eLIFE
elifesciences.org

# Figures and figure supplements

Extracellular matrix signatures of human mammary carcinoma identify novel metastasis promoters

**Alexandra Naba, et al.**

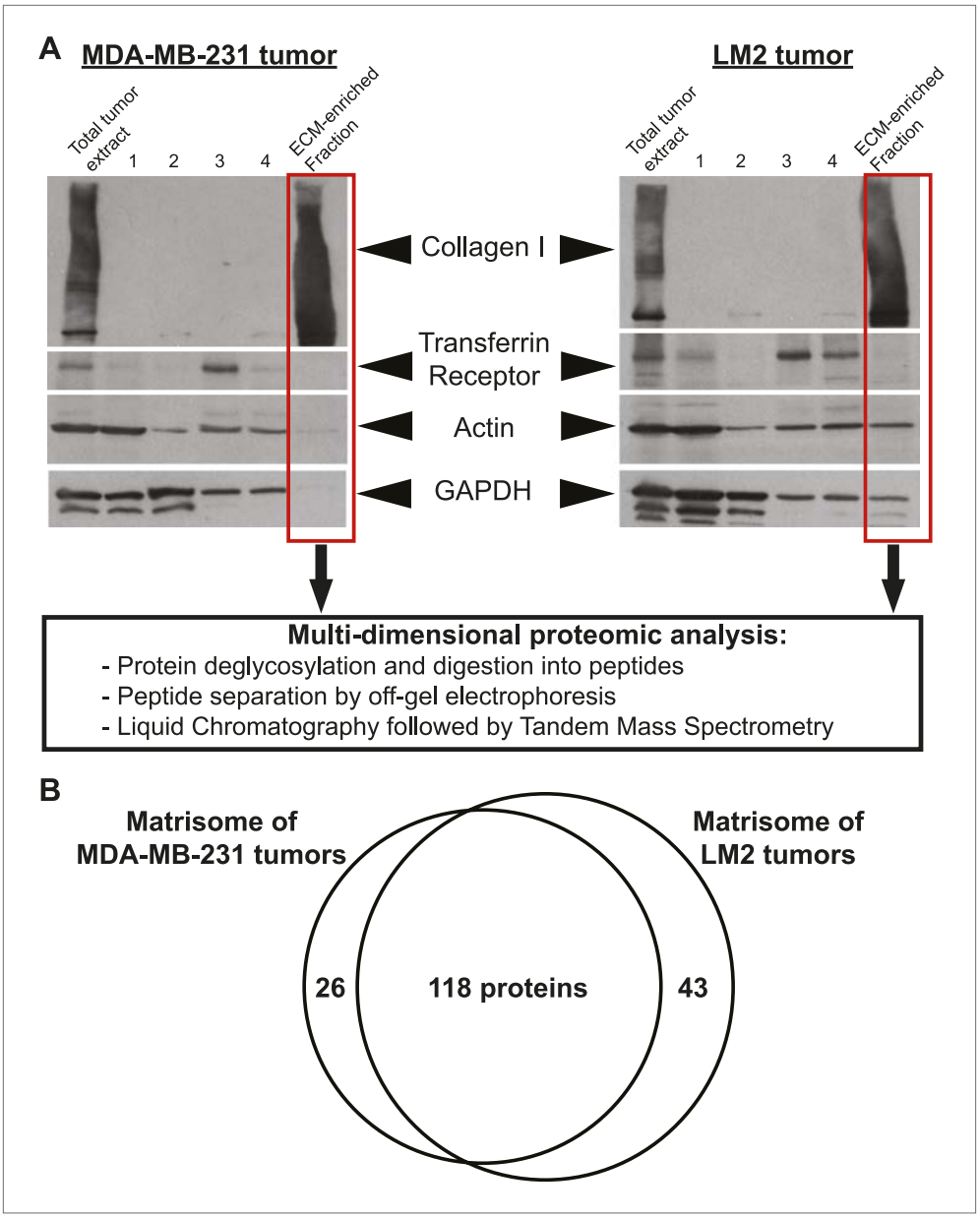



**Figure 1**. Enrichment of extracellular matrix proteins from human mammary tumor xenografts. (**A**) The sequential extraction of intracellular components was monitored by immunoblotting for GAPDH (cytosol), the transferrin receptor (plasma membrane), and actin (cytoskeleton). The remaining insoluble fraction was highly enriched for ECM proteins (collagen I panel) and largely depleted for intracellular components. The ECM-enriched fraction obtained is subsequently submitted to multidimensional proteomic analysis and the matrisome (ECM composition) of each tumor type is defined as the ensemble of proteins present in two replicate samples and with at least two peptides in one of the two replicates. (**B**) Venn diagram represents the comparison of the matrisomes of MDA-MB-231 and LM2 tumors. In addition to 118 ECM proteins detected in both tumor types, we identified 26 proteins specific to poorly metastatic (MDA-MB-231) tumors and 43 proteins characteristic of highly metastatic tumors (LM2).
DOI: 10.7554/eLife.01308.003

**Core Matrisome**

**ECM Glycoproteins**

| Entrez Gene Symbol | MW (kDa) | MDA-MB-231 | LM2 |
|---|---|---|---|
| Fn1 | 266.5 | 127 | 145 |
| Fbn1 | 332.9 | 87 | 107 |
| Tnc | 237.5 | 81 | 93 |
| TGFBI | 75.3 | 48 | 40 |
| Fgb | 55.4 | 44 | 46 |
| Postn | 90.9 | 43 | 26 |
| Fga | 61.8 | 31 | 34 |
| Fgg | 51.0 | 31 | 29 |
| Lamb2 | 203.7 | 28 | 19 |
| LAMA5/A3 | 412.3 | 22 | 34 |
| Nid1/Nid2 | 139.5 | 22 | 22 |
| THBS1 | 133.4 | 22 | 19 |
| Lamc1 | 182.9 | 20 | 15 |
| Emilin1 | 108.9 | 20 | 11 |
| PXDN | 167.9 | 17 | 30 |
| ECM1 | 65.2 | 15 | 27 |
| LTBP1 | 195.2 | 15 | 26 |
| ELN | 69.2 | 13 | 14 |
| Lamb1-1 | 209.7 | 13 | 14 |
| LAMB3 | 133.5 | 11 | 14 |
| Vtn | 55.6 | 10 | 12 |
| IGFBP7 | 30.2 | 8 | 12 |
| Fbln2 | 132.2 | 8 | 6 |
| Lama4 | 204.5 | 7 | 9 |
| Dpt | 24.6 | 6 | 2 |
| Efemp1 | 57.2 | 5 | 10 |
| LTBP4 | 182.6 | 4 | 10 |
| Vwa5a | 87.9 | 4 | 7 |
| THSD4 | 115.9 | 4 | 5 |
| IGFBP3 | 32.7 | 4 | 4 |
| LAMC2 | 134.9 | 4 | 4 |
| MFAP2 | 21.6 | 4 | 3 |
| LTBP2 | 204.2 | 3 | 18 |
| CYR61 | 44.2 | 3 | 6 |
| Mfap5 | 19.0 | 3 | 4 |
| GAS6 | 81.7 | 2 | 2 |
| Tnxb | 454.3 | 23 | - |
| Emilin2 | 118.6 | 7 | - |
| Lama2 | 351.9 | 5 | |
| Thbs2 | 133.8 | 3 | |
| FBLN5 | 52.5 | 3 | |
| Hmcn1 | 608.1 | 3 | |
| CTGF | 40.3 | - | 8 |
| Vwf | 322.6 | - | 7 |
| TINAGL1 | 53.8 | - | 6 |
| Vwa1 | 44.8 | - | 3 |
| Papln | 145.4 | - | 3 |
| EMID2 | 46.1 | - | 2 |
| SNED1 | 158.3 | | 8 |
| LTBP3 | 146.5 | | 6 |
| AGRN | 223.0 | | 6 |
| SRPX | 52.8 | | 5 |
| IGFBP4 | 29.1 | | 2 |
| EFEMP2 | 51.8 | | 2 |
| MFGE8 | 43.9 | | 2 |

**Collagens**

| Entrez Gene Symbol | MW (kDa) | MDA-MB-231 | LM2 |
|---|---|---|---|
| Col6a3 | 288.4 | 144 | 123 |
| Col1a1/2a1 | 139.1 | 129 | 150 |
| Col12a1 | 341.4 | 120 | 114 |
| Col3a1 | 140.2 | 104 | 104 |
| Col1a2 | 130.1 | 95 | 116 |
| Col5a2 | 146.0 | 43 | 37 |
| Col6a1 | 109.6 | 41 | 37 |
| Col4a2 | 168.5 | 37 | 33 |
| Col5a1/11a1-a2 | 184.4 | 35 | 28 |
| COL7A1 | 296.2 | 29 | 44 |
| Col6a2 | 111.5 | 29 | 23 |
| Col4a1/a3/a5 | 161.8 | 26 | 22 |
| Col5a3 | 172.7 | 23 | 20 |
| Col14a1 | 194.2 | 17 | 4 |
| COL18A1 | 154.4 | 11 | 22 |
| Col16a1 | 157.6 | 9 | 7 |
| Col10a1 | 66.8 | 7 | 17 |
| Col15a1 | 140.9 | 7 | 5 |
| COL13A1 | 70.4 | 6 | 8 |
| COL4A4 | 165.9 | 6 | 3 |
| COL19A1 | 116.0 | 8 | |
| Col6a6/Gm7455 | 248.1 | 7 | |
| COL17A1 | 150.8 | 3 | |
| Col28a1 | 119.7 | 2 | |
| COL22A1 | 162.2 | - | 4 |
| COL24A1 | 176.3 | | 7 |
| Col4a6 | 165.3 | | 6 |

**Proteoglycans**

| | | | |
|---|---|---|---|
| Hspg2 | 479.7 | 80 | 86 |
| Bgn | 42.1 | 12 | 9 |
| Dcn | 40.2 | 6 | 5 |
| Aspn | 43.0 | 8 | - |
| Lum | 38.7 | 6 | - |
| Prelp | 43.6 | 9 | - |
| OGN | 34.3 | 3 | |
| PRG4 | 152.3 | | 3 |

**Matrisome-associated**

**ECM-affiliated Proteins**

| Entrez Gene Symbol | MW (kDa) | MDA-MB-231 | LM2 |
|---|---|---|---|
| ANXA1 | 38.9 | 29 | 41 |
| ANXA6 | 76.2 | 28 | 25 |
| ANXA2 | 40.7 | 22 | 24 |
| ANXA3 | 36.5 | 15 | 15 |
| LGALS1 | 15.1 | 12 | 10 |
| ANXA5 | 36.0 | 11 | 10 |
| LMAN1 | 57.8 | 8 | 11 |
| ANXA11 | 54.7 | 8 | 11 |
| LGALS3 | 26.2 | 7 | 5 |
| ANXA7 | 53.0 | 5 | 7 |
| Lgals9 | 40.4 | 3 | 5 |
| LGALS8 | 40.4 | 2 | 5 |
| Cspg4 | 253.1 | 2 | 2 |
| C1qa | 26.3 | | 3 |
| Plxnb2 | 209.1 | | 3 |
| C1qc | 26.4 | | 2 |

**Secreted Factors**

| | | | |
|---|---|---|---|
| HCFC1 | 210.7 | 15 | 18 |
| S100A6 | 10.2 | 6 | 5 |
| S100A11 | 11.9 | 5 | 4 |
| S100A9 | 13.3 | 5 | 3 |
| S100A4 | 12.0 | 4 | 4 |
| CRLF3 | 50.3 | 3 | 5 |
| WNT16 | 25.0 | 2 | 3 |
| FLG2 | 249.4 | 3 | |
| S100A16 | 11.9 | 2 | |
| S100A2 | 11.3 | - | 2 |
| S100a8 | 10.4 | - | 1 |
| ANGPTL4 | 45.6 | | 6 |
| Il16 | 143.1 | | 4 |
| HCFC2 | 87.7 | | 3 |
| S100A10 | 11.3 | | 2 |

**ECM Regulators**

| Entrez Gene Symbol | MW (kDa) | MDA-MB-231 | LM2 |
|---|---|---|---|
| TGM2 | 78.5 | 45 | 48 |
| Plg | 93.5 | 24 | 34 |
| Itih1 | 101.5 | 21 | 5 |
| PLOD1 | 84.1 | 20 | 13 |
| SERPINH1 | 46.6 | 19 | 23 |
| Pzp | 167.2 | 18 | 11 |
| F2 | 71.7 | 14 | 26 |
| F13a1 | 83.7 | 12 | 8 |
| P4HA1 | 61.3 | 11 | 19 |
| PLOD2 | 87.8 | 11 | 7 |
| SERPINB9 | 43.0 | 11 | 7 |
| ADAMTSL1 | 199.4 | 9 | 17 |
| SERPIND1 | 57.2 | 9 | 10 |
| PLOD3 | 85.4 | 8 | 6 |
| MMP14 | 66.2 | 7 | 10 |
| LEPRE1 | 91.8 | 7 | 8 |
| Serpinf2 | 55.2 | 7 | 7 |
| Itih2 | 106.3 | 7 | 6 |
| Loxl1 | 67.2 | 7 | 4 |
| Kng1 | 54.3 | 7 | 3 |
| Serpina3k | 47.1 | 6 | 9 |
| P4HA2 | 61.3 | 6 | 6 |
| SERPINB6 | 43.0 | 6 | 4 |
| PLAU | 45.7 | 5 | 7 |
| Bmp1/Tll1/Tll2 | 113.9 | 4 | 9 |
| CTSD | 45.1 | 4 | 7 |
| Serping1 | 55.9 | 3 | 6 |
| Hrg | 60.1 | 3 | 4 |
| LOX | 47.6 | 2 | 6 |
| CSTB | 11.2 | 2 | 3 |
| NGLY1 | 75.5 | 4 | - |
| MMP1 | 54.2 | 3 | - |
| MMP19 | 57.6 | 2 | - |
| Timp3 | 24.9 | 2 | - |
| Itih3 | 99.7 | 8 | |
| PLAT | 65.1 | 2 | |
| F10 | 56.7 | - | 8 |
| Itih4 | 104.8 | - | 7 |
| Serpinc1 | 52.5 | - | 5 |
| EGLN1 | 46.9 | - | 3 |
| F13b | 78.4 | - | 3 |
| Serpine2 | 44.3 | - | 3 |
| CTSB | 38.8 | | 9 |
| LOXL2 | 88.8 | | 7 |
| CTSC | 52.6 | | 7 |
| ADAM10 | 86.2 | | 6 |
| TIMP1 | 23.9 | | 5 |
| Habp2 | 64.4 | | 5 |
| Serpina1b | 46.1 | | 5 |
| P4HTM | 63.7 | | 3 |
| ADAM9 | 93.1 | | 3 |
| HTRA1 | 52.2 | | 3 |
| CTSF | 54.0 | | 3 |
| LEPREL2 | 82.6 | | 3 |
| CST3 | 16.0 | | 2 |

Number of peptides: Low to High

No peptides detected

\- Detected in only one of the 2 replicate samples

**Figure 2**. Comparison of the matrisomes of MDA-MB-231 tumors and LM2 tumors identifies ECM proteins characteristic of poorly and highly metastatic tumors. Color code represents the number of unique peptides for each protein from poorly metastatic (MDA-MB-231) or highly metastatic (LM2) human mammary tumors. Values used to generate the figure were extracted from ***Figure 2—source data 1***, columns P and AA (number of peptides). Grayed

*Figure 2. Continued on next page*

*Figure 2. Continued*

cells indicate that no peptides were detected in either of the two replicate samples. A dash (–) indicates that the protein was detected in only one of the two replicate samples of a given tumor type or with only one peptide in both replicate samples.

DOI: 10.7554/eLife.01308.004

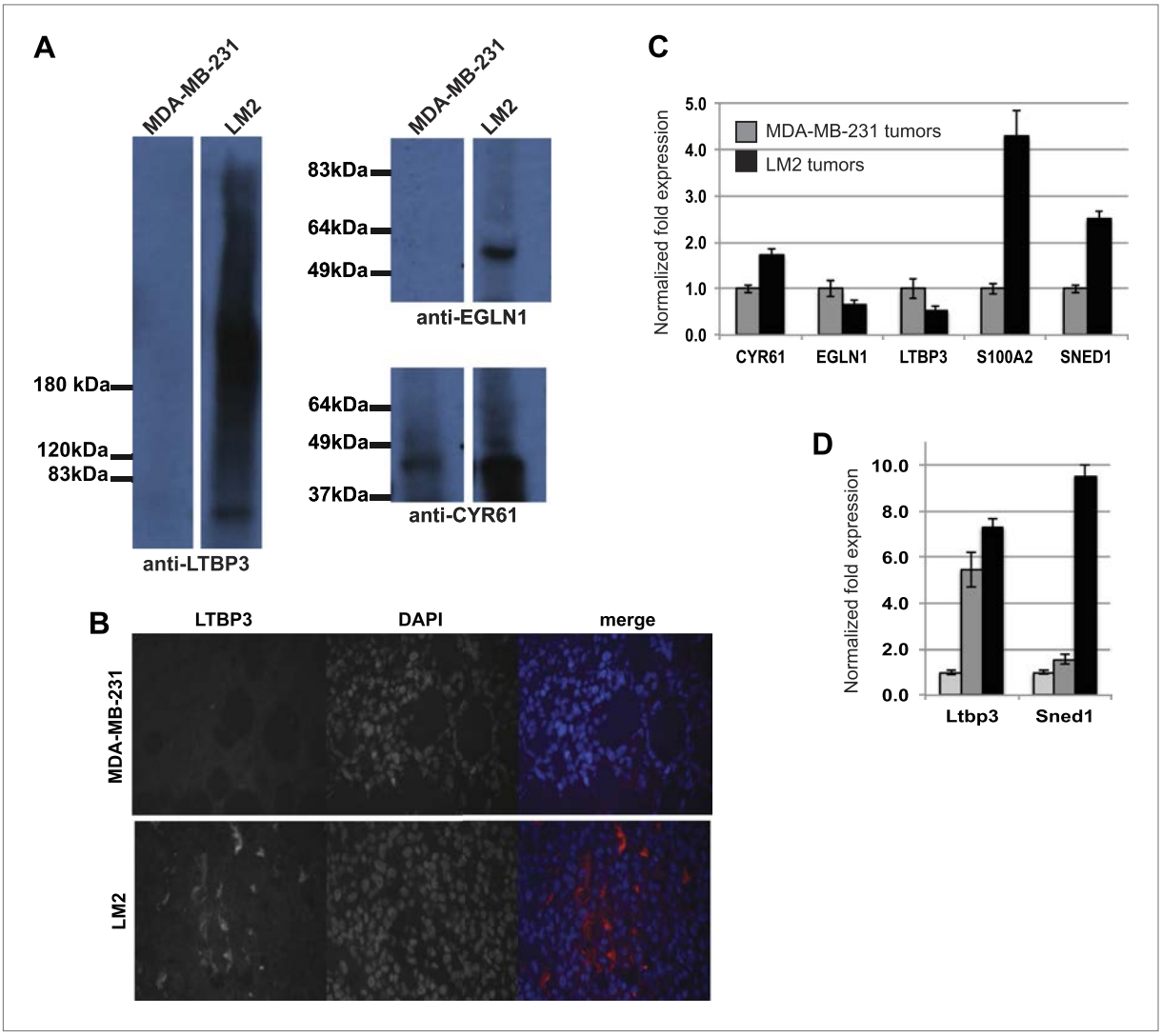



**Figure 2—figure supplement 1**. Validation of the differential expression of proteins identified by proteomics.

DOI: 10.7554/eLife.01308.006

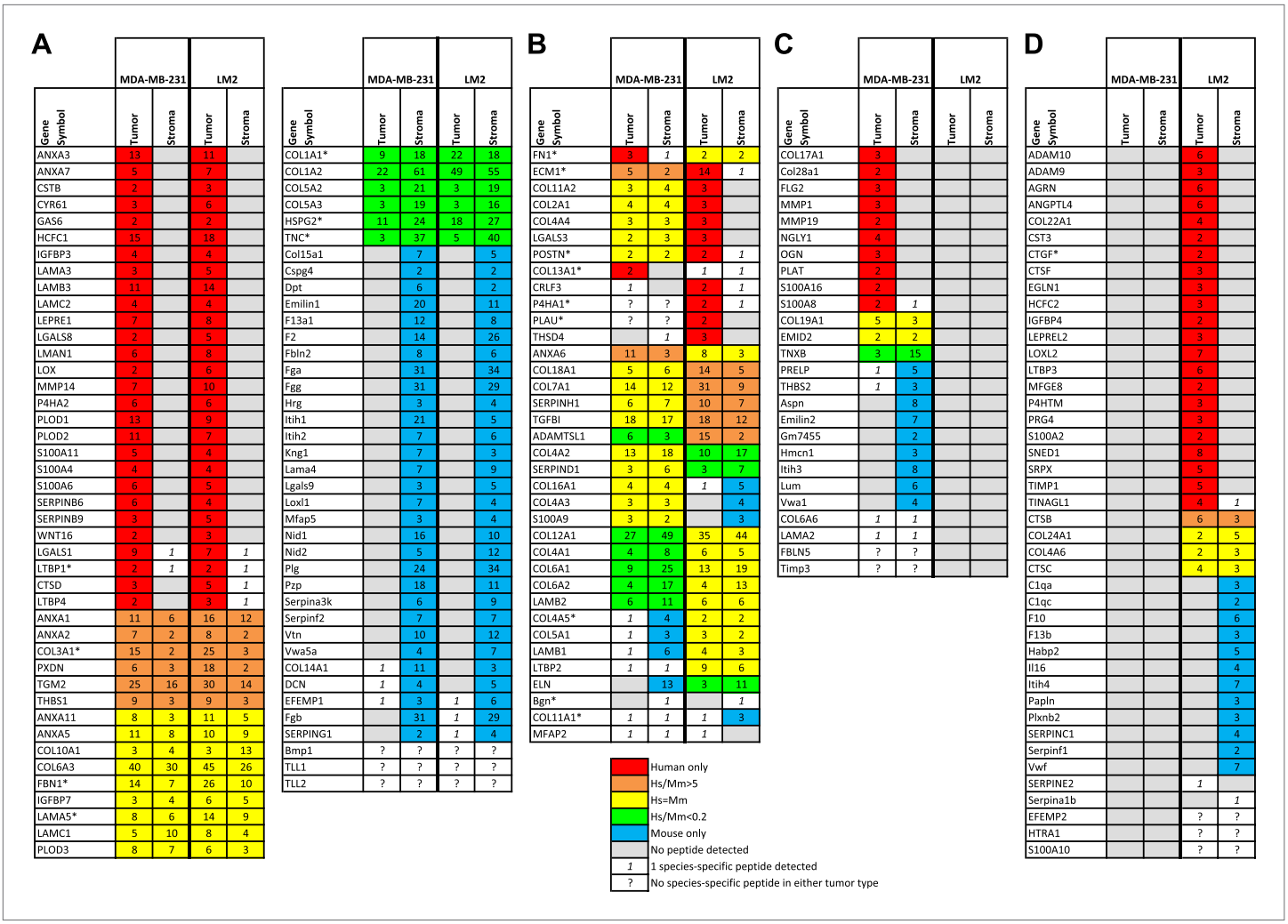



**Figure 3**. The tumor extracellular matrix is secreted by both tumor cells and stromal cells and differs with the tumor's metastatic potential. (**A**) Proteins expressed by both tumor types and by the same compartment in the two tumor types. (**B**) Proteins expressed by both tumor types but by different compartments. (**C**) Proteins secreted by MDA-MB-231 tumors and not by LM2 tumors. (**D**) Proteins secreted by LM2 tumors and not by MDA-MB-231 tumors. The number of peptides detected for each protein is indicated. For the proteins secreted by both the tumor cells and the stromal cells, the number of peptides listed corresponds to the number of human (tumor-derived)- or murine (stroma-derived)-specific peptides. For the proteins secreted by only one compartment, the number of peptides includes both species-specific and indistinguishable peptides. Proteins are sorted by tumor type and by their origins: tumor (red), stroma (blue), or both (yellow: similar abundance of the human and mouse proteins, orange: human form is at least five times more abundant than the mouse form, green: the mouse form is at least five times more abundant). To determine the relative contributions of the tumor and stromal cells to the secretion of ECM proteins, human-to-mouse peptide abundance ratios were calculated using the values indicated in column P and AB for the MDA-MB-231 and LM2 tumors respectively (***Figure 3—source data 1***). Proteins for which different isoforms have been detected are indicated with an asterisk (*) and, for simplicity, isoforms are combined here, their UniProt accession numbers can be found in ***Figure 3—source data 1***, column AP. In a few instances, the origin of the protein could not be determined due to the lack of species-specific peptides; these proteins are indicated with a question mark (?).

DOI: 10.7554/eLife.01308.007

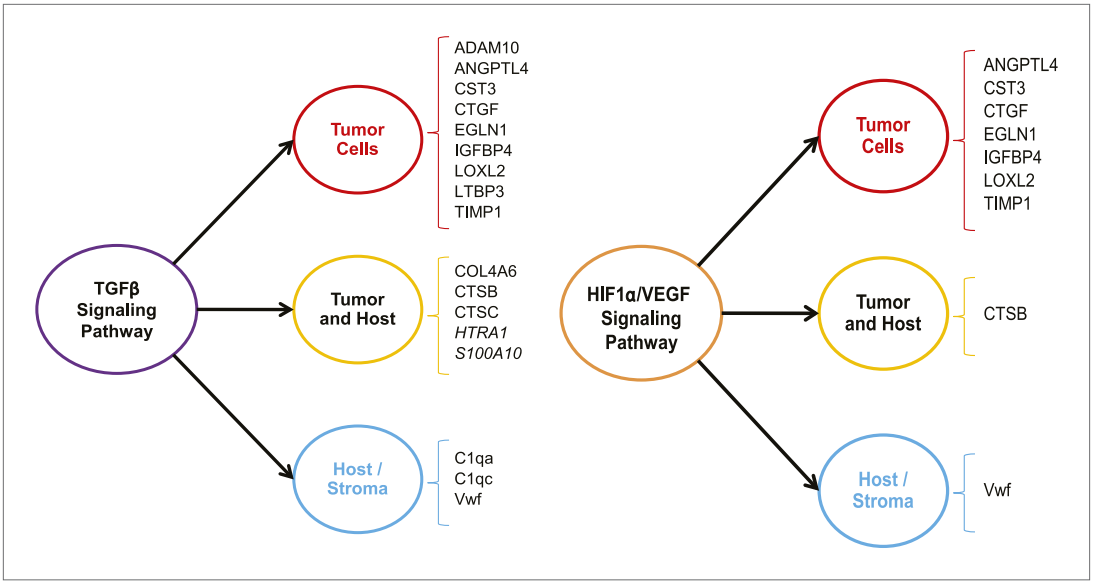


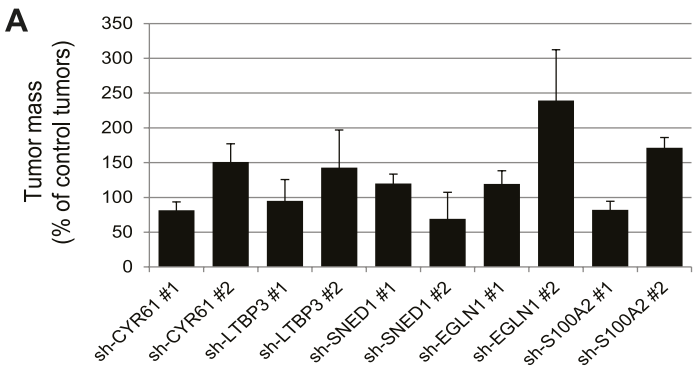

**Figure 4**. The TGFβ and the HIF1α/VEGF pathways are up-regulated in highly metastatic mammary tumors. The 43 ECM and ECM-associated proteins (tumor- and/or stroma-derived) unique to highly metastatic mammary tumors were uploaded to the Ingenuity Pathway Analysis software and queried for common upstream regulators. The analysis revealed an enrichment of TGFβ and HIF1α/VEGF targets within our ECM signature of highly metastatic mammary tumors.

DOI: 10.7554/eLife.01308.010

| Upstream Regulator | Molecule Type | Number of Proteins | Target molecules in dataset |
|---|---|---|---|
| TGFB1 | growth factor | 17 | ADAM10,ANGPTL4,C1QA,C1QC,COL4A6,CST3,CTGF,CTSB,CTSC,EGLN1,HTRA1,IGFBP4,LOXL2,LTBP3,S100A10,TIMP1,VWF |
| Vegf / HIF1a | group | 10 | ADAM10,ANGPTL4,CST3,CTGF,CTSB,EGLN1, IGFBP4,LOXL2, TIMP1,VWF |
| IFNG | cytokine | 10 | AGRN,ANGPTL4,C1QA,C1QC,CTGF,CTSB,CTSC,IGFBP4,S100A10,TIMP1 |
| TNF | cytokine | 10 | ANGPTL4,CTGF,CTSB,CTSC,CTSF,IGFBP4,IL16,PLXNB2,SERPINF1,TIMP1 |
| TP53 | transcription regulator | 9 | C1QC,CTGF,CTSB,CTSF,IGFBP4,PLXNB2,S100A2,SERPINC1,TINAGL1 |
| IL4 | cytokine | 7 | CTGF,CTSC,EGLN1,IGFBP4,IL16,S100A10,TIMP1 |
| IL1B | cytokine | 7 | ANGPTL4,CTSB,CTSF,IGFBP4,IL16,S100A10,TIMP1 |
| SMARCA4 | transcription regulator | 6 | CTGF,CTSB,IGFBP4,LOXL2,MFGE8,S100A2 |
| IL13 | cytokine | 6 | CTGF,CTSB,CTSC,HTRA1,SERPINF1,TIMP1 |
| CTNNB1 | transcription regulator | 6 | AGRN,COL4A6,CTGF,HTRA1,MFGE8,TIMP1 |
| JUN | transcription regulator | 5 | COL24A1,CTGF,IGFBP4,S100A10,TIMP1 |
| IL6 | cytokine | 5 | CST3,CTGF,CTSC,IGFBP4,TIMP1 |
| IL17A | cytokine | 4 | CTGF,IL16,PLXNB2,VWF |
| PRL | cytokine | 4 | CST3,CTSB,MFGE8,TIMP1 |
| FGF2 | growth factor | 4 | IGFBP4,S100A10,TIMP1,VWF |
| SEMA7A | transmembrane receptor | 3 | CTGF,CTSB,TIMP1 |
| Smad | complex | 3 | ANGPTL4,CTGF,EGLN1 |
| TGFB2 | growth factor | 3 | ANGPTL4,CTGF,TIMP1 |
| TGFA | growth factor | 3 | CTGF,IGFBP4,S100A10 |
| INHA | growth factor | 3 | COL4A6,CTGF,IGFBP4 |
| SMAD7 | transcription regulator | 3 | CTGF,LTBP3,TIMP1 |
| KLF2 | transcription regulator | 3 | CTGF,TIMP1,VWF |
| EDN1 | cytokine | 3 | CST3,CTGF,VWF |
| TP73 | transcription regulator | 3 | IGFBP4,S100A2,SERPINF1 |
| IGF1R | transmembrane receptor | 3 | HTRA1,IGFBP4,IL16 |
| STAT6 | transcription regulator | 3 | CTSB,S100A10,SERPINF1 |

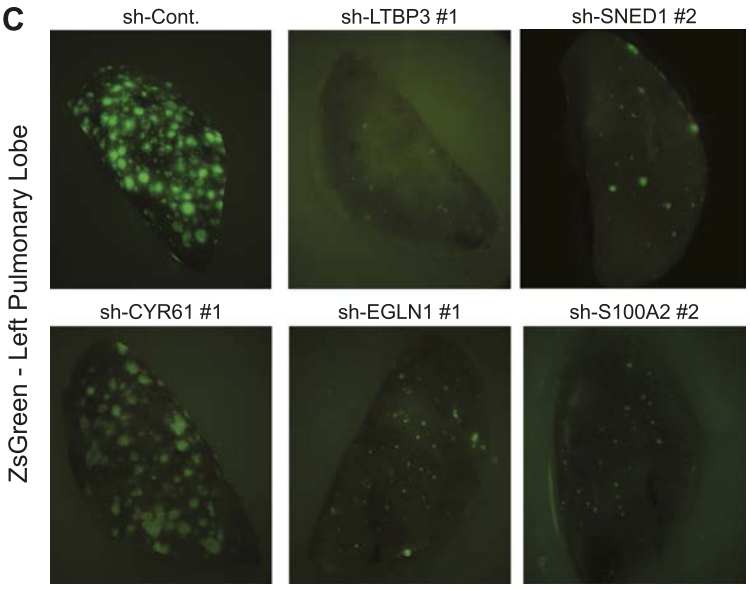
**Figure 4—figure supplement 1**. Ingenuity Pathway Analysis.

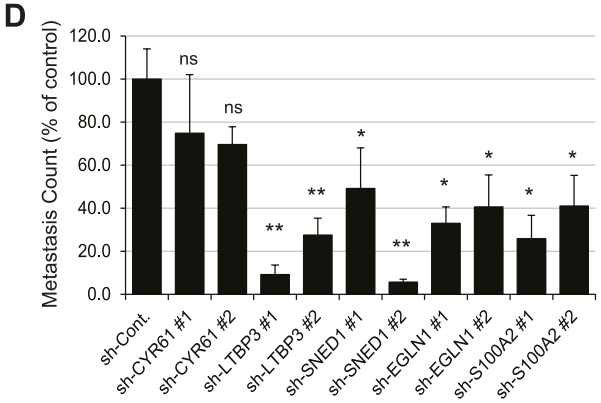
DOI: 10.7554/eLife.01308.011

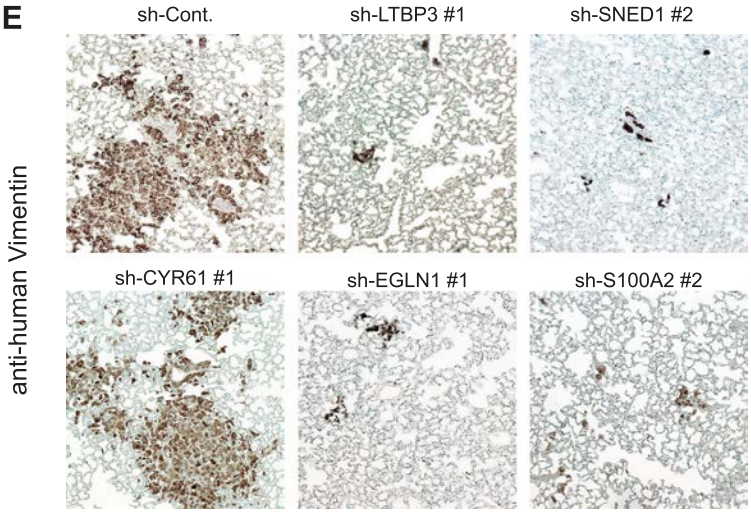



B

| | Mice with Visible Lung Metastases | Mice with Visible Liver Metastases | Mice with Visible Spleen Metastases |
|---|---|---|---|
| sh-Cont. | 10/13 | 6/13 | 5/13 |
| sh-CYR61 #1 | 3/5 | 4/5 | 2/5 |
| sh-CYR61 #2 | 3/3 | 0/3 | 0/3 |
| sh-LTBP3 #1 | 1/5 | 2/5 | 2/5 |
| sh-LTBP3 #2 | 2/10 | 0/10 | 0/10 |
| sh-SNED1 #1 | 4/8 | 0/8 | 0/8 |
| sh-SNED1 #2 | 3/8 | 0/8 | 0/8 |
| sh-EGLN1 #1 | 2/5 | 0/5 | 0/5 |
| sh-EGLN1 #2 | 1/4 | 0/4 | 0/4 |
| sh-S100A2 #1 | 1/4 | 2/4 | 2/4 |
| sh-S100A2 #2 | 2/5 | 1/5 | 0/5 |



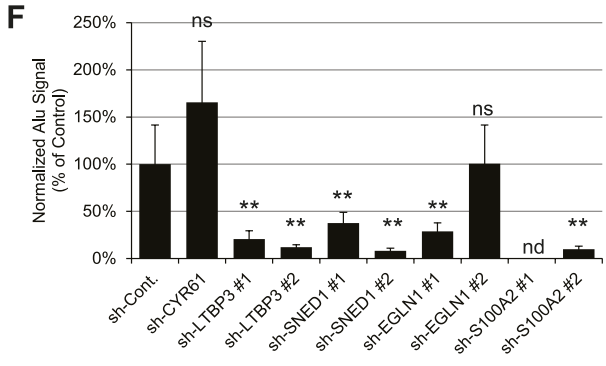




**Figure 5**. Tumor-cell-derived ECM proteins influence the metastatic dissemination of tumor cells to distant organs. Mice were injected orthotopically with control or knockdown LM2 cells. Tumors were allowed to grow for 7 ± 0.5 weeks. Number of mice per condition is indicated in ***Figure 5B***. (**A**) ECM protein knockdown does not inhibit primary tumor growth. At sacrifice, control and knockdown primary tumors were weighed. Bar chart represents the mass of knockdown tumors as a percentage of that of control tumors ± SEM. Student's *t* test was performed and none of the genes affected significantly and consistently primary tumor growth. (**B**) LM2 control tumors metastasize to the lungs, liver and spleen. The number of mice that presented with visible

*Figure 5. Continued on next page*

*Figure 5. Continued*

metastases in the indicated organs is indicated. (**C**) Representative pictures of whole left pulmonary lobe from LM2 (control or knockdown)-tumor-bearing mice with ZsGreen-positive metastatic foci. (**D**) Numbers of ZsGreen-positive metastatic foci in the left pulmonary lobe were counted. Data are presented as percentage of control ± SEM (Student's *t* test, *p<0.05, and **p<0.01). Number of animals per group is indicated in ***Figure 5B***. (**E**) Lung sections were stained with a human-specific anti-vimentin antibody to detect human tumor cells in the murine lung. (**F**) Alu PCR was performed on genomic DNA extracted from the lungs of control or knockdown tumor-bearing mice. Data are presented as normalized Alu signal as compared to murine actin signal and as percentage of control ± SEM (Student's *t* test, *p<0.05, **p<0.01, ns: not significant, and nd: not determined). Number of animals per group is indicated in ***Figure 5B***.

DOI: 10.7554/eLife.01308.012

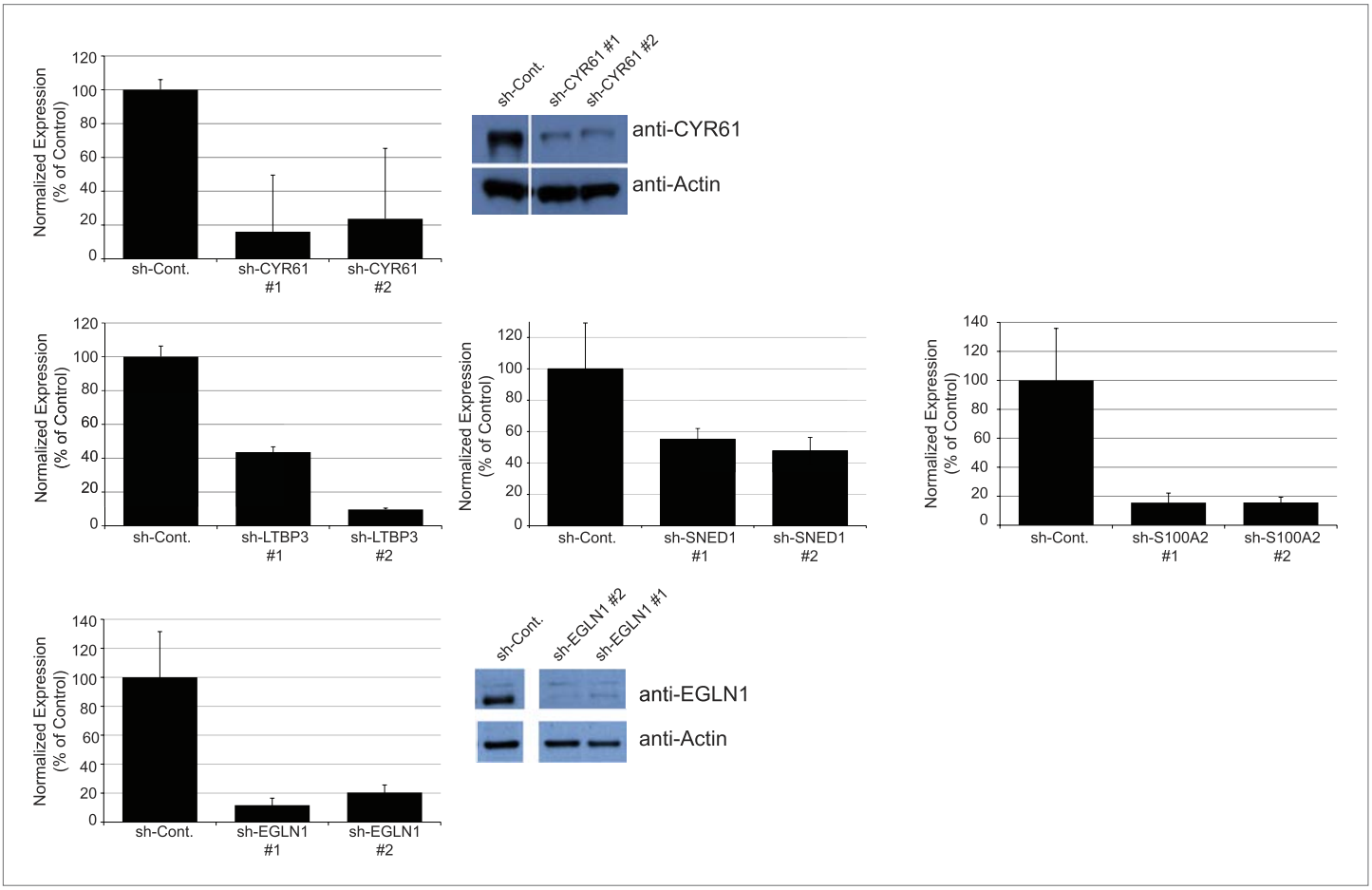



**Figure 5—figure supplement 1**. Establishment of stable LM2 knockdown cells.

DOI: 10.7554/eLife.01308.013

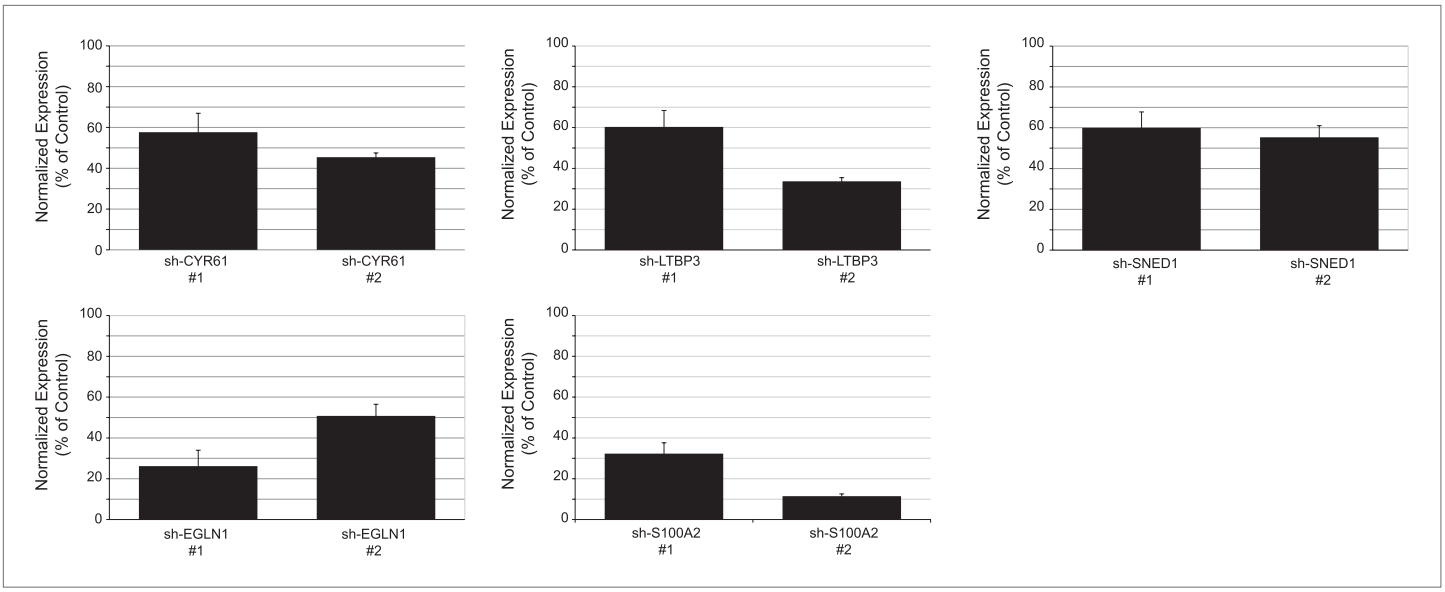



**Figure 5—figure supplement 2**. Persistence of gene expression knockdown in tumors.
DOI: 10.7554/eLife.01308.014

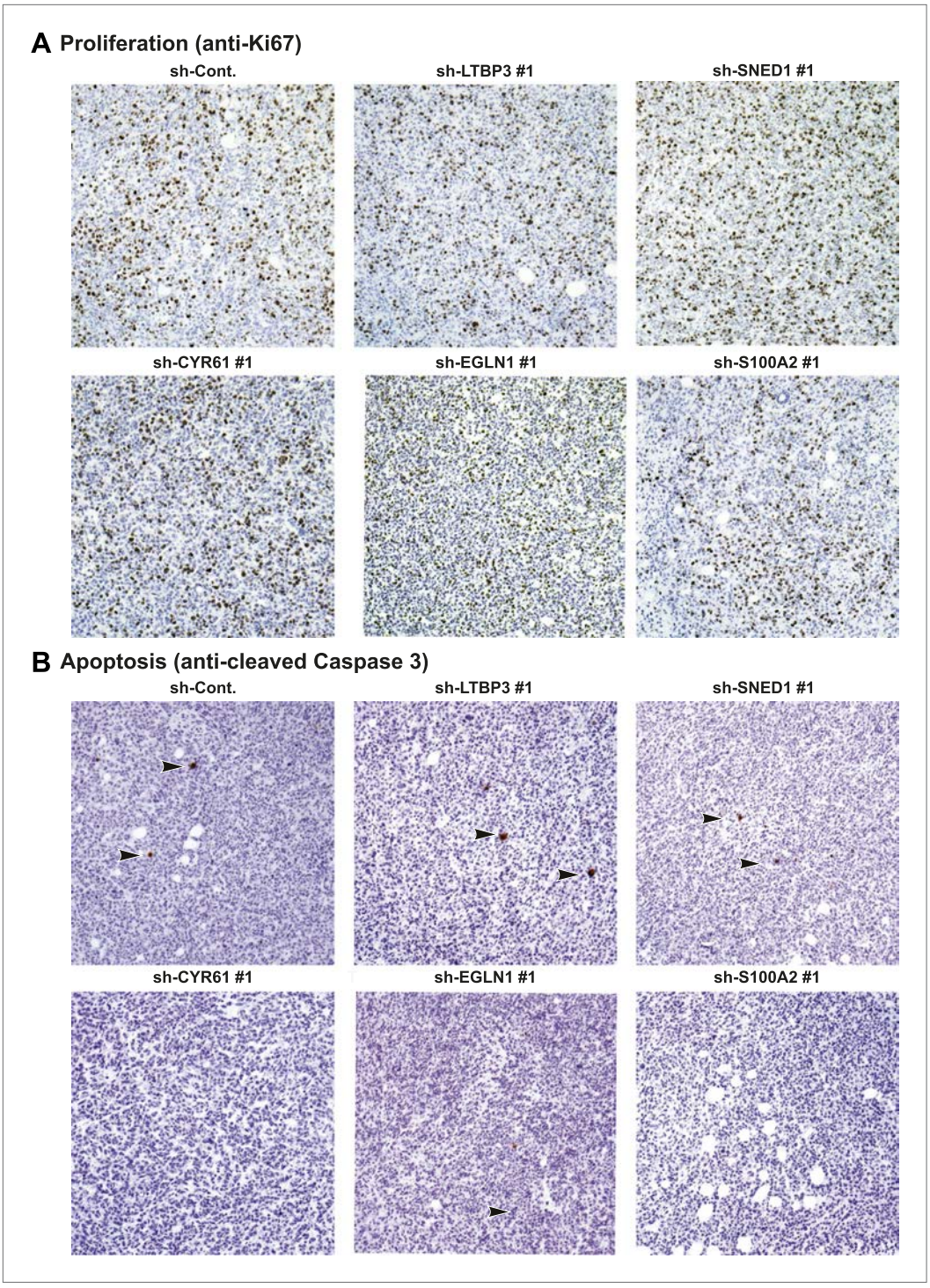



**Figure 5—figure supplement 3**. Tumor-cell-derived ECM proteins do not influence proliferation or apoptosis of primary mammary tumors.

DOI: 10.7554/eLife.01308.015

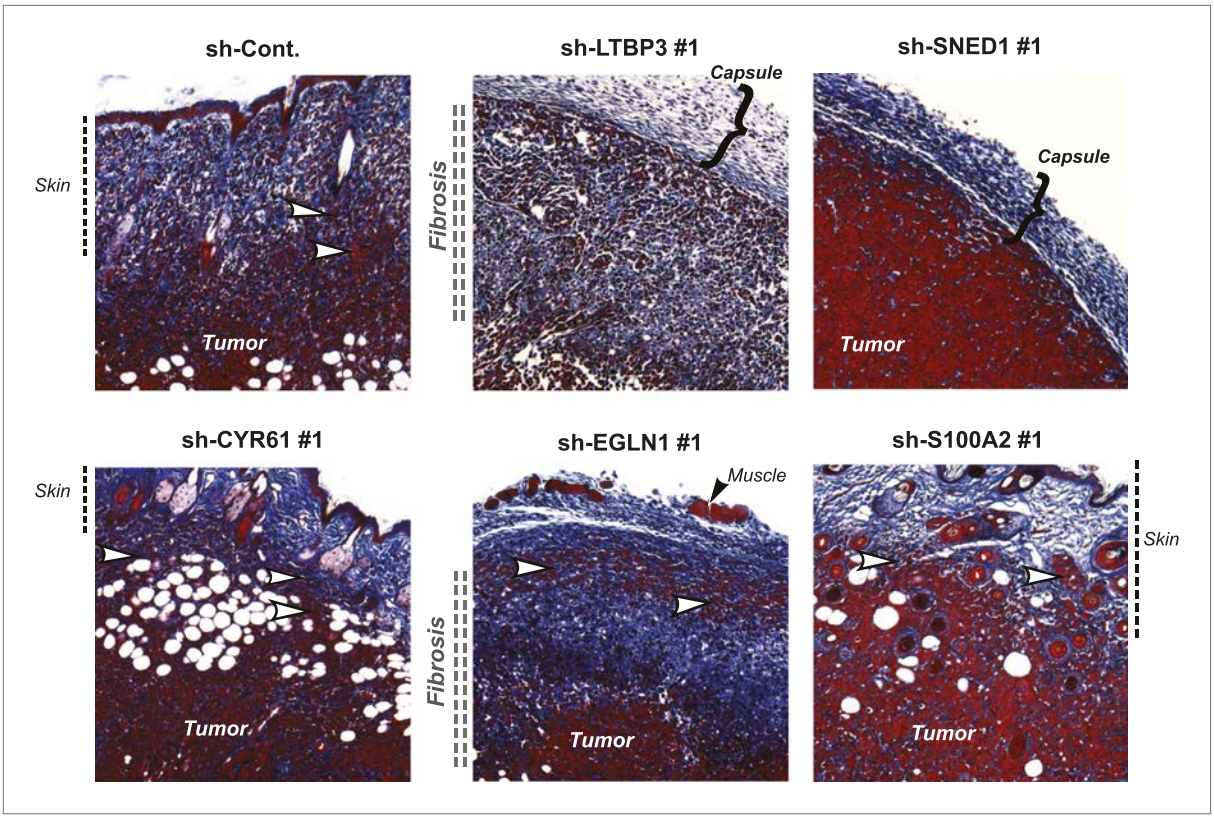



**Figure 6**. Tumor-cell-derived ECM proteins influence the invasiveness of primary mammary tumors. Primary tumor sections were stained with Masson's trichrome (blue: collagen fibers, red: cells) to evaluate fibrosis (indicated by vertical double-dashed line) and encapsulation (bracket) or invasiveness (white arrowhead) into the skin of the primary tumors. Note that tumors in which LTBP3 or SNED1 are knocked down are less invasive and more encapsulated than in the control tumors. CYR61, EGLN1, or S100A2 knockdown did not affect tumors' invasiveness.

DOI: 10.7554/eLife.01308.016

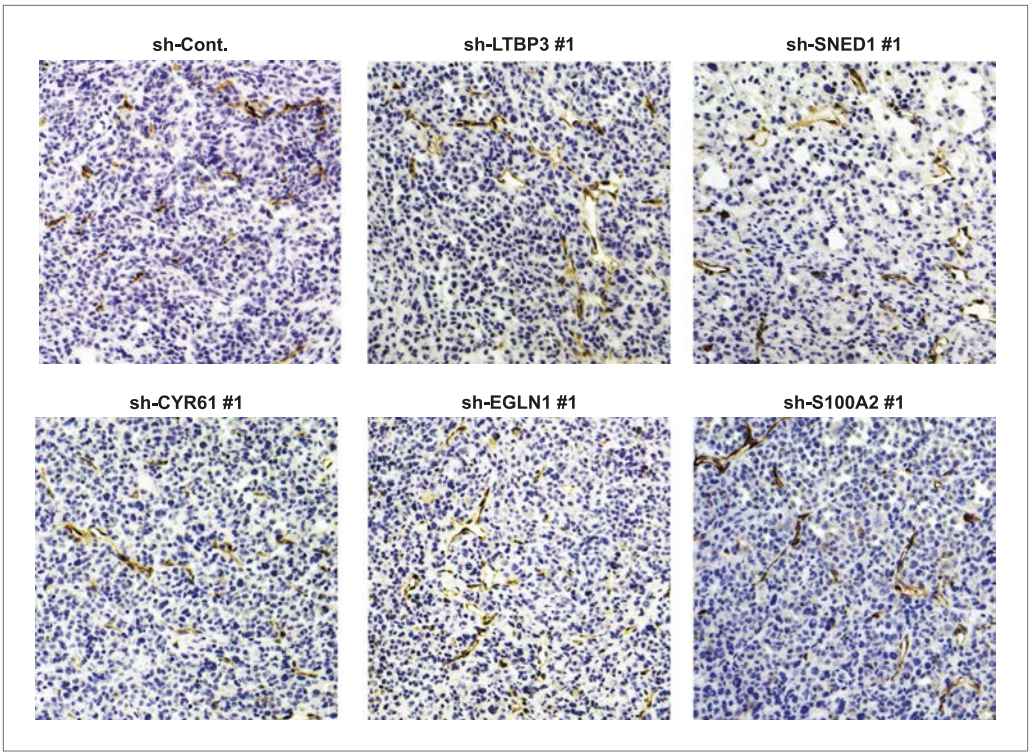



**Figure 6—figure supplement 1**. Tumor-cell-derived ECM proteins do not influence vascularization of primary mammary tumors.

DOI: 10.7554/eLife.01308.017

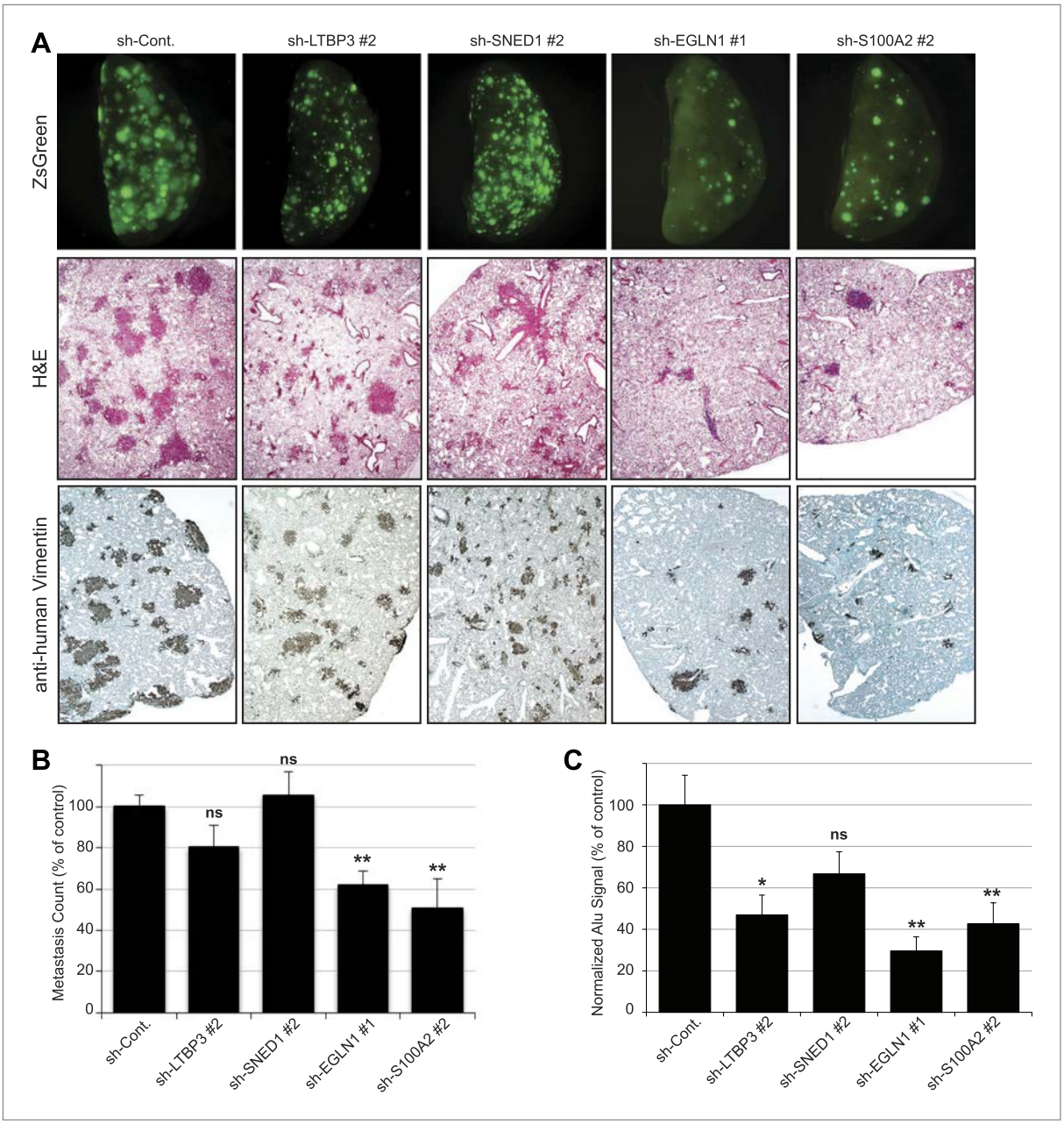



**Figure 7**. Tail vein metastasis assay. Control or knockdown cells were injected via the tail vein and the formation of lung metastases was evaluated. (**A**) Upper panel: representative pictures of the whole left pulmonary lobe with ZsGreen-positive metastatic foci from mice injected with LM2 (control or knockdown) cells (upper panel). Lung sections were stained with hematoxylin and eosin (H&E, middle panel) or with human-specific anti-vimentin antibody (lower panel) to visualize the metastatic foci. (**B**) Numbers of ZsGreen-positive metastatic foci in the left pulmonary lobe. Data are presented as percentage of control ± SEM (Student's *t* test, *p<0.05, **p<0.01, ns: not significant). Number of animals per group: sh-Cont.: 17 mice, sh-LTBP3: 10 mice, sh-SNED1: 10 mice, sh-EGLN1: 8 mice, sh-S100A2: 8 mice. (**C**) Alu PCR was performed to monitor the presence of human tumor cells in the murine lung. Data are presented as percentage of control ± SEM (Student's *t* test, *p<0.05, **p<0.01, ns: not significant). Number of animals per group: sh-Cont.: 17 mice, sh-LTBP3: 10 mice, sh-SNED1: 10 mice, sh-EGLN1: 8 mice, sh-S100A2: 8 mice.

DOI: 10.7554/eLife.01308.018

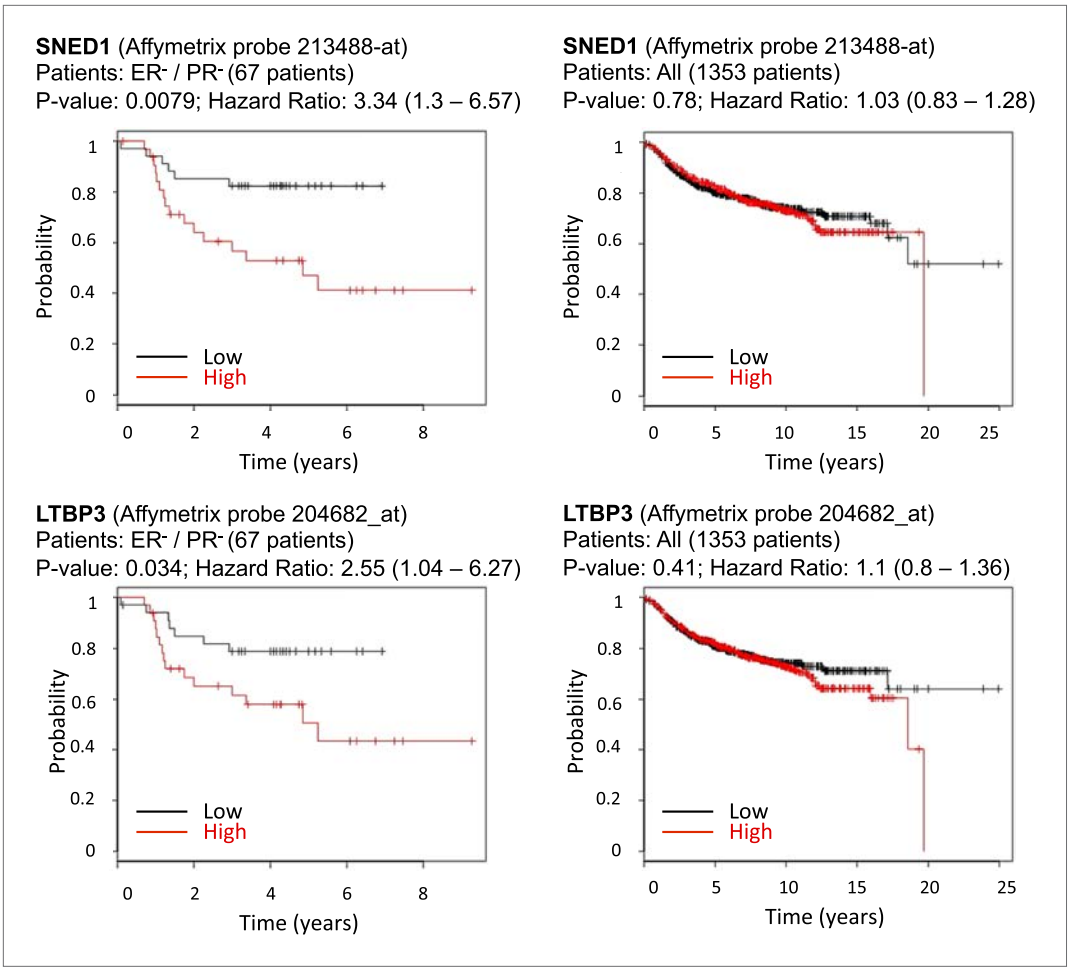



**Figure 8**. Correlation between SNED1 and LTBP3 expression and distant-metastasis-free survival in breast cancer patients. We tested the predictive value of the four ECM genes studied using the online assessment tool Kaplan–Meier Plotter. Whereas the expression of none of the genes studied correlated with distant metastasis-free survival for the entire population of breast cancer patients, SNED1 and LTBP3 expression inversely correlated with the survival of estrogen-receptor-negative and progesterone-receptor-negative ($ER^-/PR^-$) breast cancer patients.

DOI: 10.7554/eLife.01308.019