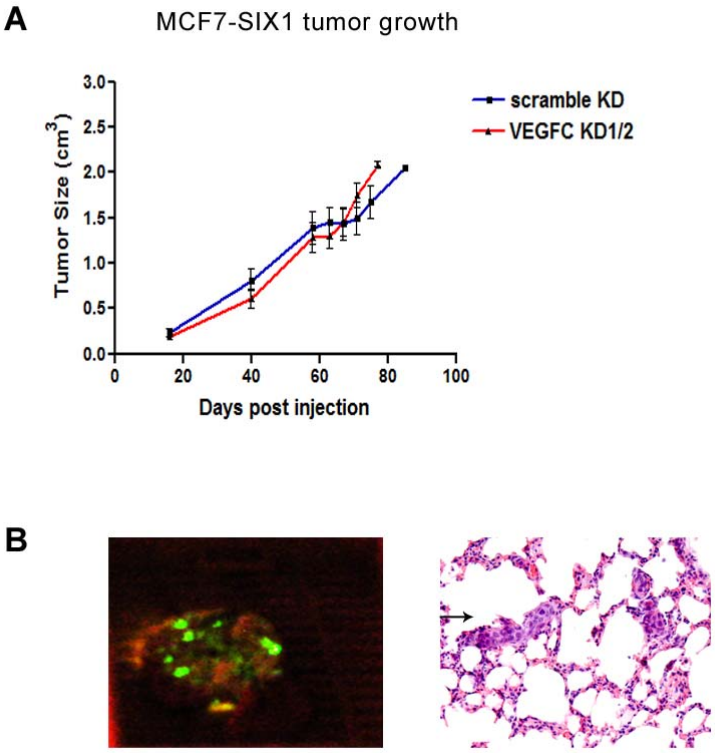
Supplementary Figure 2. MCF7-SIX1 tumors display an increased number of intra-tumoral lymphatic vessels when compared to MCF7-Ctrl tumors in NOD/SCID mice. Immunohistochemical staining was used to detect Lyve-1 in MCF7-Ctrl and MCF7-SIX1 tumors and sections were scored in a blinded manner by a pathologist. Error bars represent the standard error of the mean. Asterisks denote significant difference from control group **, $P<0.01$.



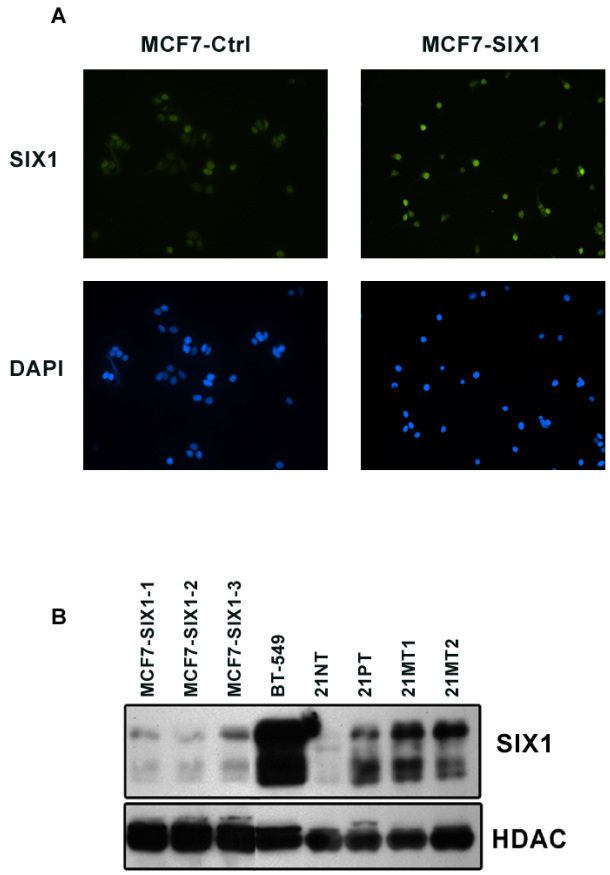
Supplementary Figure 3. *VEGF-A* and *VEGF-D* expression in MCF7-Ctrl and MCF7-SIX1 cells was detected by real-time PCR. Values represent the level of *VEGF-A* or *VEGF-D* transcript normalized to the level of the loading control *CYCLOPHILIN B*. Error bars represent the standard error of the mean.



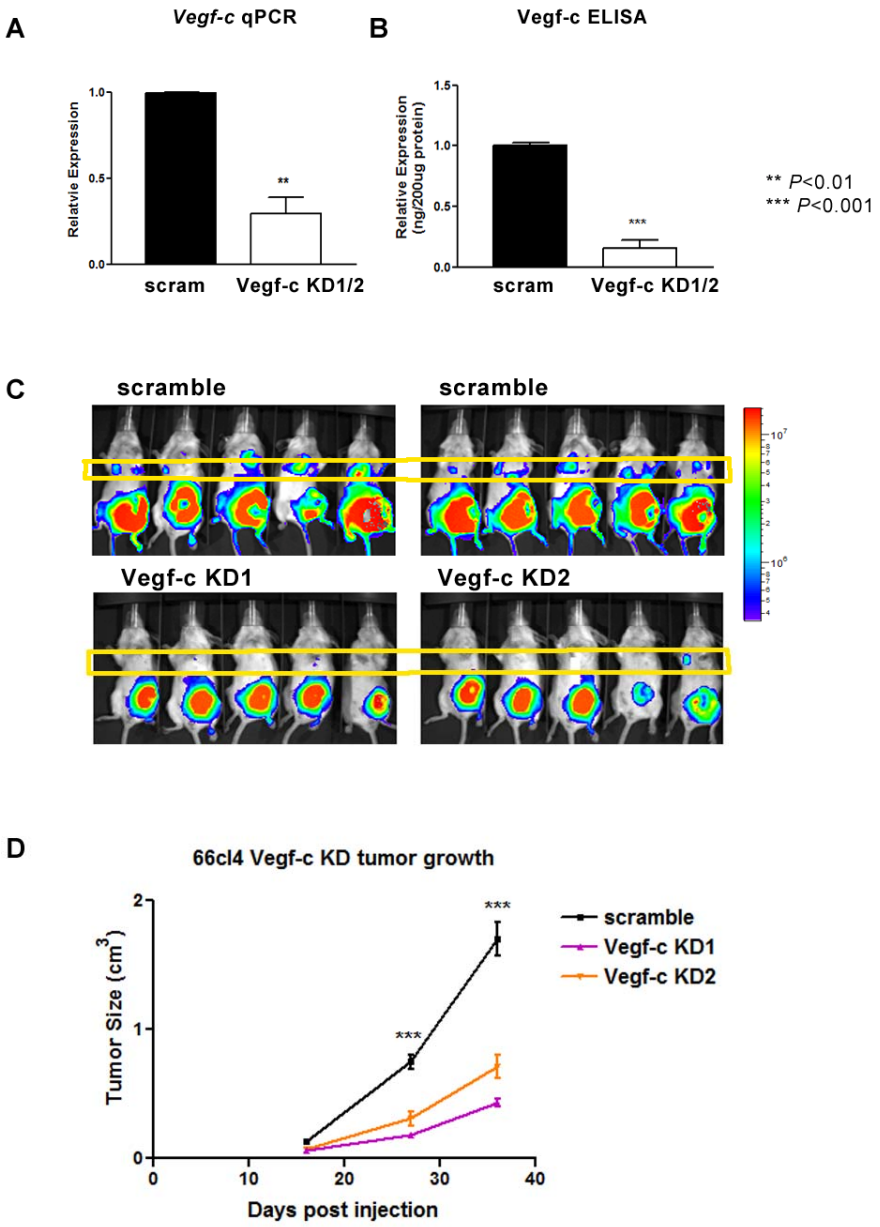
Supplementary Figure 4. Examination of SIX1 binding to different sites on the *VEGF-C* promoter using the electrophoretic mobility shift assay (EMSA). Purified protein SIX1 and its cofactor EYA were incubated with biotin labeled 30 base pair double stranded oligonucleotides spanning the predicted SIX1 binding sites at -2034, -1755, and -1496 alone, or in the presence of unlabeled competitor DNA or mutated (scrambled the core sequence) unlabeled competitor DNA in a 200-fold excess. The EMSA was performed according to the manufacturer's protocol (Thermo scientific).



Supplementary Figure 5. (A) Primary tumor growth in MCF7-SIX1-scramble KD and MCF7-SIX1-VEGFC KD1/2 (two shRNAs were used and combined in the figure). Tumor size in the animals was measured using calipers, and calculated according to the formula $V=1/2(W)^2(L)$. (B) Left panel shows representative Zsgreen positive lung from MCF7-SIX1-scramble KD tumor bearing mouse. Right panel shows histologic confirmation of lung metastasis by H& E staining.

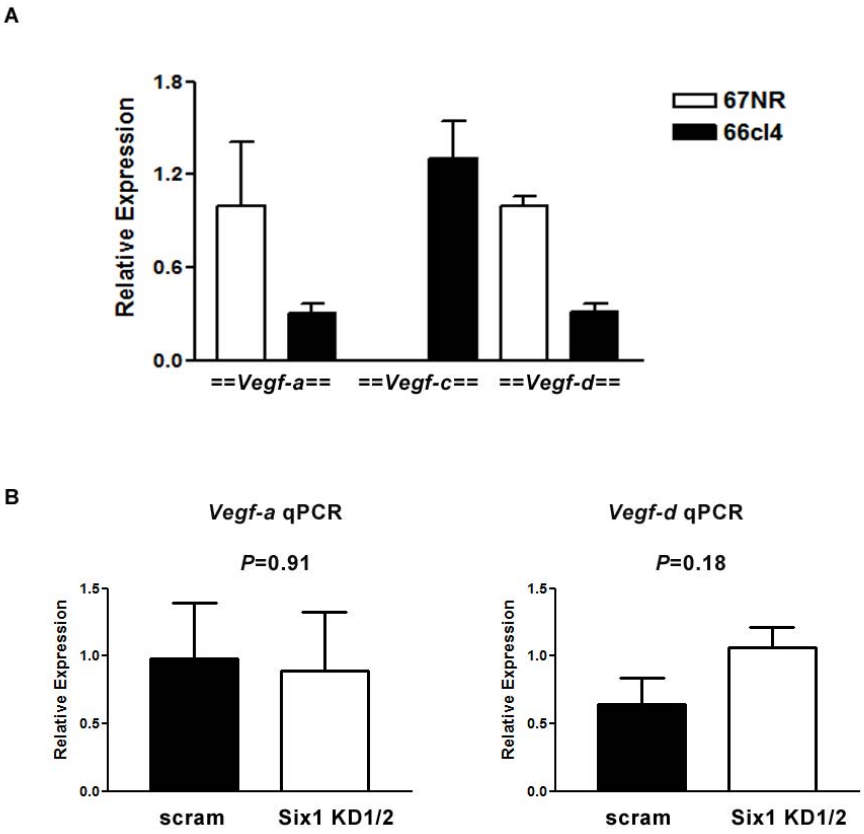


Supplementary Figure 6. (A) Representative picture showing SIX1 expression levels detected in MCF7-Ctrl and MCF7-SIX1 cells by immunocytochemistry (ICC). (B) Western blotting was performed to detect SIX1 expression in MCF7-SIX1 cells compared to endogenous SIX1 expression in numerous breast cancer cell lines.

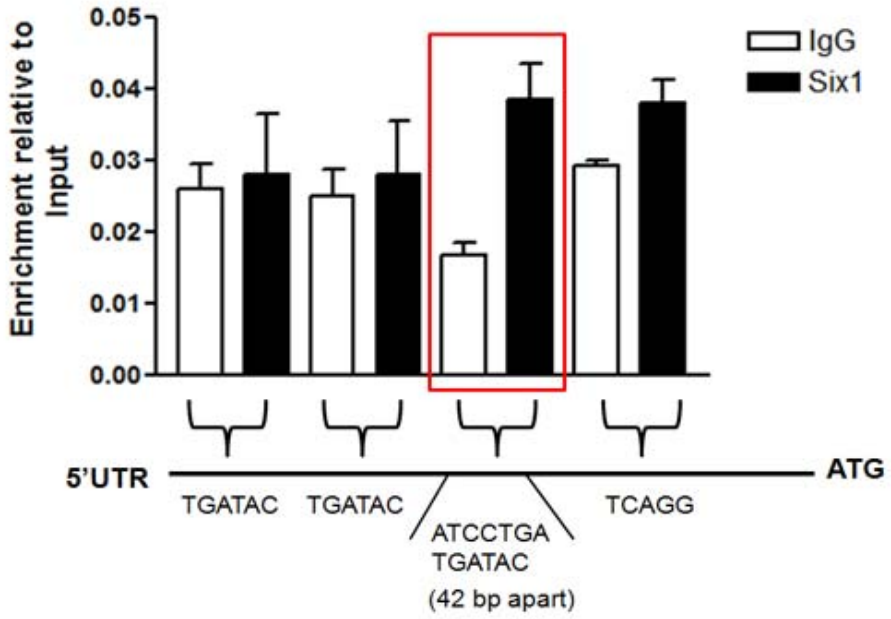


Supplementary Figure 7. Vegf-c expression is critical for metastasis in the 66cl4 mammary carcinoma model. (A) Detection of *Vegf-c* expression by real-time PCR in scramble control (scram) and two Vegf-c knockdown lines (Vegf-c KD1/2). Two different shRNAs (combined as Vegf-c KD1/2) were used to knockdown Vegf-c in

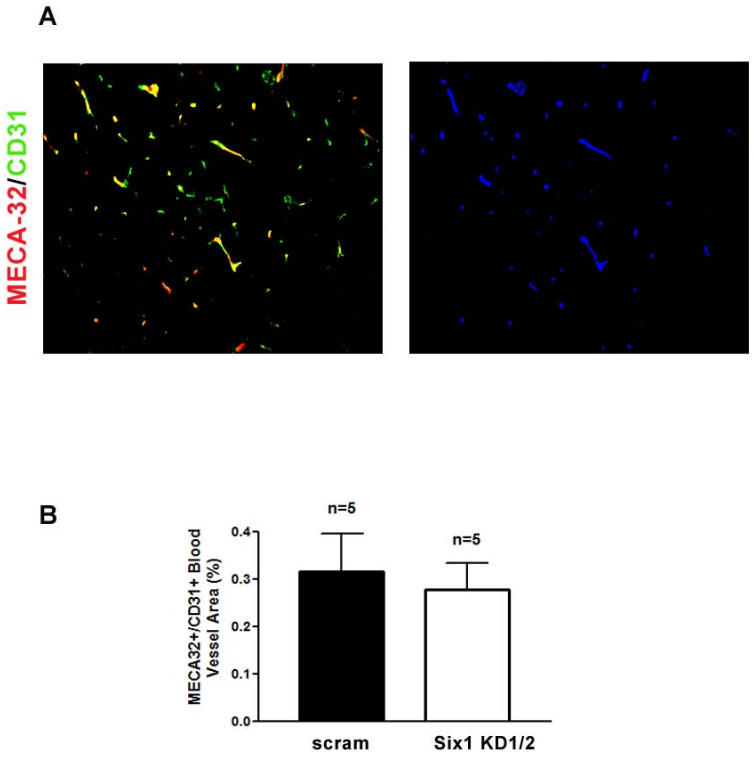
66cl4 cells, and pooled populations were labeled with luciferase for in vivo imagining.(B) Secreted Vegf-c was measured by ELISA in conditioned medium from 66cl4-scramble or 66cl4-Vegf-c KD cells. (C)Bioluminescent imaging of Balb/c mice 40 days after injection of 66cl4-scramble or 66cl4-Vegf-c KD cells into the 4^{th} mammary fat pad. The luminescence signal is represented by the overlaid false-color image and the intensity of the signal is indicated by the scale. The anesthetized mice were imaged with the IVIS200, 10 minutes after intraperitoneal injection of D-luciferin. Metastatic region is highlight by the yellow box. (D) Primary tumor growth in mice bearing 66cl4-scramble or Vegf-c KD tumors. Tumor size in the animals was measured using calipers according to the formula $V=1/2(W)^2(L)$. Error bars represent the standard error of the mean. Asterisks denote significant difference from control group. **, $P<0.01$; ***, $P<0.001$



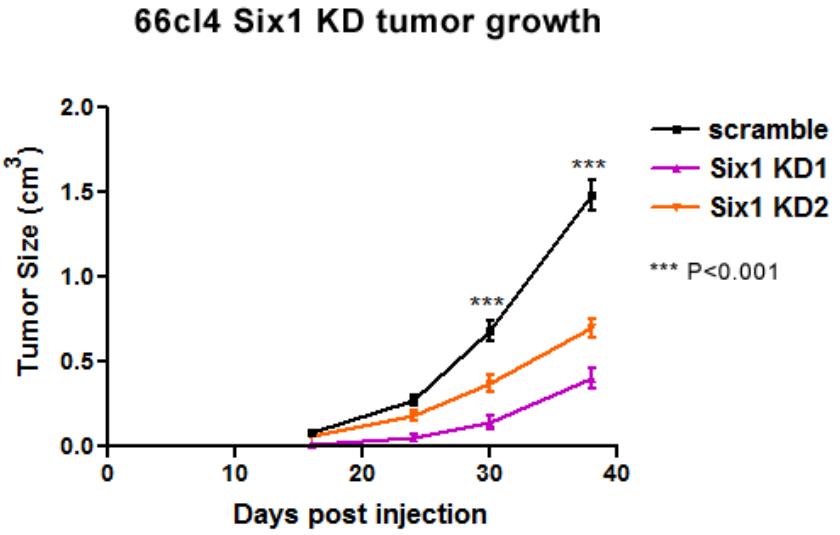
Supplementary Figure 8. (A) *Vegf-a*, *Vegf-c* and *Vegf-d* expression in 67NR and 66cl4 cells was detected by real-time PCR. (B)*Vegf-a* and *Vegf-d* expression in 66cl4-scramble and 66cl4-Six1 KD cells was detected by real-time PCR. Values represent the level of *Vegf-a* or *Vegf-d* transcript normalized to the level of the control *Cyclophilin b* mRNA. Error bars represent the standard error of the mean.



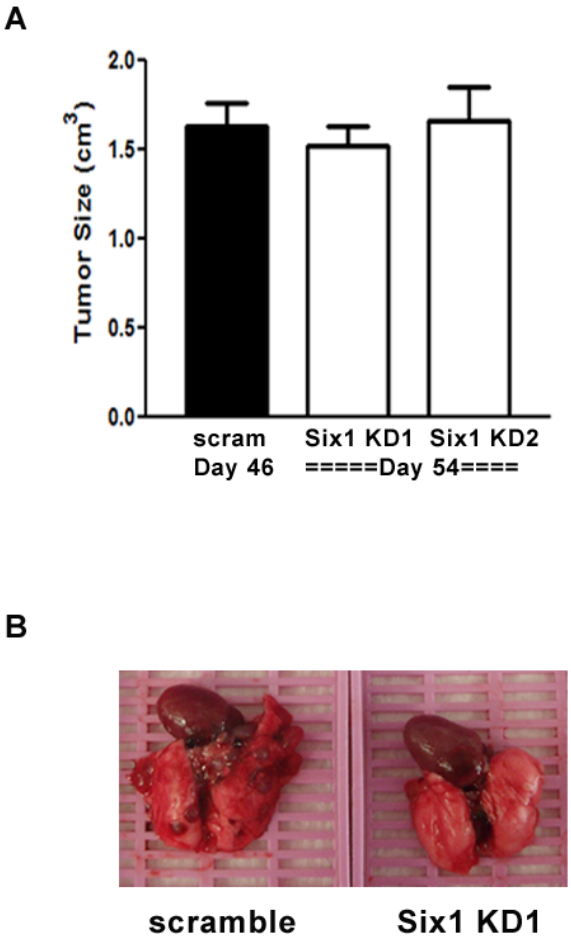
Supplementary Figure 9. Chromatin immunoprecipitation (ChIP) was performed to detect endogenous Six1 presence on the *Vegf-c* promoter in 66cl4 cells.



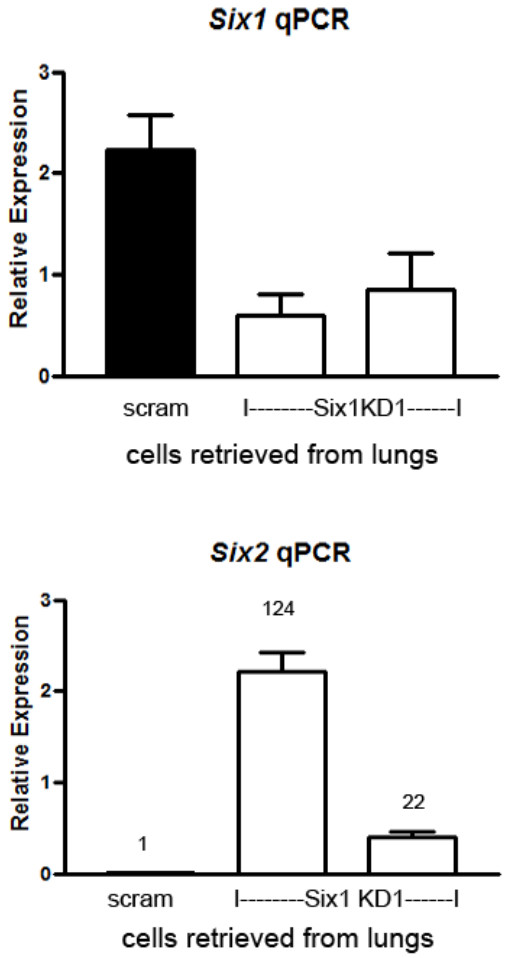
Supplementary Figure 10. (A) Representative picture showing blood vessel (angiogenesis) staining by MECA-32 and CD31 in a 66cl4 tumor (Left panel). Slidebook software was utilized to transform the signal from double stained regions (Right panel). (B) Quantification of blood vessels in 66cl4-scramble and 66cl4-Six1 KD1/2 tumors. Five individual tumor sections were stained in each group.



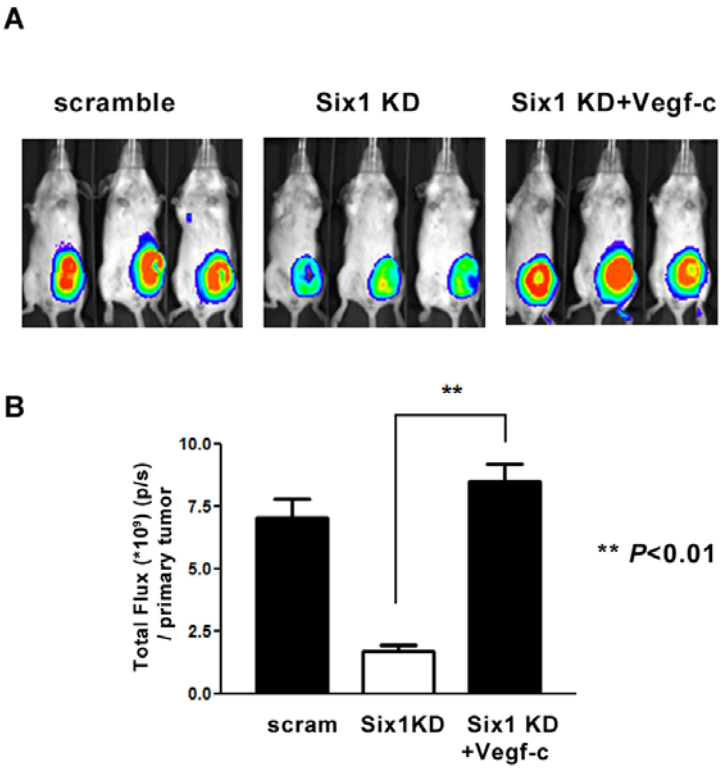
Supplementary Figure 11. Primary tumor growth in mice bearing 66cl4-scramble or 66cl4-Six1 KD tumors. Tumor size in the animals was measured using calipers and calculated using the formula $V=1/2(W)^2(L)$. Error bars represent the standard error of the mean. Asterisks denote significant difference from control group ***, $P<0.001$.



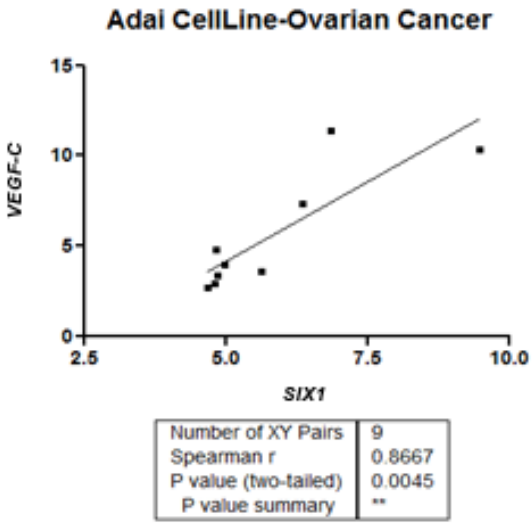
Supplementary Figure 12. (A) Tumor sizes of 66cl4-scramble at Day 46 and 66cl4-Six KD at Day 54. Tumor size in the animals was measured using calipers and calculated using the formula $V=1/2(W)^2(L)$. (B) Representative pictures showing the lungs from mice bearing 66cl4-scramble and 66cl4-Six1 KD1 tumors when mice had tumors of comparable sizes.



Supplementary Figure 13. (A) *Six1* expression was measured by real-time PCR in Six1 KD1 cells retrieved from the lungs of two individual 66cl4-Six1 KD1 tumor bearing mice, and 66cl4-scramble control cells were retrieved from the lung of one animal. (B) *Six2* expression was measured by real-time PCR in Six1 KD1 cells retrieved from the lungs of two individual 66cl4-Six1 KD1 tumor bearing mice, and from 66cl4-scramble control cells retrieved from the lung of one animal.



Supplementary Figure 14. (A) Representative bioluminescent imaging of Balb/c mice injected with 66cl4-scramble, 66cl4-Six1 KD, and 66cl4-Six1 KD+Vegf-c cells into the 4th mammary fat pad. (B) Quantitation of the luminescent signal in primary tumors from 66cl4-scramble, 66cl4-Six1 KD and 66cl4-Six1 KD+Vegf-c groups. Error bars represent the standard error of the mean. Asterisks denote significant difference **, P<0.01.

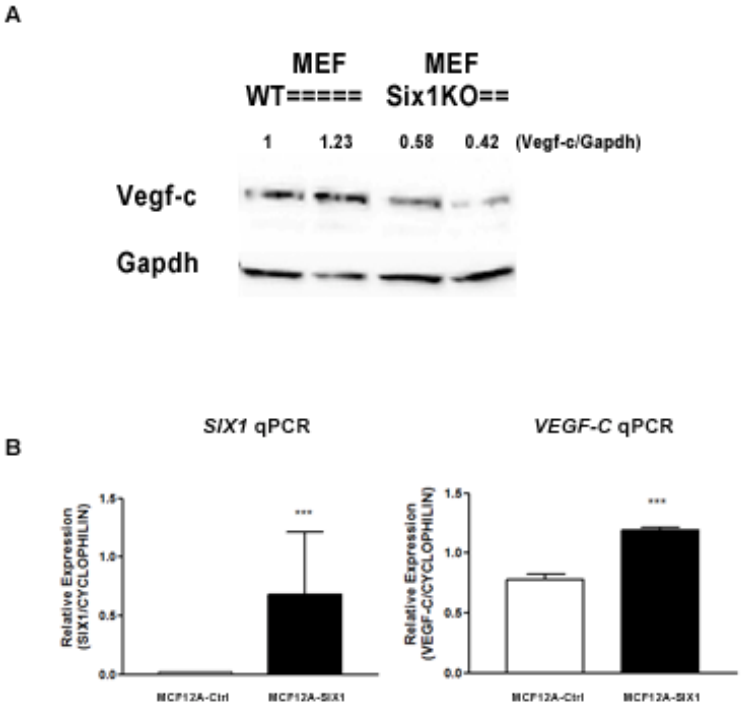


Supplementary Figure 15. *SIX1* and *VEGF-C* expression are correlated in ovarian cancer cell lines.

	SIX1 (+)	SIX1 (-)
VEGF-C (+)	42/110 (38%)	21/110 (19%)
VEGF-C (-)	24/110 (21%)	23/110 (20%)

Chi-square test, one-tailed
P=0.049 *

Supplementary Figure 16. Immunohistochemical staining for SIX1 and VEGF-C on 110 cases of invasive ductal carcinoma were quantified by a pathologist and their correlation in breast tumors was analyzed by Chi-square test.



Supplementary Figure 17. (A) Vegf-c levels were detected by Western blotting in WT and Six1 knock out mouse embryonic fibroblast cells (MEF). (B) *SIX1* and *VEGF-C* expression in MC12A-Ctrl and MC12A-SIX1 cells was detected by real-time PCR. Values represent the level of *SIX1* or *VEGF-C* transcript normalized to the level of the control *CYCLOPHILIN B*. Error bars represent the standard error of the mean. Asterisks denote significant difference from control group ***, $P<0.001$.

Supplementary Methods

Cell lines

MCF7-Ctrl and MCF7-SIX1 stable cell lines were generated as previously described [1]. MCF7-Ctrl and MCF-SIX1 cells used within the study are clonal isolates, and three clonal isolates from each line (MCF7-Ctrl and MCF7-SIX1) were used in our experiments. The results of all three clonal isolates in each group are combined to show the groups as MCF7-Ctrl and MCF7-SIX1. MCF7-Ctrl and MCF7-SIX1 cells were fluorescently tagged with Zsgreen to enhance our ability to follow metastasis. To generate VEGF-C knockdown in MCF7-SIX1 cells, shRNAs against *VEGF-C* were purchased from *Open Biosystems* and delivered retrovirally according to the manufacture's protocol. Cells containing the knockdown constructs were selected using puromycin (2.5ug/ml), and the two stable knockdown clones that most efficiently reduced VEGF-C levels were selected for subsequent studies. The 67NR and 66cl4 mammary carcinoma cell lines were generously provided by Fred Miller [2]. 66cl4 cells were tagged with luciferase for in vivo imaging of metastases. Vegf-c and Six1 were knocked down in 66cl4 cells using two different shRNAs from *Open Biosystems* that were lentivirally delivered according to the manufacturer's protocol. Stable knockdown cells were selected with puromycin. To restore Vegf-c expression, a Vegf-c expression vector (CMV-sport6) was purchased from *Open Biosystems* and

the full length cDNA was cloned into pcDNA3.1 (Invitrogen) and transfected into 66cl4 Six1 knockdown cells. Stable clones with Vegf-c restored were selected via treatment with neomycin (400ug/ml).

In vitro assays

TaqMan Analysis: TaqMan primers used against *VEGF-A*, *VEGF-C*, *VEGF-D*, and *Lyve-1* were purchased from *Applied Biosystems*. All analyses were performed using the Bio-Rad CFX-96 real time system machine and iScript cDNA synthesis kits (*Bio-Rad*) as well as TaqMan Fast universal PCR master mix (*Applied Biosystems*) according to the manufacturer's recommended protocols.

Luciferase reporter assay: Using a Fugene lipid-based transfection reagent (*Roche*), MCF7 Cells were transfected with a pGL3-VEGF-C promoter luciferase construct (0.5ug) (a kind gift from Kari Alitalo) along with increasing amounts of pcDNA3.1-SIX1 (up to 0.9ug) and a constant amount of pcDNA-EYA2 (0.2ug). A Renilla luciferase construct was also transfected as an internal control. After transfection, the cells were incubated for 48 hours and then harvested in passive lysis buffer (*Promega*). Firefly and Renilla luciferase activity were measured using the Dual Luciferase Reporter Assay System (*Promega*) on a Modulus Luminometer (*Turner BioSystems*). Luciferase activity was normalized to renilla activity, then the reading was analyzed by the following equation:

(SIX1/EYA+pVEGFC)-(SIX1/EYA+pGL3)/(pcDNA+pVEGFC)-(pcDNA+PGL3).

Enzyme-linked immunosorbent assay (ELISA): Cells were cultured under serum starvation for 48 hours and medium was collected and concentrated on an Amicon Ultra (*Millipore*) column. Equal amounts of protein were loaded and measured by ELISA plate (*ELISA Tech*) according to the manufacturer's recommended protocol. The VEGF-C antibody from *R&D systems* (BAF752 and MAB752) was used (information provided by *ELISA Tech*).

Chromatin immunoprecipitation (ChIP): MCF7 cells were transfected with pcDNA3.1-SIX1 and pcDNA3.1-EYA2. After cross-linking and sonication, Protein-DNA complexes were precipitated with a SIX1 specific antibody made as previously described [3] as well as a control Rabbit IgG antibody (*cell signaling*). Real-time PCR was utilized to amplify regions around predicted SIX1 binding sites, as well as one upstream region with no predicted SIX1 binding site (no ATCCTGA or TGATAC core sequence) on the *VEGF-C* promoter as a negative control. 66cl4 cells were used to determine if endogenous Six1 binds to the mouse *Vegf-c* promoter. ChIP procedures are as described above except that no transfections were performed in the 66cl4 cells. The y-axis (enrichment relative to input) is a the percentage of *VEGF-C* promoter IP'ed using the Six1 antibody or the percentage of the *VEGF-C* promoter IP'ed using a control IgG antibody normalized to the amount of *VEGF-C* promoter

region found in the input DNA.

Histology, immunohistochemistry and immunofluorescence

Tumors, lymph nodes and lungs from animals were fixed in 4% paraformaldehyde, embedded in paraffin, and cut into 5-μm sections. Sections were stained with H&E to confirm metastases as previously described [4]. Primary antibodies against VEGF-C (1:200, *ZYMED laboratories*) and Lyve-1 (1:200, *Angiobio*), MECA32 (1:50, *BD*), Podoplanin (1:800, *developmental studies hybridoma bank*), CD31 (1:50, *Abcam*) were used on tissue sections. Tumor arrays (*US Biomax*) were stained using an anti-Six1 antibody (1:100; *Atlas antibodies*) and an anti-VEGF-C antibody (1:200, *ZYMED laboratories*). For immunofluorescence staining, corresponding secondary antibodies were labeled with AlexaFluor488 or AlexaFluor594 (*Invitrogen*) and images were captured using Nikon Diaphot fluorescence microscope and analyzed by Slidebook software (*Intelligent Imaging Innovations, Inc., Denver, CO*); images were converted to TIFF files and processed by Photoshop.

ShRNA hairpin sequence used to knock down VEGF-C and Six1.

shRNA target	**Hairpin sequence**
VEGFC (Homo sapiens)	TGCTGTTGACAGTGAGCGCCGCGACAAACACCTTCTTTAATAGTGAAGC CACAGATGTATTAAAGAAGGTGTTTGTCGCGATGCCTACTGCCTCGGA

VEGFC (Homo sapiens)	TGCTGTTGACAGTGAGCGCGCACATTATAATACAGAGATCTAGTGAAGCC ACAGATGTAGATCTCTGTATTATAATGTGCTTGCCTACTGCCTCGGA
VEGFC (Mus musculus)	CCGGCCCAAGTCTGTGTTTATTGAACTCGAGTTCAATAAACACAGACTTG GGTTTTTG
VEGFC (Mus musculus)	CCGGCCCTAATTCATGTGGAGCCAACTCGAGTTGGCTCCACATGAATTAG GGTTTTTG
Six1 (Mus musculus)	CCGGAGCTCAAACTATTCTCTTCCACTCGAGTGGAAGAGAATAGTTTGA GCTTTTTTG
Six1 (Mus musculus)	CCGGCTCCAACAAGCAGAATCAACTCTCGAGAGTTGATTCTGCTTGTTG GAGTTTTTG

Color Codes: sense loop antisense

(Sequence obtained from *Open Biosystems*)

Primer sequence designed for the detection of predicted SIX1 binding sites on human *VEGF-C* promoter.

Detected region	Primer sequence
-3447~ -3340	5'CTGTTTGCCAAAGGAAGCAC3'(forward) 5'CCCCATCTCTGCAATCAATC3'(reverse)

-2076~ -1985	5’GTTGGAAATGCCTGTGGTTC3’(forward) 5’GCCTTTCTGGCTTAACTGGTC3’(reverse)
-1776~ -1694	5’AGCAGGAAGACAGCAAGAGAGTGA3’(forward) 5’GACTCGGTCCATGATCTAGCCTGT3’(reverse)
-1547~ -1463	5’GTGAGTCAACACTGTGATTTGGCTGC3’(forward) 5’GCATGCAGAATGGACAGTATGCCTTTG3’(reverse)
-673~ -576	5’TAGCATCCATCCCAACAGC3’(forward) 5’GGTCCCCTCTCTCCCTTG3’(reverse)
-4486~ -4483	5’TCAAGAGGCCTCAGTGATTTG3’(forward) 5’ TCGGGTGTACAACAACTTGG3’ (reverse)