SUPPLEMENTAL INFORMATION

METHODS

Antibodies

Rat anti-CD31 (T-2001, Bachem). Rabbit anti-MEK1/2 (#8727), rabbit anti-MEK1/2 phospho Ser217/221 (#9154) and rabbit anti-p130Cas phospho Y249 (#4014, Cell Signaling). Mouse anti-FAK clone 4.47 (#05-537) and mouse anti-SRC (#05-184, Upstate/Millipore). Rabbit anti-FAK phospho Y397 (#44-624), rabbit anti-PXN phospho Y118 (#44-722) and rabbit anti-SRC phospho Y418 (#44-660, Biosource). Mouse anti-PXN (P-13520) and mouse anti-p130Cas (P-27820, Transduction). Rabbit anti-Ki67 (NCL-Ki67p, Novocastra).

Cell culture

4T1 (ATCC-CRL-2539), MCF7 (ATCC-HB-22), HCC70 and Hs578t cells (originally purchased from ATCC and described in Hollestelle et al. (22)) were cultured in RPMI (GIBCO, Life Technologies) supplemented with 10% FBS (PAA Laboratories) and 100 International Units/mL penicillin and 100 μg/mL streptomycin (Invitrogen).

Proliferation assay

Cells were plated onto micro-clear 96-well plates (Greiner Bio-one). After 5 days of incubation, the cells were stained with Hoechst 33342 and fixed with TCA (Trichloroacetic acid) allowing both a nuclear counting and/or Sulforodamine B (SRB) readout. Whole wells were imaged using epi-fluorescence and the number of nuclei was determined using a custom-made ImagePro macro. Plates were further processed for SRB staining as described earlier (Zhang et al (1)). SRB data showed a complete overlap with the nuclear count so the latter measure is used in all figures.

Stable shRNA mediated gene knockdown in 4T1

First, 4T1 cells were transduced with a lentiviral GFP construct. After that, cells were transduced with Sigma MISSION shRNA constructs coding a non-targeting control sequence and 2 independent sequences for SRPK1 (shSRPK1-A: CCG GCC TGA TAG GAT CTG GCT ACA ACT CGA GTT GTA GCC AGA TCC TAT CAG GTT TTT, shSRPK1-B: CCG GGC GCC AGA GAT TAA TTG CAA TCT CGA GAT TGC AAT TAA TCT CTG GCG CTT TT). After puromycin selection, cells were FACS sorted for homogenous GFP expression (in collaboration with Dr. Hoeben, LUMC).

Experimental metastasis assay

For the experimental lung metastasis assay, 2×10^5 MDA-MB-417.5 cells were injected into the lateral tail veins of 5- to 7-week-old female *RAG2*$^{-/-}$//*IL2RG*$^{-/-}$ mice. During the three to four weeks following injection, animals were monitored for tumor burden using bioluminescence imaging every week. After approximately four weeks, the mice were euthanized, and the lungs were removed and subjected to bioluminescent imaging, Indian ink injection (for counting of surface metastases) and histologic examination.

Orthotopic mouse breast cancer model

Experiment with 4T1 was done as previously described (2). In short, female Rag2$^{-/-}$//common-$\gamma^{-/-}$ mice were orthotopically injected with 100,000 4T1-GFP cells in the fourth mammary gland (10-11 animals per group). Tumor growth was monitored over a 3-week period and then mice were sacrificed and tumor and lungs were isolated. Lungs were injected with Indian ink to count surface metastases. The small lobe was processed for immunohistological analysis with H&E staining.

Gene expression profiling human breast cancer patients

A total of 344 lymph node negative (LNN) breast cancer (BC) patients were used for associating SRPK1 with the molecular subtypes. Details on data processing are available in the main methods. The dataset is available in National Center for Biotechnology Information/Gene Expression Omnibus (GEO), entries GSE2034 and GSE5327 and subtypes were assigned using a subtype clustering model present in the "genefu" R package (Desmedt et al 2008 Clinical Cancer Res 14:5158-5165). These data were used for supplemental figures 3,4 and 7. Publicly available breast cancer data were used for supplemental figure 5. These samples are derived from 3 different cohorts, GSE2990 (n=189), GSE7390 (n=198) and GSE11121 (n=200). Raw data were downloaded and processed using the R package "fRMA" (McCall MN, Bolstad BM and Irizarry RA (2010). "Frozen robust multiarray analysis (fRMA)." Biostatistics, 11(2), pp. 242–253). Batch effects were removed using ComBat (Johnson, WE, Rabinovic, A, and Li, C (2007). Adjusting batch effects in microarray expression data using Empirical Bayes methods. Biostatistics 8(1):118-127). All the patients in these cohort are LNN, additional patient characteristics can be obtained via GEO: http://www.ncbi.nlm.nih.gov/geo/.

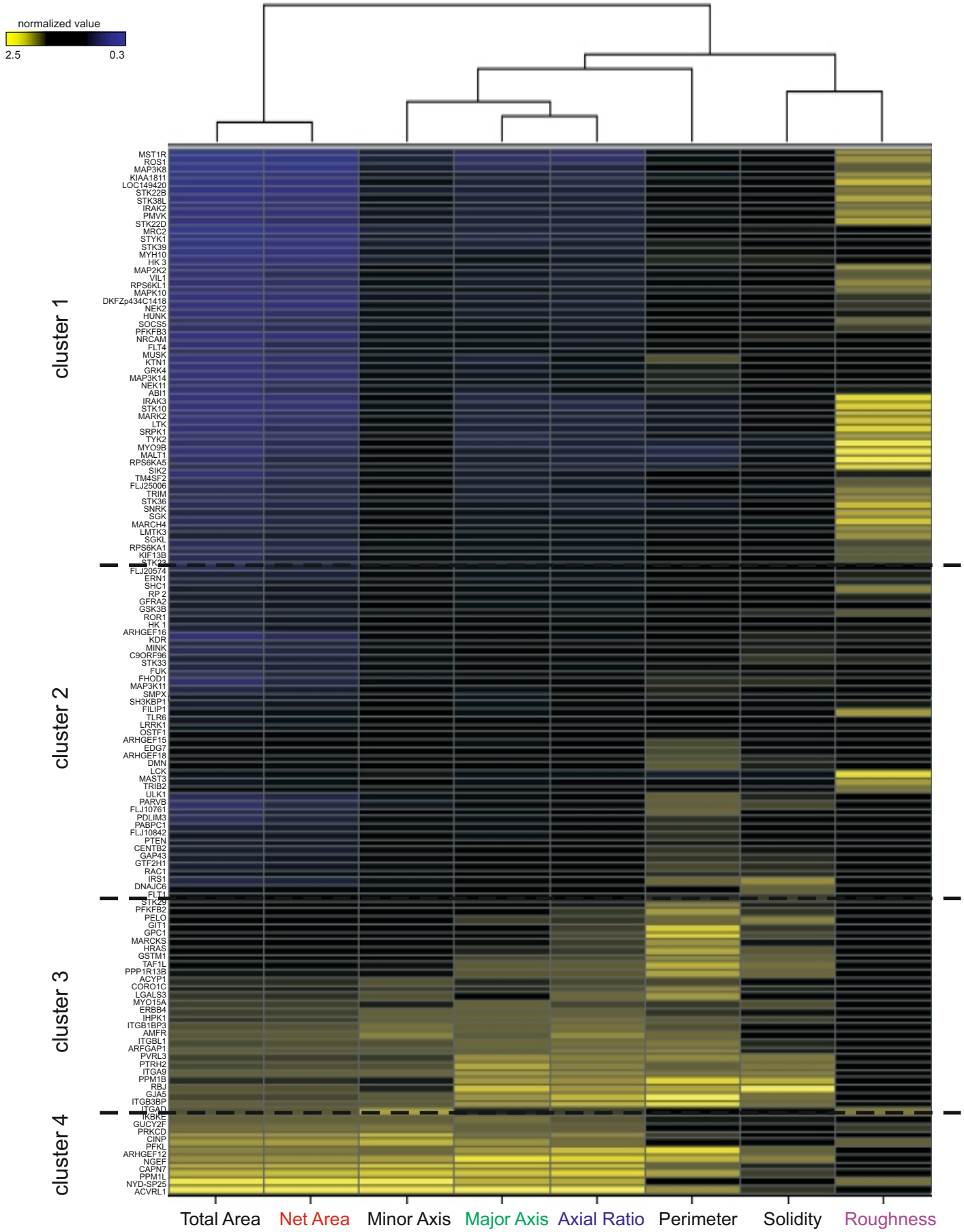
Declaration of ethical approval

Mouse experiments and housing were performed according to the Dutch guidelines for the care and use of laboratory animals (UL-DEC-14027).

REFERENCES

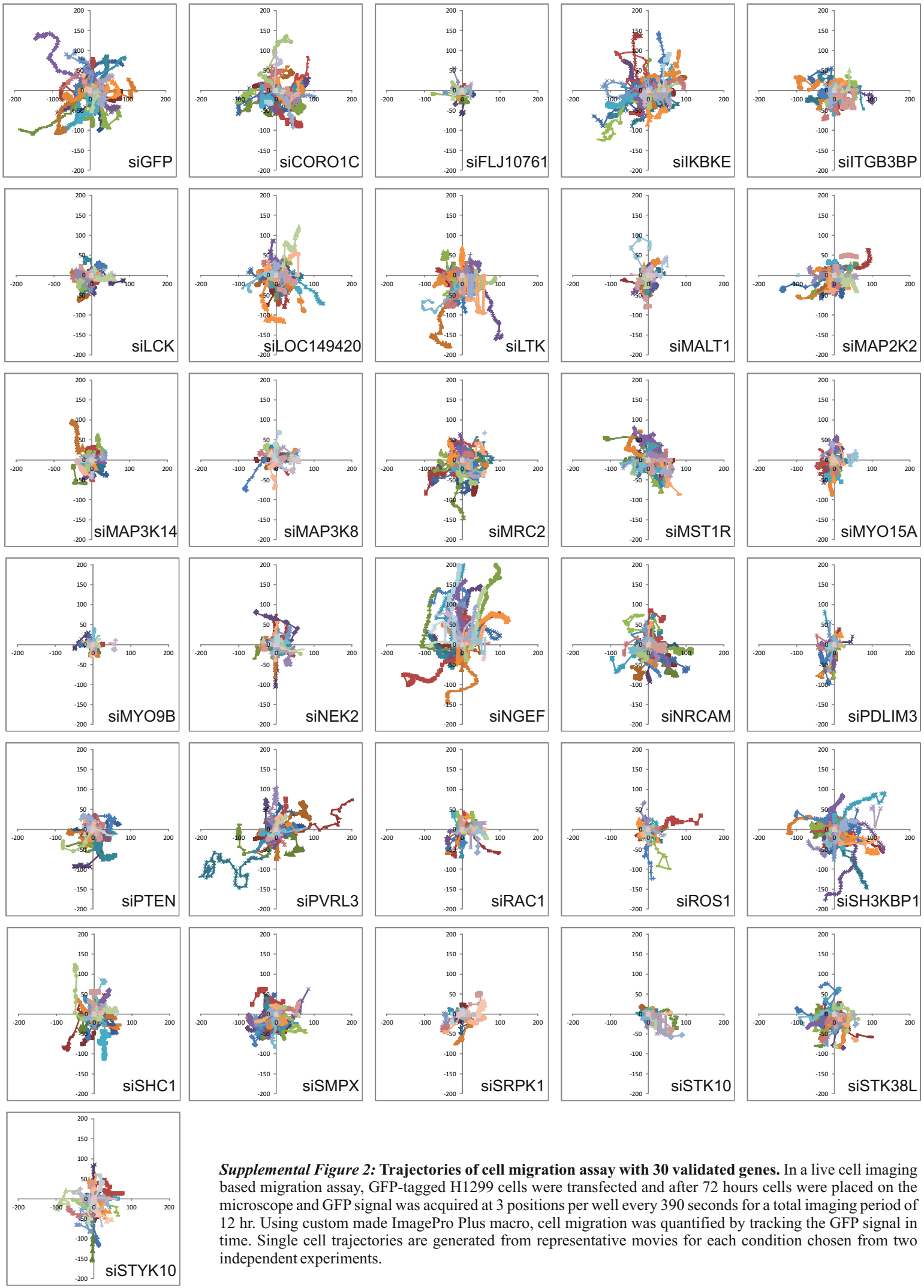
1. Zhang, Y., Moerkens, M., Ramaiahgari, S., de Bont, H., Price, L., Meerman, J., and van de Water, B. 2011. Elevated insulin-like growth factor 1 receptor signaling induces antiestrogen resistance through the MAPK/ERK and PI3K/Akt signaling routes. *Breast Cancer Res* 13:R52.
2. Le Devedec, S.E., van Roosmalen, W., Maria, N., Grimbergen, M., Pont, C., Lalai, R., and van de Water, B. 2009. An improved model to study tumor cell autonomous metastasis programs using MTLn3 cells and the Rag2(-/-) gammac (-/-) mouse. *Clin Exp Metastasis* 26:673-684.

Supplemental Figure 1 - Van Roosmalen et al



Supplemental Figure 1: **Unsupervised hierarchical clustering of 136 genes identified in the primary screen.** Unsupervised hierarchical clustering on the 136 identified genes revealed 4 clusters with signatures for specific migratory behavior. Cluster 1 comprises genes that after knockdown result in impaired cell migration with increased protrusive activity. Cluster 2 reflects impaired cell migration whereas cluster 3 includes genes that after knockdown result in increased directional cell migration. Cluster 4 contains genes that upon depletion lead to increased cell migration with increased cell area.

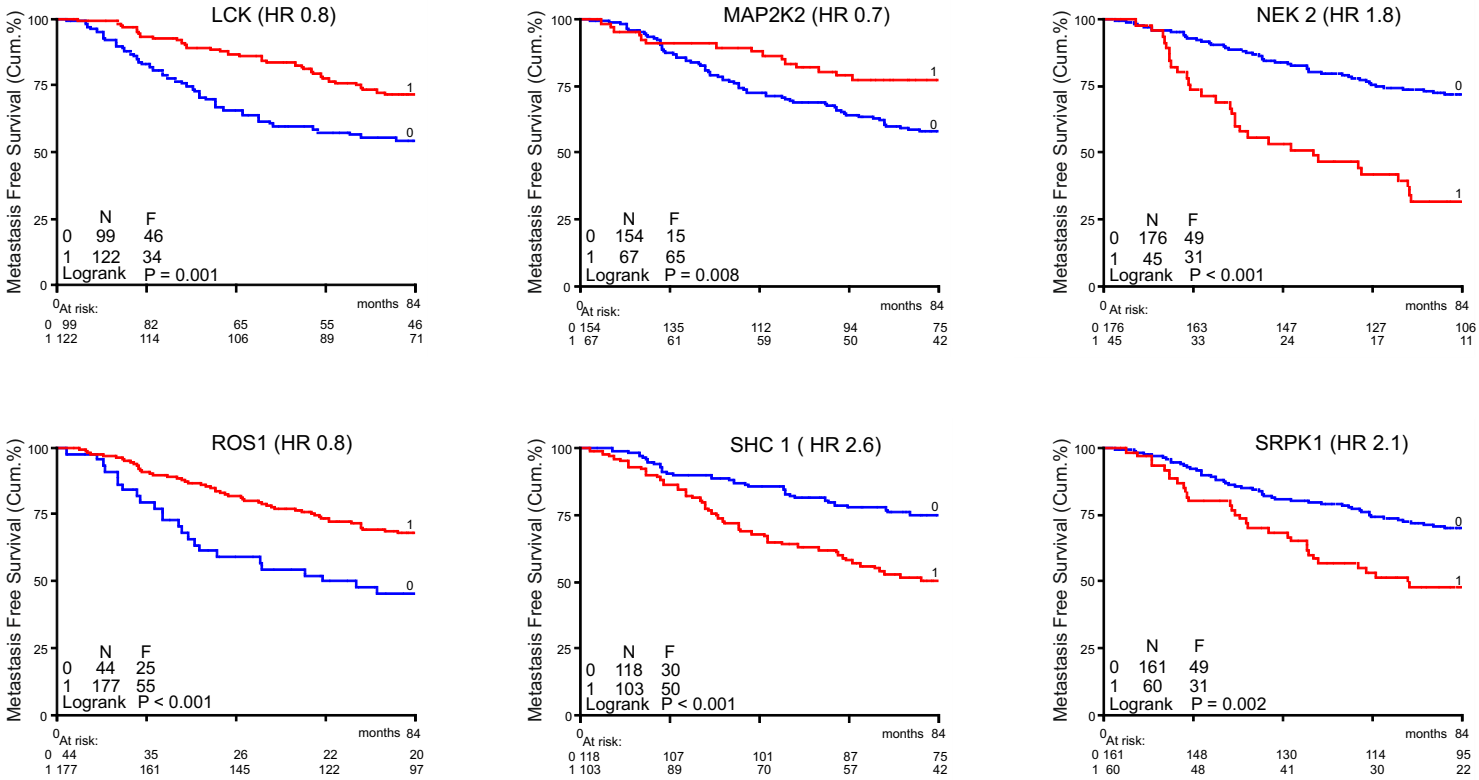
Supplemental Figure 2 - Van Roosmalen et al



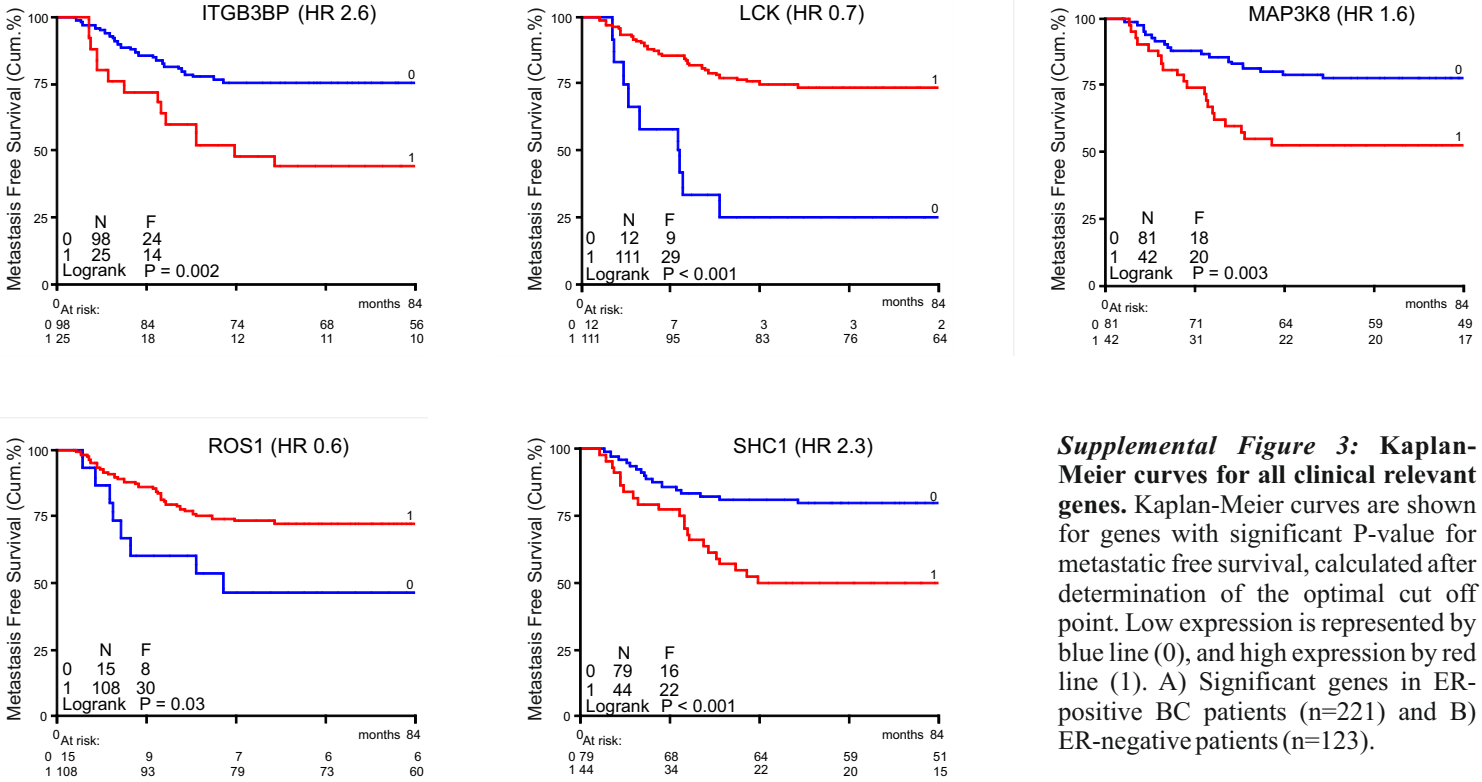
Supplemental Figure 2: **Trajectories of cell migration assay with 30 validated genes.** In a live cell imaging based migration assay, GFP-tagged H1299 cells were transfected and after 72 hours cells were placed on the microscope and GFP signal was acquired at 3 positions per well every 390 seconds for a total imaging period of 12 hr. Using custom made ImagePro Plus macro, cell migration was quantified by tracking the GFP signal in time. Single cell trajectories are generated from representative movies for each condition chosen from two independent experiments.

Supplemental Figure 3 - Van Roosmalen et al

A) ER-POS



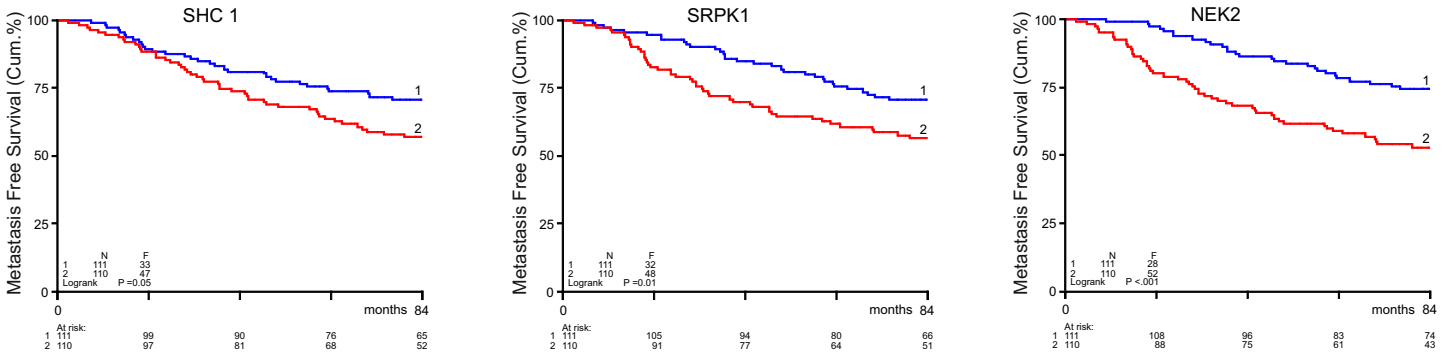
B) ER-NEG



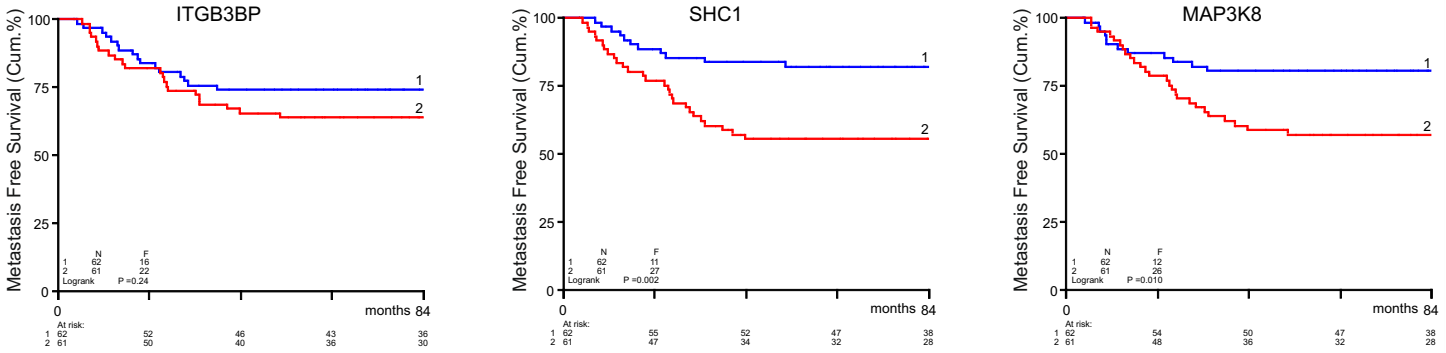
Supplemental Figure 3: **Kaplan-Meier curves for all clinical relevant genes.** Kaplan-Meier curves are shown for genes with significant P-value for metastatic free survival, calculated after determination of the optimal cut off point. Low expression is represented by blue line (0), and high expression by red line (1). A) Significant genes in ER-positive BC patients (n=221) and B) ER-negative patients (n=123).

Supplemental Figure 4 - Van Roosmalen et al

ER-POS



ER-NEG

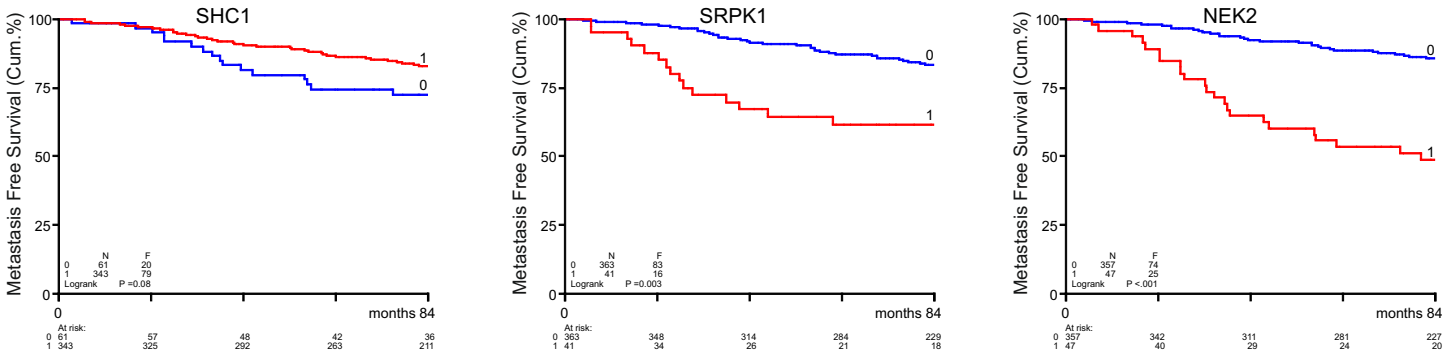


Supplemental Figure 4: **Clinical relevance of candidate tumor cell migration hits in breast cancer metastasis free survival.** Kaplan-Meier curves are shown for genes in Figure 3 for MFS drawn after median split of patient group (ER-positive n=221, ER-negative n=123). Low expression is represented by blue line (1), and high expression by red line (2).

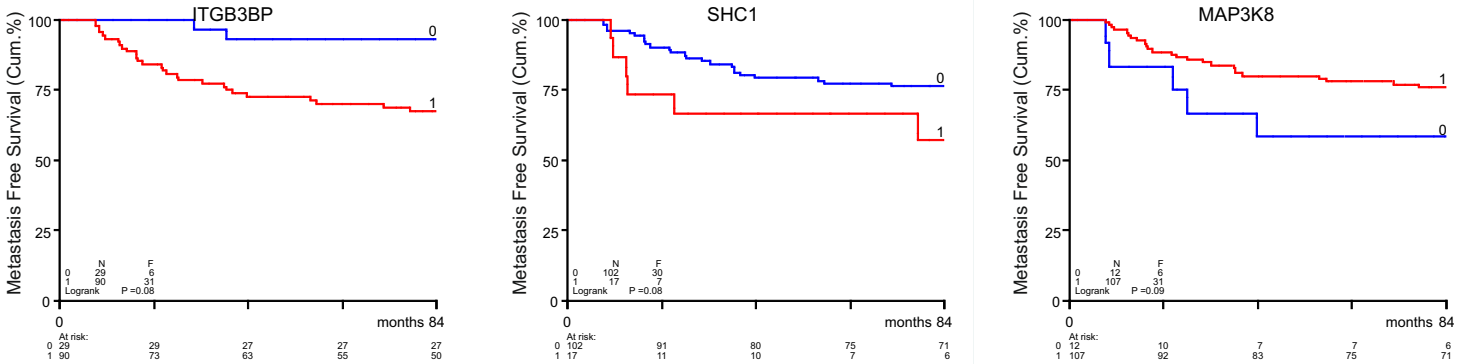
Supplemental Figure 5 - Van Roosmalen et al

A) OPTIMAL CUT OFF POINT

ER-POS

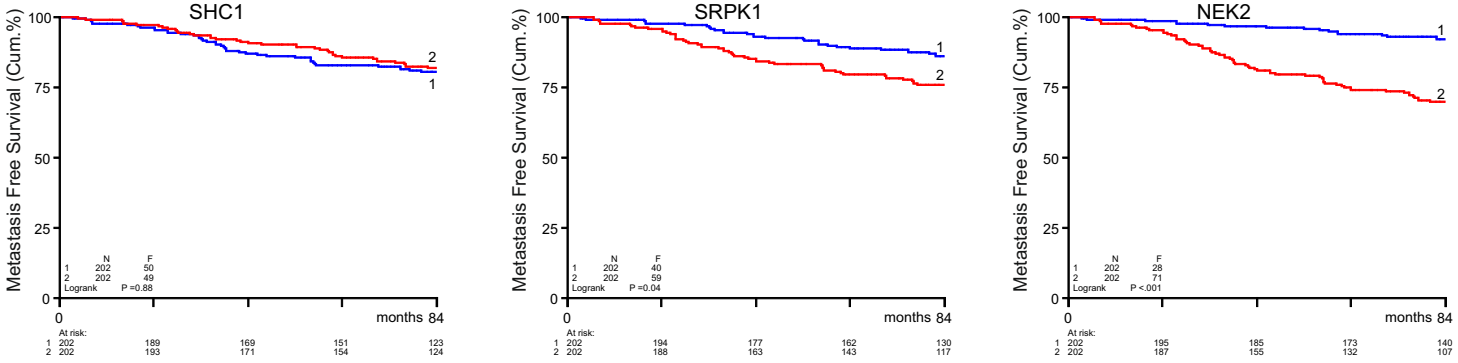


ER-NEG

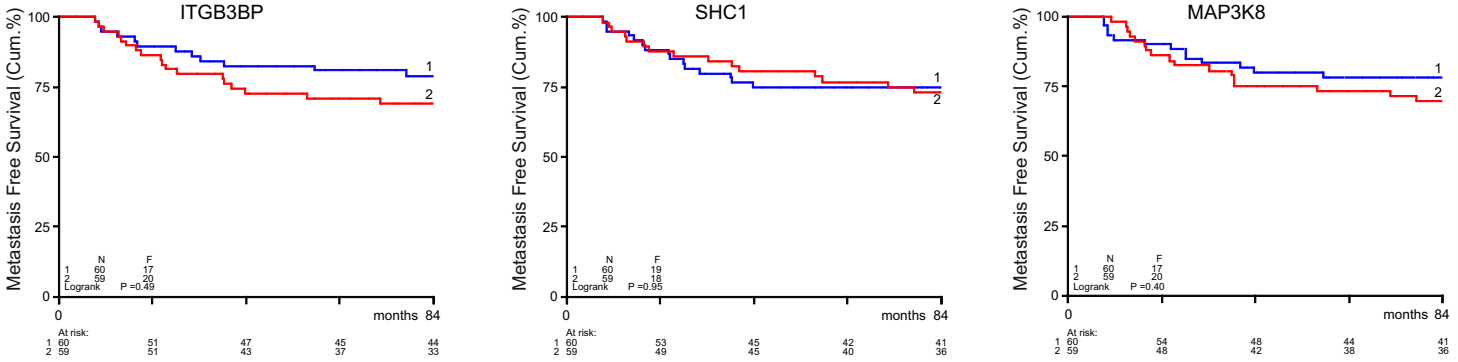


B) MEDIAN SPLIT

ER-POS

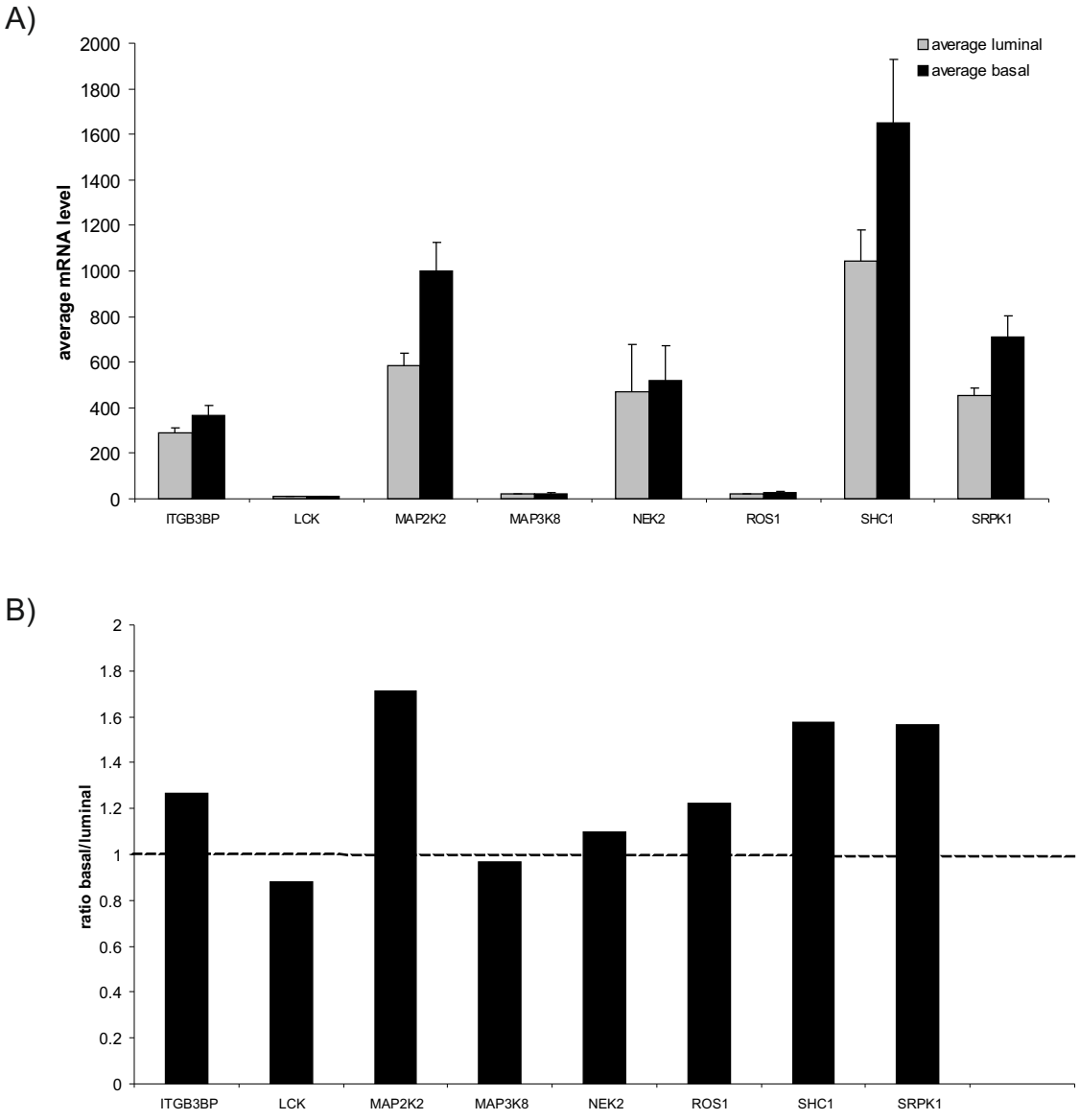


ER-NEG



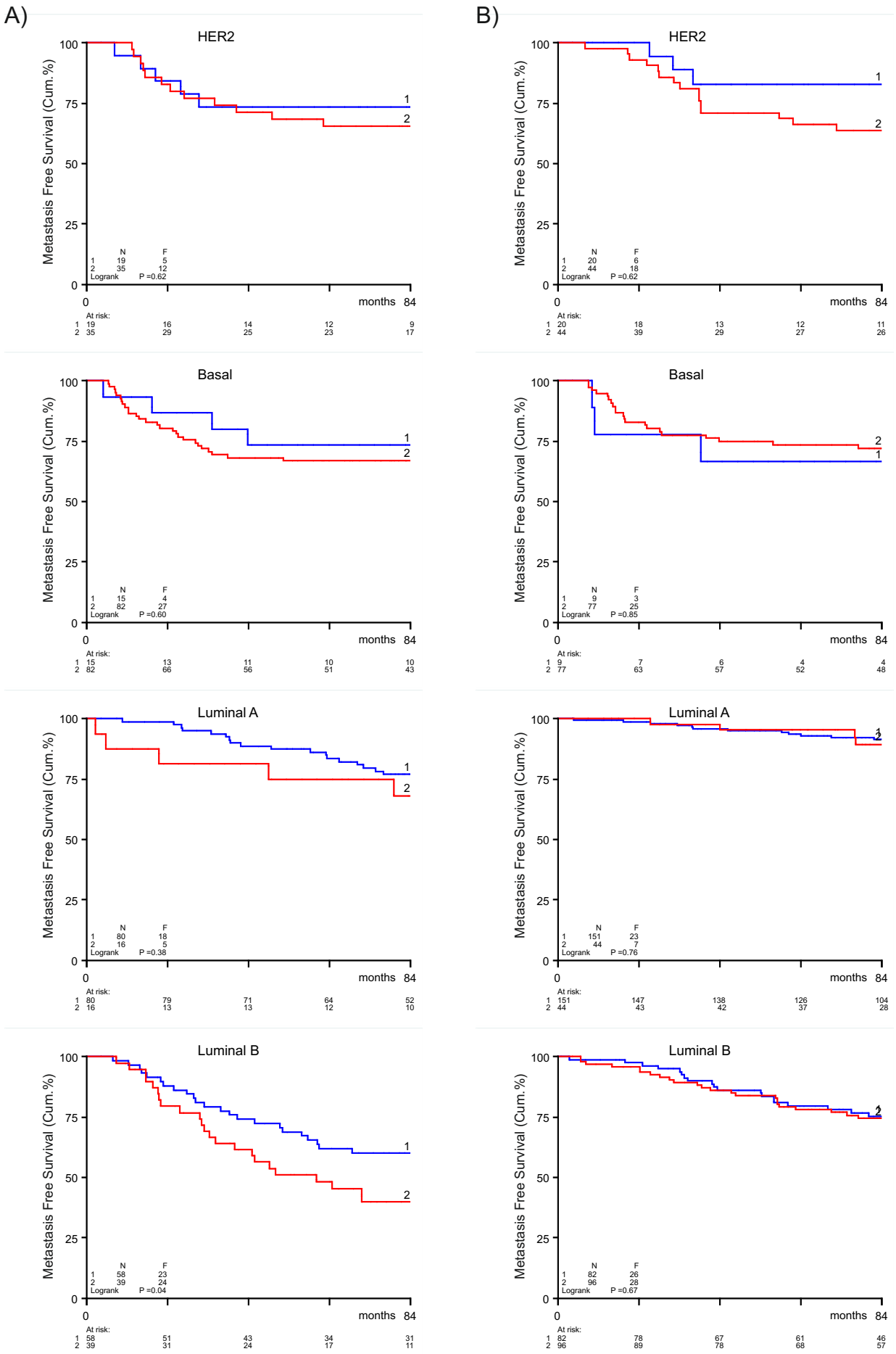
Supplemental Figure 5: **Clinical relevance of candidate tumor cell migration hits in breast cancer metastasis free survival in public dataset.** Analysis done on gene expression data combined from GSE2990 (n=189), GSE7390 (n=198) and GSE11121 (n=200). Kaplan-Meier curves are shown for genes in Figure 3 for MFS drawn after determination of the optimal cut off point (A, blue line is low expression (0), red line is high expression (1)) and median split of patient group (B, blue line is low expression (1), red line is high expression (2)).

Supplemental Figure 6 - Van Roosmalen et al



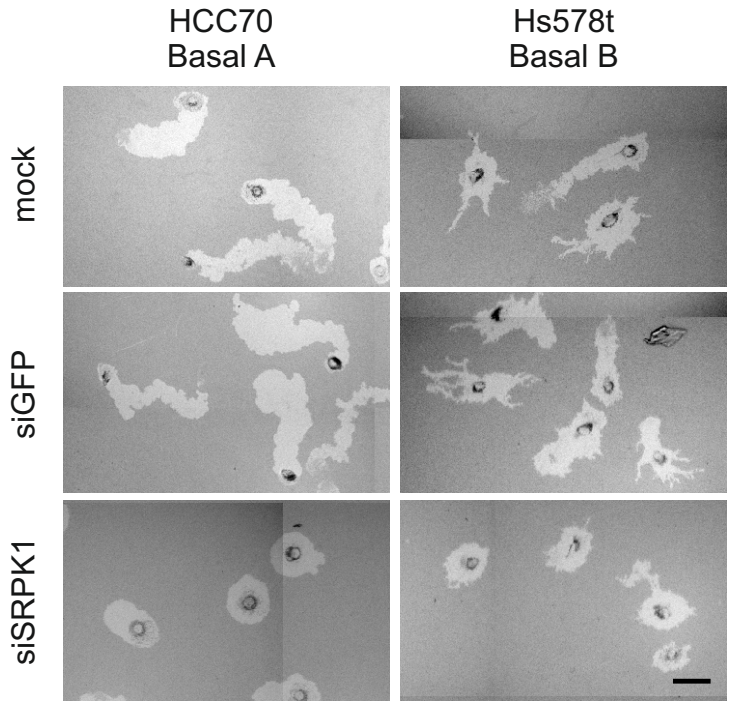
Supplemental Figure 6: **Breast cancer cell lines data mRNA expression migration hits.** The expression of the clinical relevant tumor cell migration genes in a panel of 22 luminal and 13 basal human BC cell lines was determined (A) and the expression ratio of basal over luminal cell lines was calculated (B).

Supplemental Figure 7 - Van Roosmalen et al



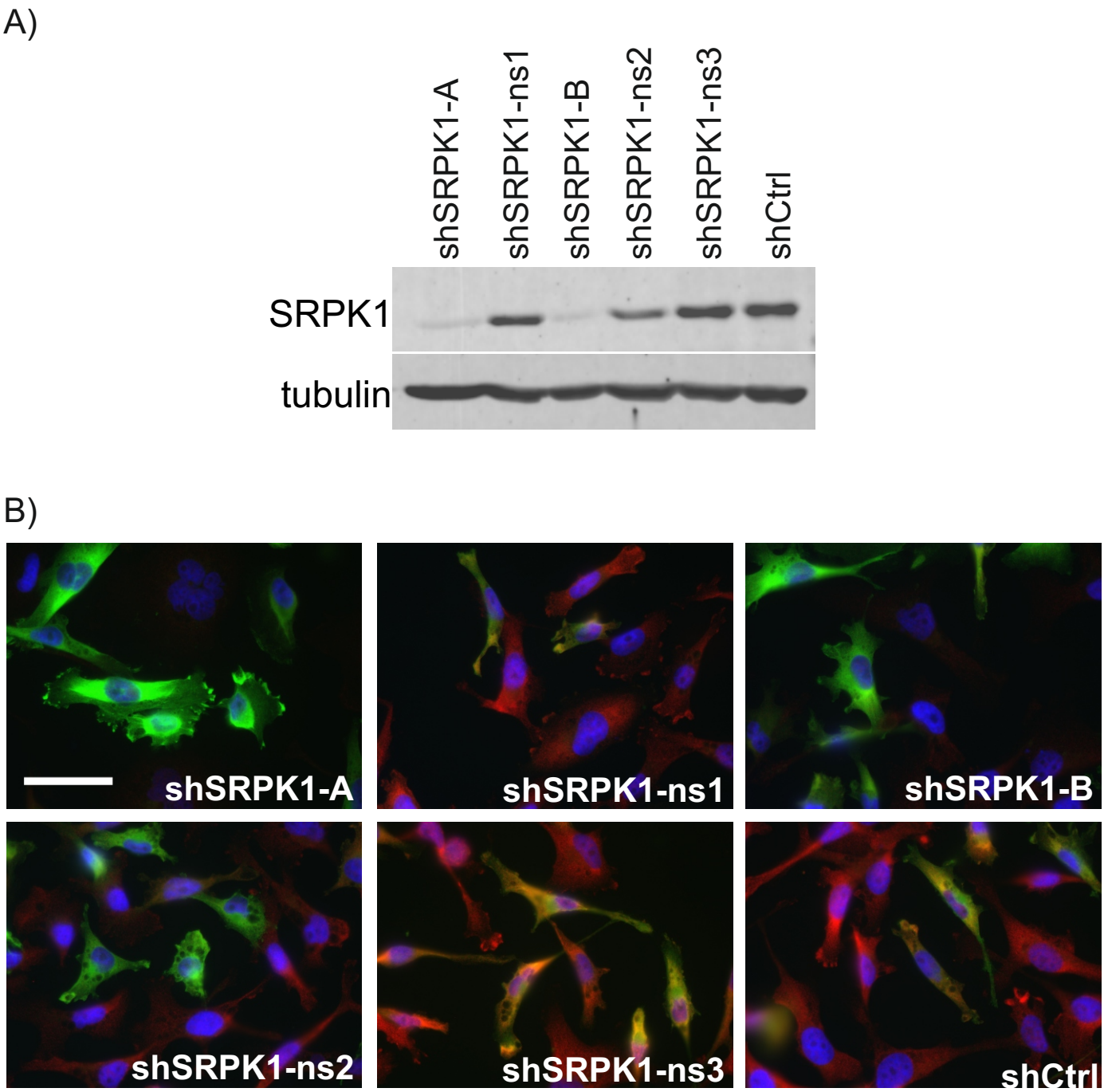
Supplemental Figure 7: **The correlation of *SRPK1* and metastasis free survival in different breast cancer subtypes .** Kaplan-Meier curves are shown for *SRPK1* expression in relation to MFS, drawn after determination of optimal cut off point (low log-rank P-value) of patient groups for different breast cancer subtypes. In A) graphs for patients as analyzed in Figure 3 are shown (n=344), in B) gene expression from public datasets (GSE2990 (n=189), GSE7390 (n=198) and GSE11121 (n=200)) is analyzed. Low expression is represented by blue line (1), and high expression by red line (2). BC subtypes were assigned using a subtype clustering model present in the 'genefu' R package (see methods section for more details).

Supplemental Figure 8 - Van Roosmalen et al



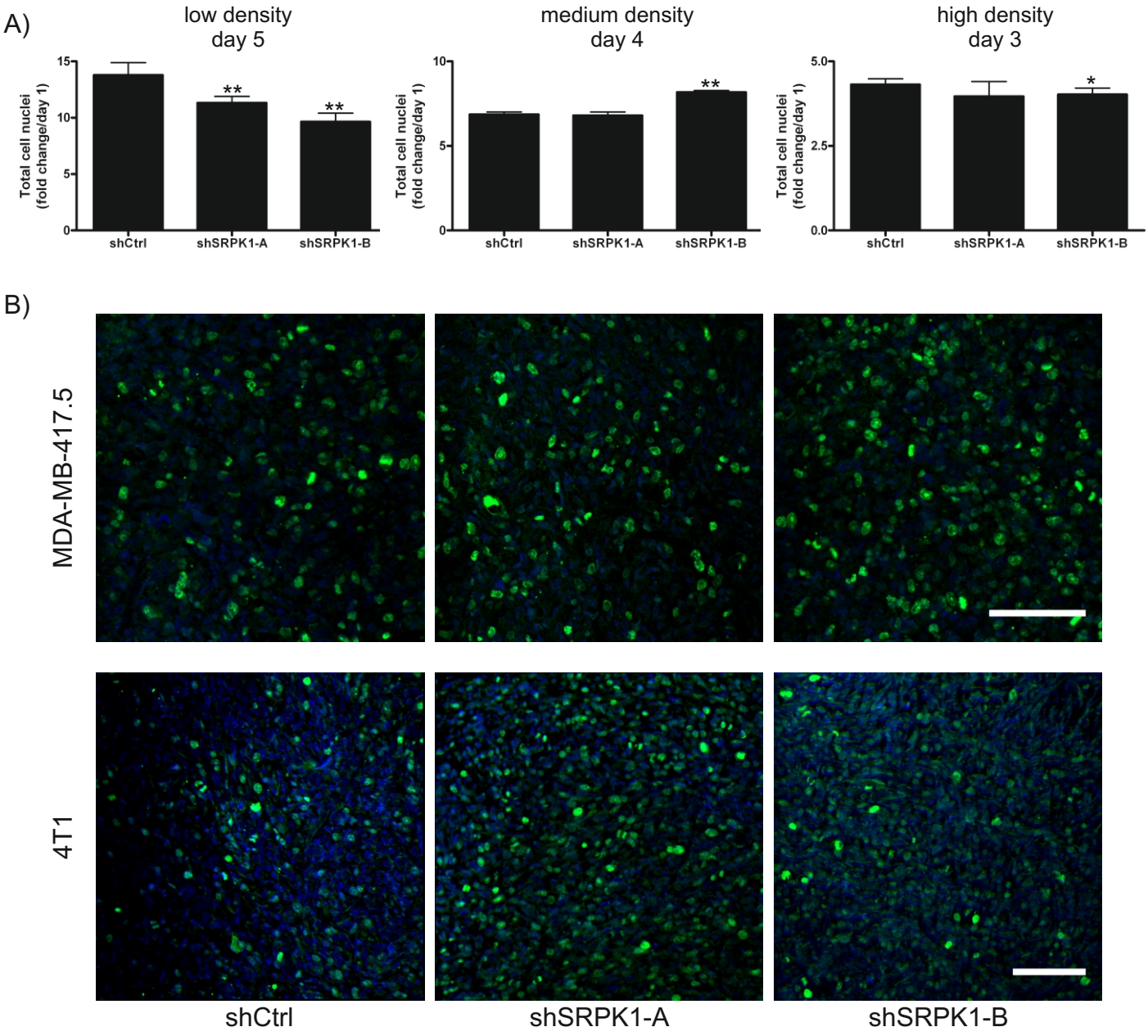
Supplemental Figure 8: **Knockdown of SRPK1 in human ER-negative BC HCC70 and Hs578t cells results in decreased cell migration.** Transient knockdown of SRPK1 in Hs578T and HCC70 blocked cell migration in a PKT assay. Cells were allowed to migrate on the bead-coated plates for 8 hr. Scale bar 50 μm. Representative images from two experiments with each four replicates are shown.

Supplemental Figure 9 - Van Roosmalen et al



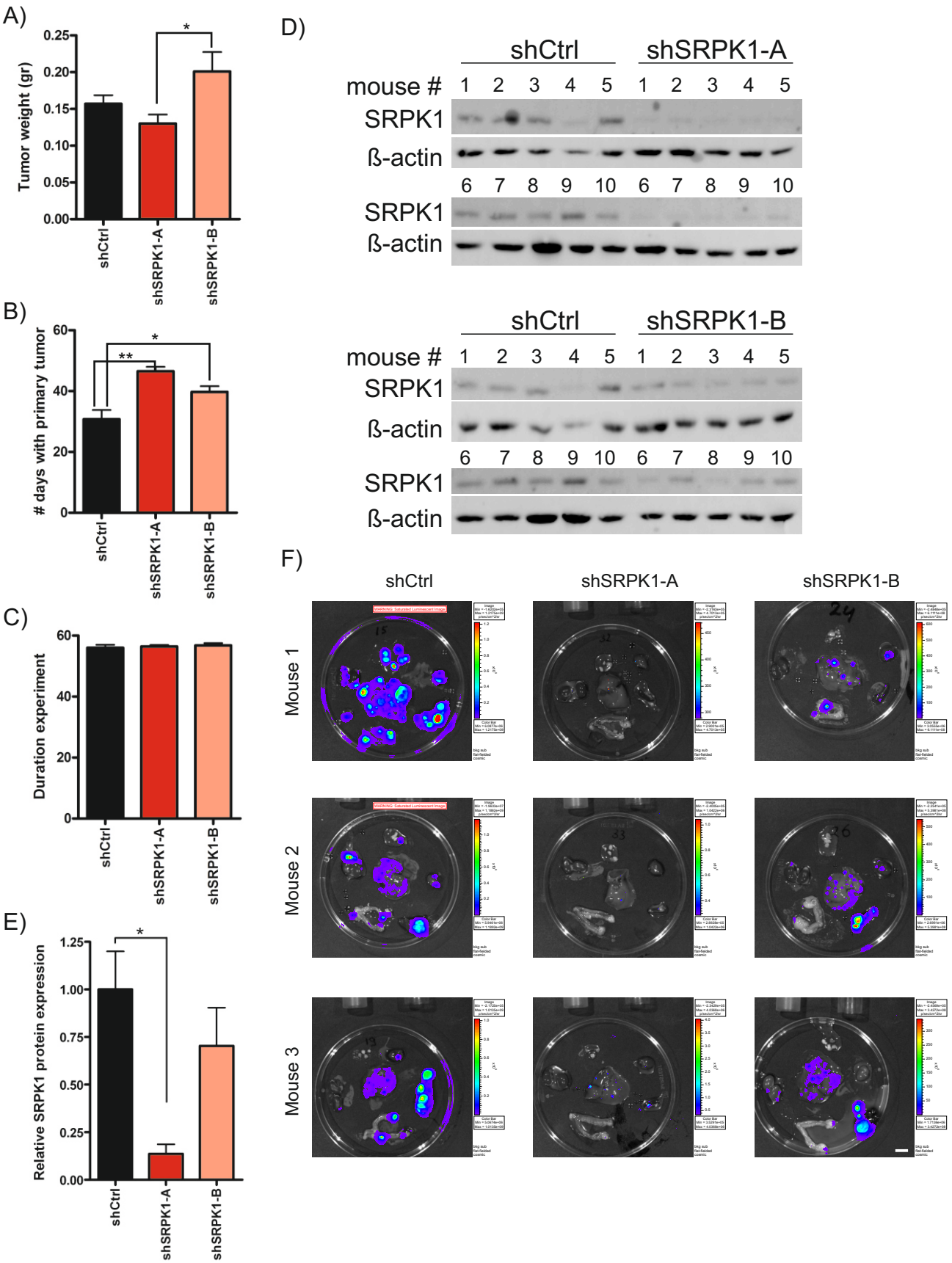
Supplemental Figure 9: **Establishment of stable SRPK1 knockdown cell lines in MDA-MB-417.5.** A) MDA-MB-417.5 cells were transduced with lentiviral shRNA specific for five independent sequences of human *SRPK1* and a non-targeting shRNA (shCtrl). SRPK1 knockdown was confirmed by Western blot. The two cell lines that showed almost complete depletion of SRPK1 were chosen for further experiments (shSRPK1-A and –B). B) Immunofluorescence staining against SRPK1 (in red) confirmed the reduced SRPK1 levels in both cell lines. GFP signal comes from MDA-MB-417.5. Cells were additionally stained with Hoechst to visualize nuclei. Scale bar 50 μm. NS = no significant knockdown effect.

Supplemental Figure 10 - Van Roosmalen et al



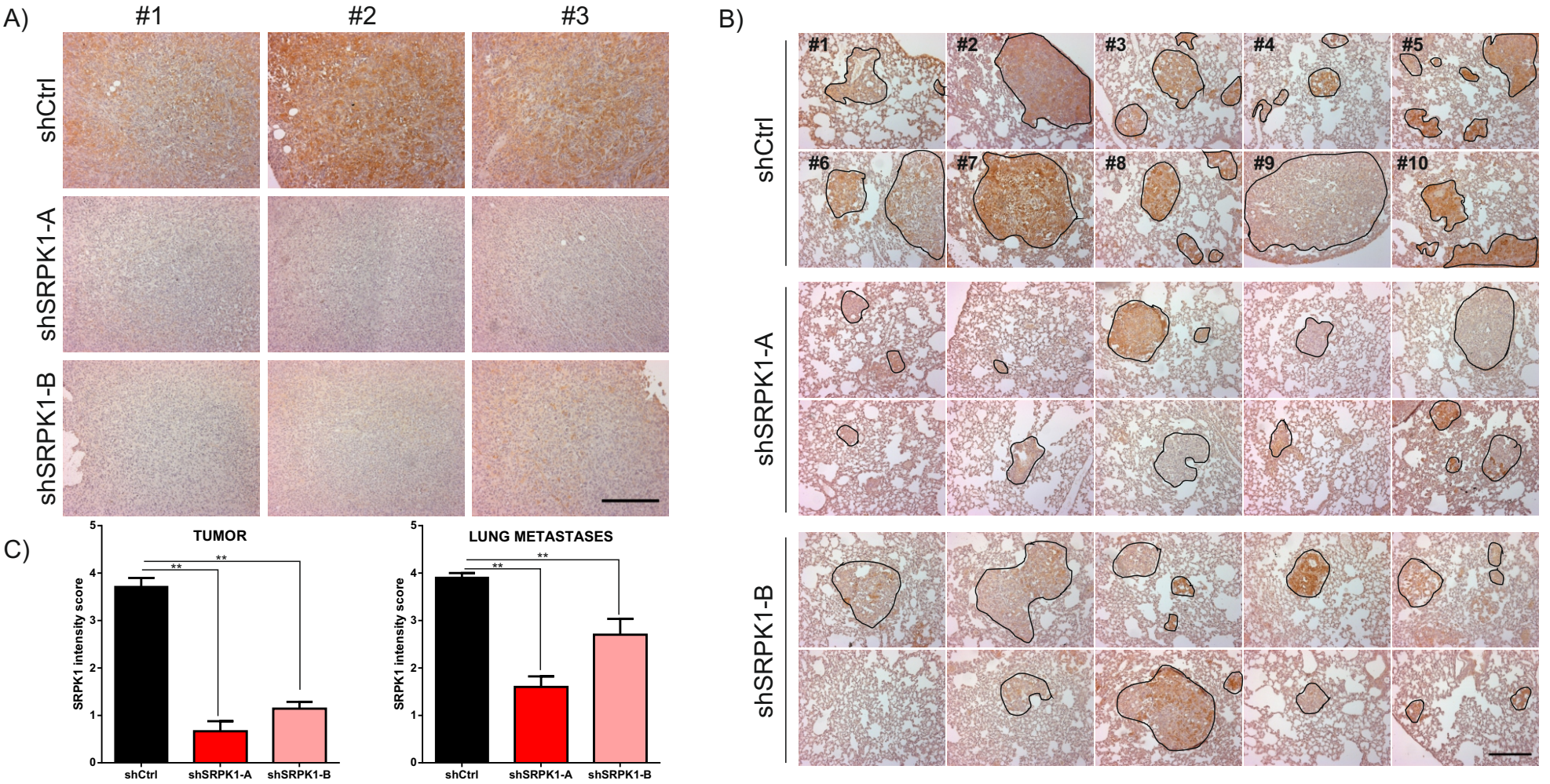
Supplemental Figure 10: **Effect of SRPK1 knockdown on cell proliferation in vitro and in vivo.** A) MDA-MB-417.5 cell numbers were determined by counting of nuclei per well after staining with Hoechst at day 5 post plating (low density), day 4 (medium density) or day 3 (high density). Values were normalized against day 1 and fold change was calculated (n=8) (two-tailed t-test, *P<0.05, **P<0.01, tested against shCtrl). B) Primary tumor sections were stained for the proliferation marker Ki67 (green) and Hoechst (blue). Representative images for each group are shown for MDA-MB-417.5 (n=10, upper images) and for 4T1 (n=10-11, lower images). Scale bar 100 μm.

Supplemental Figure 11 - Van Roosmalen et al



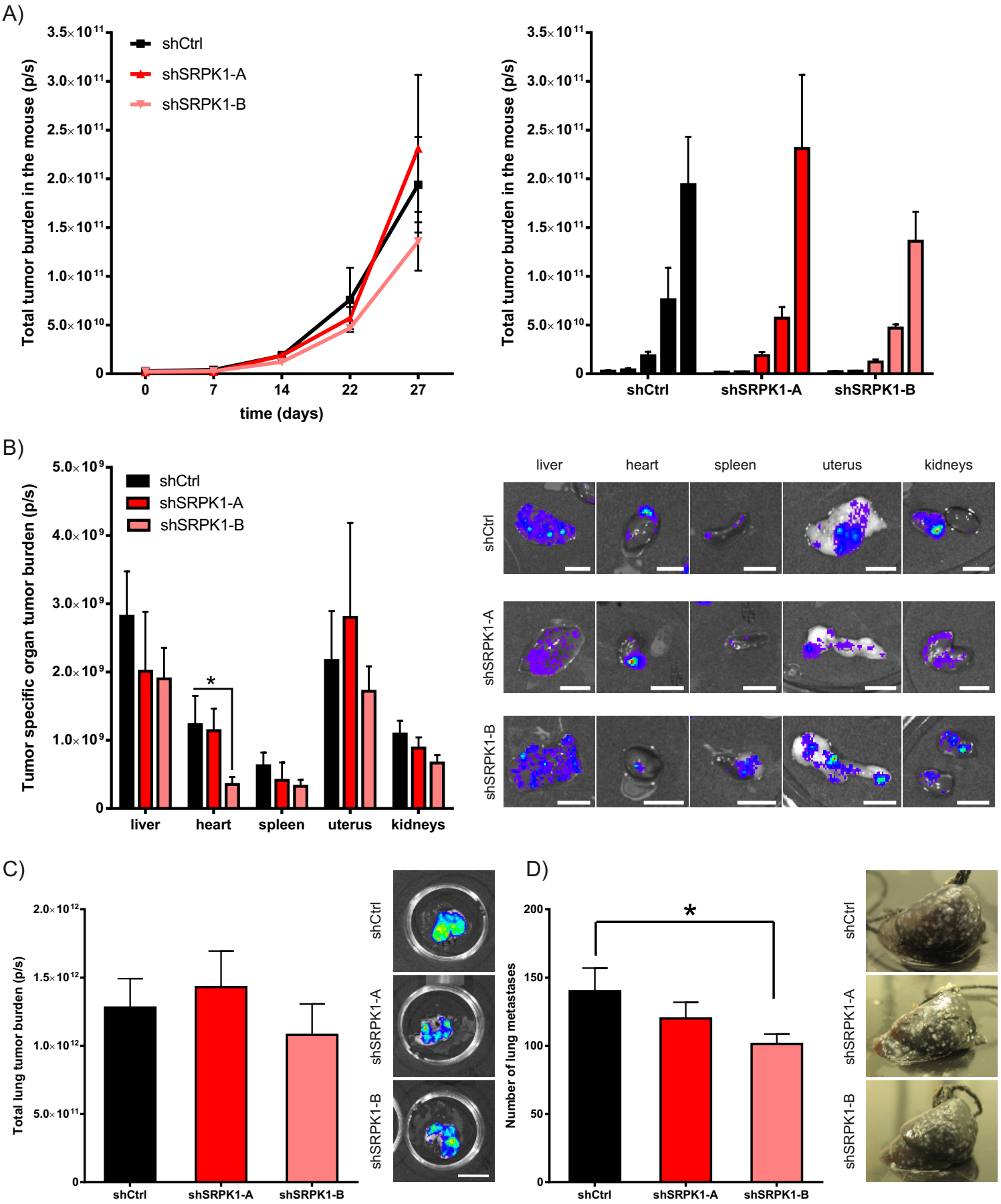
Supplemental Figure 11: **Knockdown of SRPK1 in MDA-MB-417.5 reduces their metastatic potential.** A) MDA-MB-417.5 cells were injected in 7-9 week old *RAG2-/-///IL2RG-/-* mice (n=10) and tumor growth was followed over time. Primary tumors were removed when they reached ~250 mm3 volume. At time of removal, tumor weight was similar in all groups (two-tailed t-test, $P<0.05$). B) Tumor growth was delayed in mice that received shSRPK1 cells (two-tailed t-test, $*P<0.05$, $**P<0.01$). C) However, total duration of the experiment was kept constant between different groups. D) Decreased protein expression of SRPK1 within primary tumors was detected using Western blot for all individual tumors in each group. E) Quantification of SRPK1 levels from blots shown in D). Each blot is corrected for an internal loading control and relative SRPK1 protein levels are calculated relative to the average value for shCtrl (two-tailed t-test, $P<0.05$). F) Examples of ex vivo imaging of bio-luminescence. Scale bar 1 cm. All mice (n=10) in each group were analyzed this way and average radiance per organs was quantified (as shown in Figure 6B).

Supplemental Figure 12 - Van Roosmalen et al



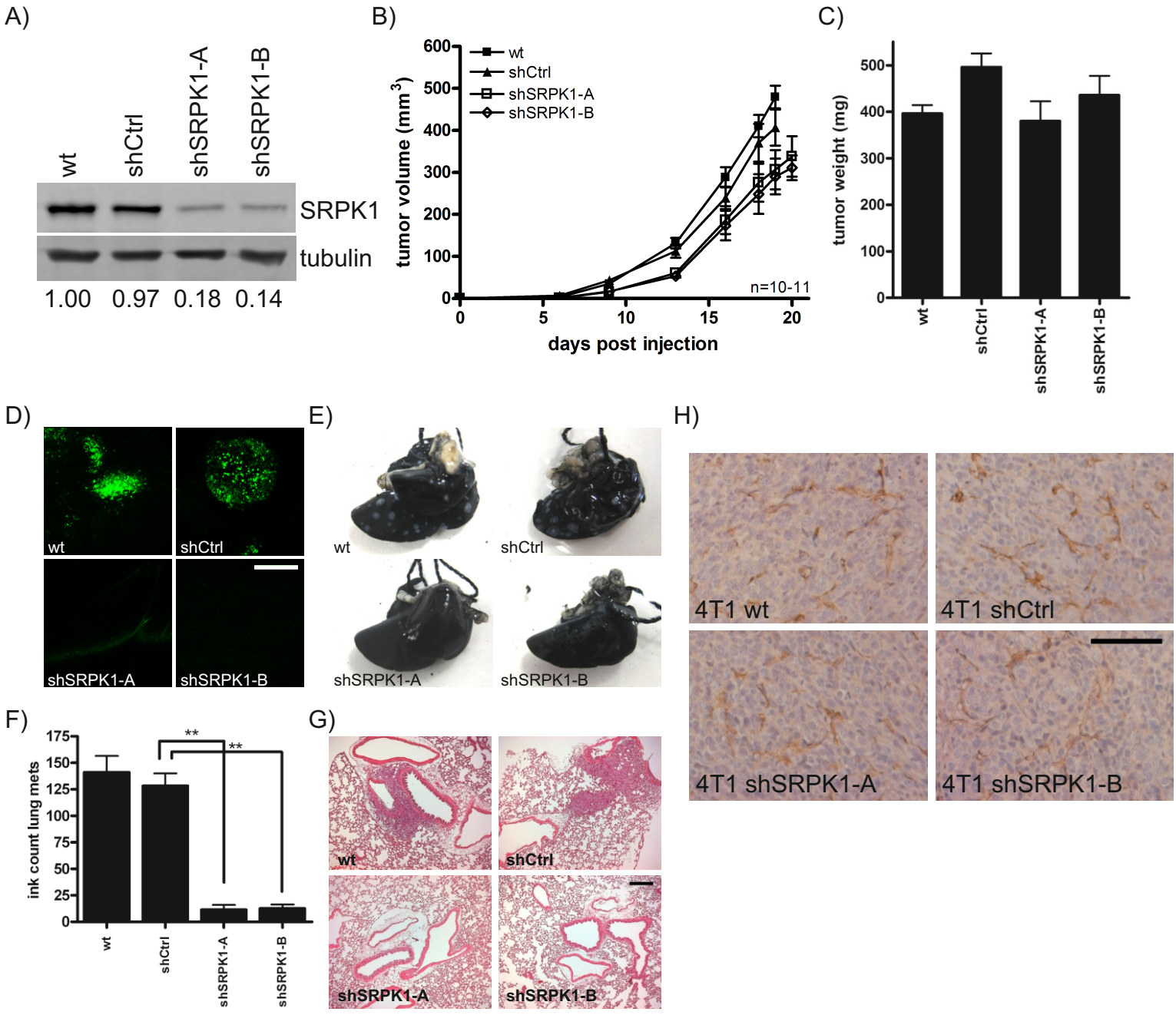
Supplemental Figure 12: **SRPK1 staining on primary tumors and lung metastases.** Primary tumors (A) and lung lobes (B) were embedded in paraffin and processed for immunohistological analysis of SRPK1 expression. SRPK1 intensity was scored in which 0 = no staining and 4 = strong staining. Scale bar 100 μm. C) Average score for each group (n=10) is shown in bar plots (right) for both tumor and lung metastases (two-tailed t-test, ** P<0.01).

Supplemental Figure 13 - Van Roosmalen et al



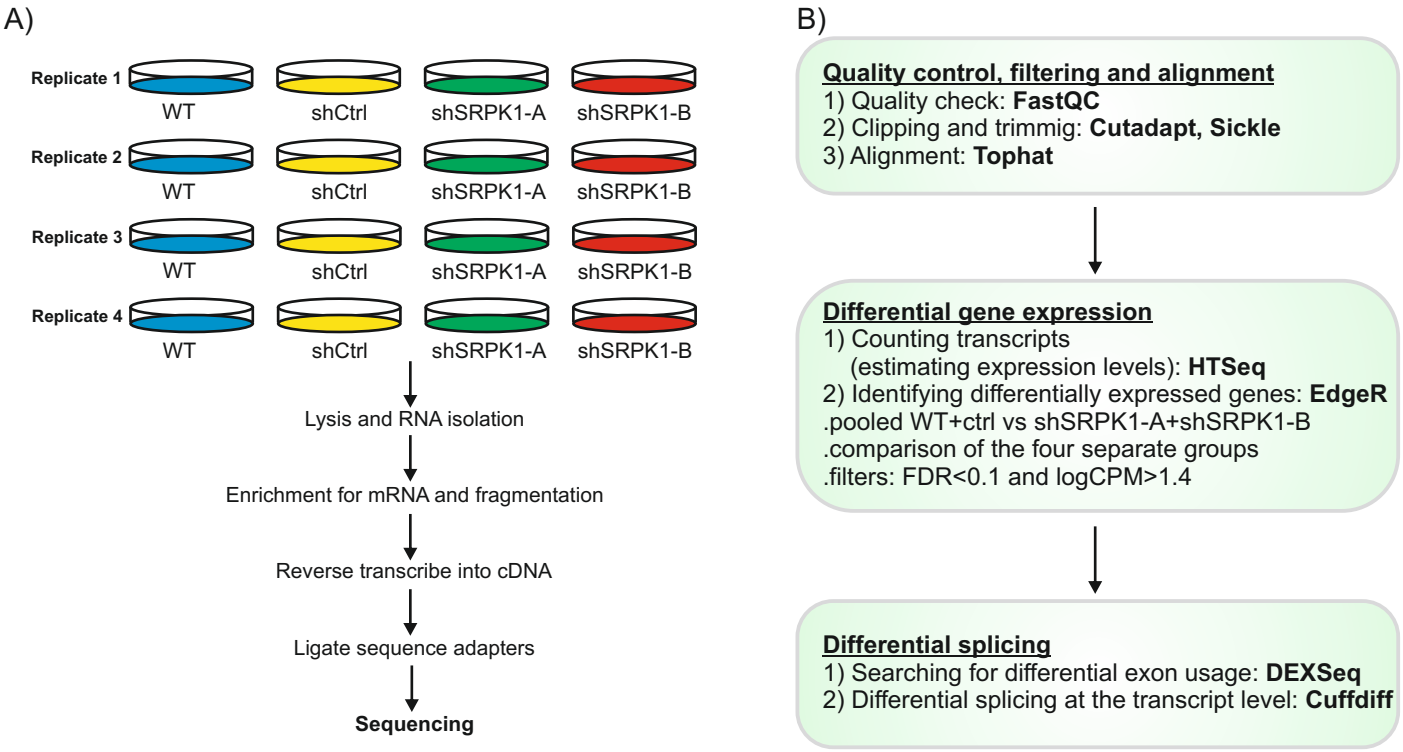
Supplemental Figure 13: **SRPK1 is not required in an experimental metastasis assay.** A) MDA-MB-417.5 cells were injected directly in the bloodstream via the tail vein of *RAG2-/-//IL2RG-/-* mice (n=7). Bio-luminescence imaging was done at day 7,14, 22 and 27 post injection. Total tumor burden measured in whole animals prior to sacrifice is plotted against time in line graphs (left) as well as bar plots including standard deviation (right). B) At day 27, animals were sacrificed and total radiance per organ was quantified ex vivo. Representative images are shown for each organ and each group. Scale bar 1 cm. C) Total radiance of lungs was measured separate from other organs due to high signal. Scale bar 1 cm. D) After imaging, lungs were injected with Indian ink and surface metastases were quantified in upper small lobe. Significant differences are highlighted in the graphs (two-tailed t-test, $P<0.05$). All other conditions are not statistically different.

Supplemental Figure 14 - Van Roosmalen et al



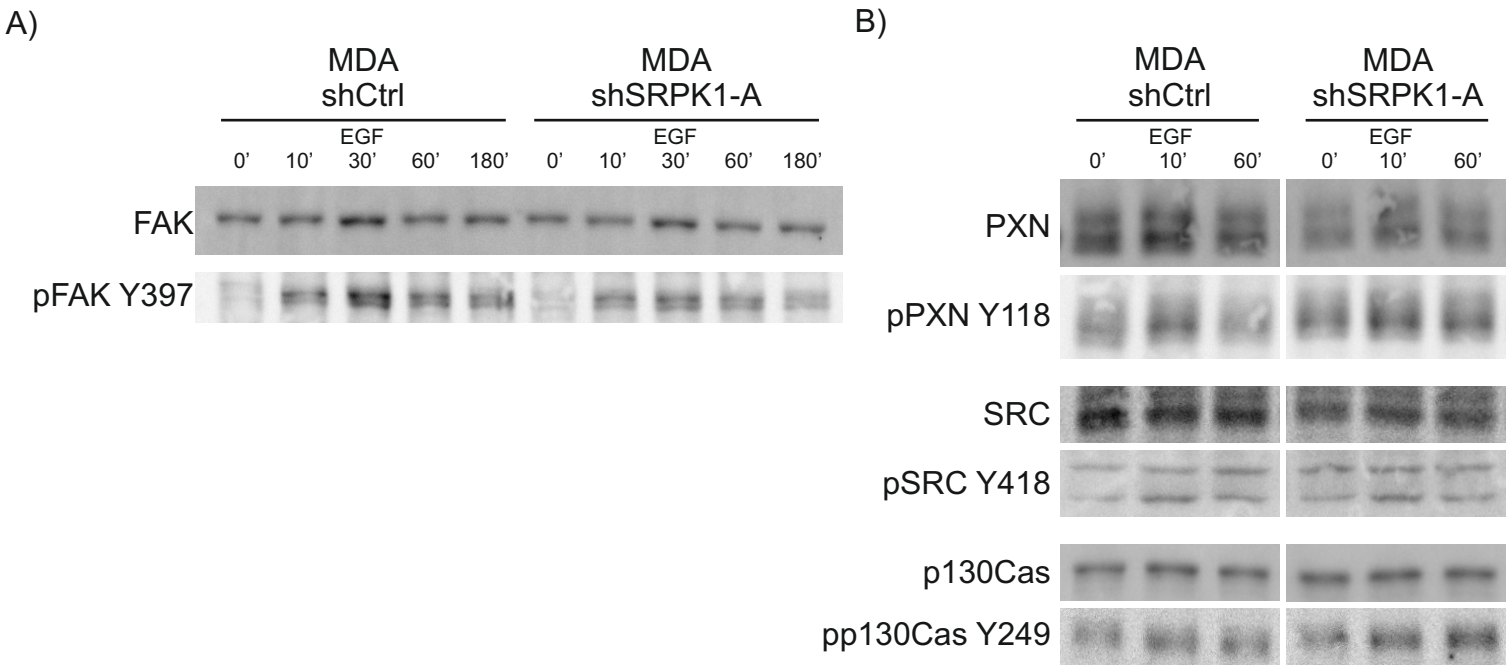
Supplemental Figure 14: **SRPK1 is required for breast cancer metastasis in vivo.** A) 4T1 cells were transduced with lentiviral shRNA specific for two independent sequences of *SRPK1* (shSRPK1-A and SRPK1-B) and a non-targeting shRNA (shCtrl). SRPK1 knockdown was confirmed by Western blot. 4T1 cells were injected in 6-8 week old *RAG2*$^{-/-}$*//IL2RG*$^{-/-}$ mice (n=10-11) and tumor growth was followed over time. Tumor growth (B) and weight (C) was similar in all groups. Three weeks post injection, mice were sacrificed and lungs were isolated. GFP positive tumor cells in the lung were visualized. Scale bar 500 µm. All lungs were injected with Indian ink (E) and surface metastases were quantified. Knockdown of SRPK1 significantly reduced the number of metastases in lung (two-tailed t-test, $P<0.01$) (F). For each animal, the small lobe was processed for immunohistological analysis and stained with H&E (G). Scale bar 100 µm. High SRPK1 expression is associated with significantly more lung metastases. H) Immunohistological analysis of CD31 expression in 4T1 control and SRPK1 knockdown tumors revealed no differences in tumor vascularization. Scale bar 100 µm.

Supplemental Figure 15 - Van Roosmalen et al



Supplemental Figure 15: **Experimental set-up and analysis pipeline for generation of RNA-seq data of Ctrl and knockdown MDA-MB-417.5 cell lines.** A) Scheme for RNA sampling. B) Steps followed for the analysis of the deep sequencing data included i) extracting genes showing significant differential expression and ii) identifying genes with potential differential splicing.

Supplemental Figure 16 - Van Roosmalen et al

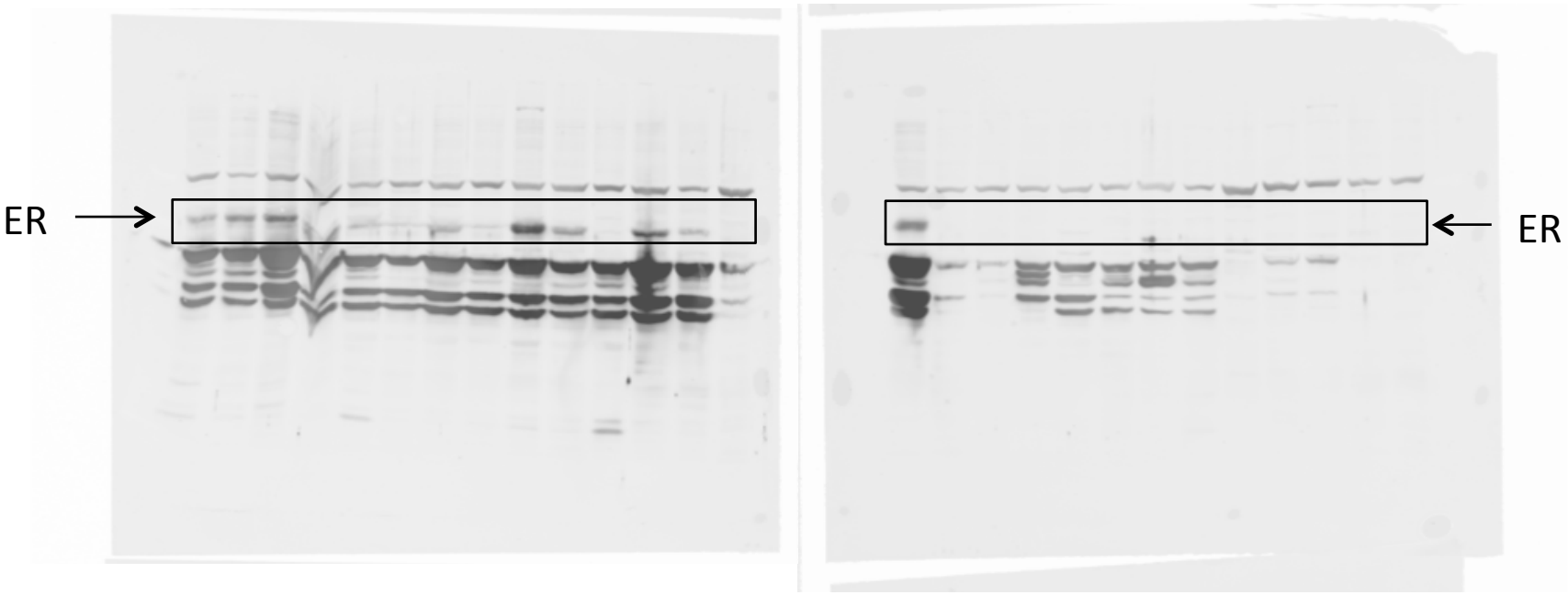


Supplemental Figure 16: **Knockdown of SRPK1 is affecting focal adhesion signaling in MDA-MB-417.5.** A) MDA-MB-417.5 cells were starved for 2 hr in medium containing 0.5% FBS and after that exposed to 100 ng/mL EGF for 0, 10, 30, 60 and 180 minutes. Cell lysates were prepared and Western blot analysis revealed decreased phosphorylation of FAK at Y397 upon stimulation with EGF in SRPK1 knockdown cells compared to control cells. Other phosphorylation sites of FAK were included in the analysis, but no difference in shCtrl versus shSRPK1 cells was found. B) Other proteins of the focal adhesion signaling cascade were examined as well. In shSRPK1 cells, the phosphorylation of PXN at Y118 following EGF stimulation is prolonged whereas no difference in phosphorylation of SRC between the two cell lines was found. The phosphorylation of p130Cas at Y249 is enhanceD in shSRPK1 cells. Cell lysates of shCtrl and shSRPK1 cells were analyzed on the same blot but were noncontiguous. This experiment has been done twice and representative blots are shown.

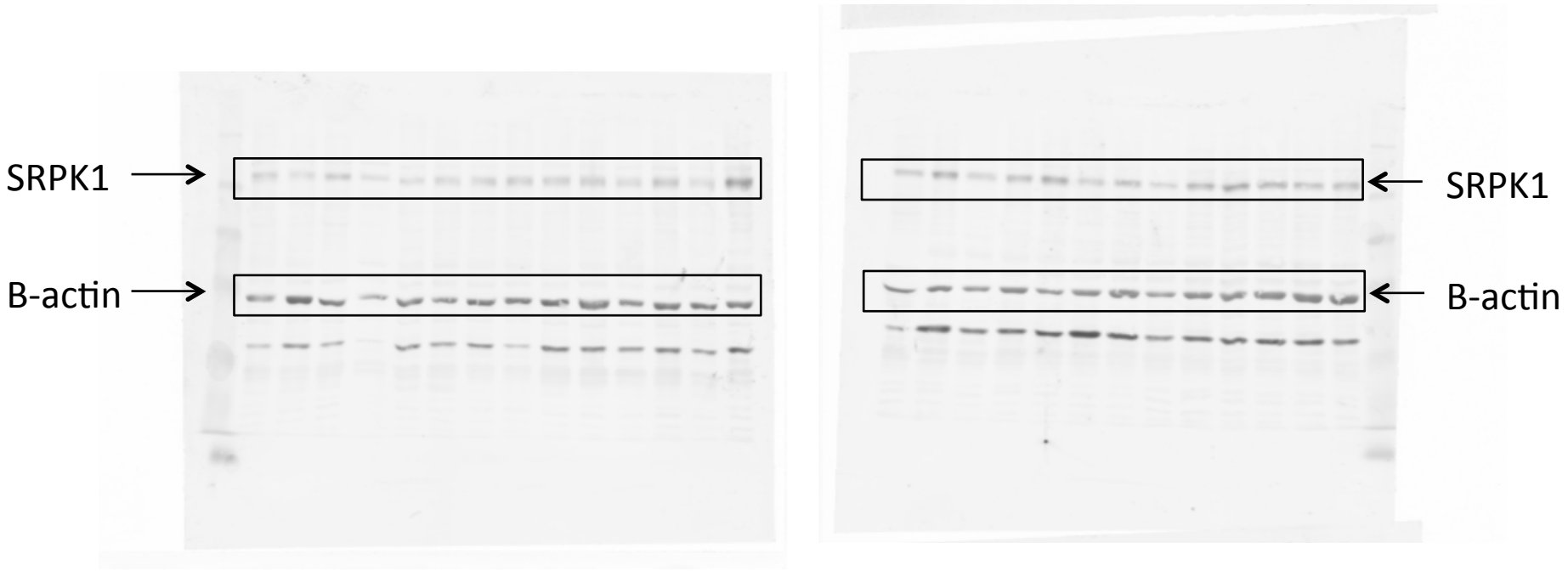
Gene Symbol	Probeset	ALL HR	ALL P-value	ER+ HR	ER+ P-value	ER- HR	ER- P-value
ITGB3BP	205176_s_at	1,2	0,227	0,9	0,811	2,6	**0,014**
LCK	204890_s_at	0,8	**<0.001**	0,8	**0,004**	0,7	**0,026**
MAP2K2	202424_at	0,8	**0,030**	0,7	**0,036**	0,8	0,597
MAP3K8	205027_s_at	1,0	0,946	0,8	0,218	1,6	**0,022**
NEK2	204641_at	1,5	**<0.001**	1,8	**<0.001**	1,4	0,077
ROS1	207569_at	0,7	**0,001**	0,8	**0,022**	0,6	**0,018**
SHC1	214853_s_at	2,4	**<0.001**	2,6	**<0.001**	2,3	**0,006**
SRPK1	202200_s_at	1,3	0,133	2,1	**0,007**	1,2	0,496

Supplemental Table 2: Clinical relevance of candidate tumor cell migration hits in breast cancer metastasis free survival. Gene expression data of a lymph node negative BC patient cohort (n=344) without prior treatment was analyzed. Expression of identified tumor cell migration genes was correlated with metastasis free survival (MFS) and Hazard Ratio (HR) as well as Cox P-values were calculated for estrogen receptor (ER)-negative (n=123) and -positive (n=221) BC patients, as well as all patients combined. Eight out of 30 validated genes showed clinical relevance ($P<0.05$, highlighted in bold).

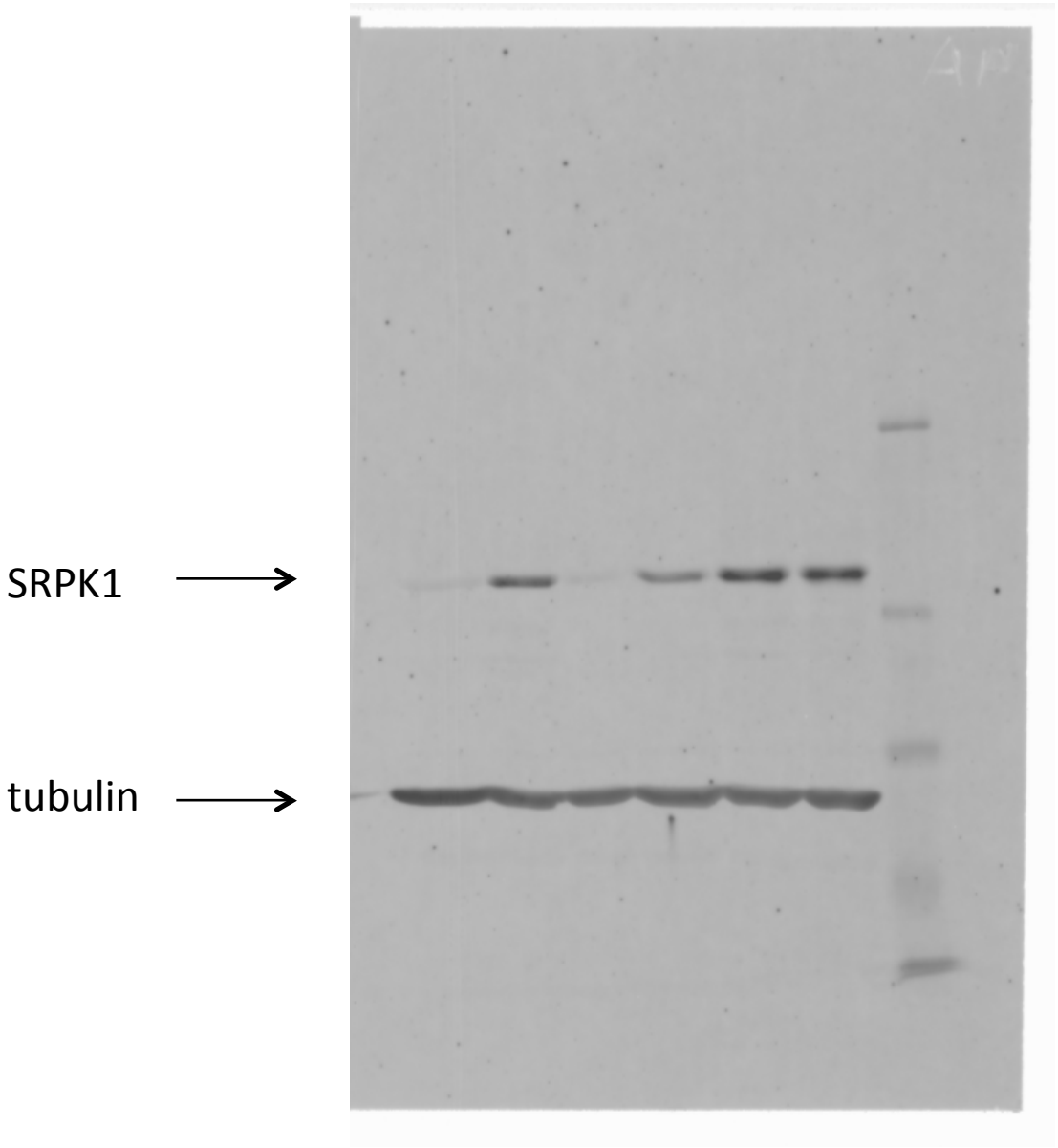
Full unedited gels for Figure 5A



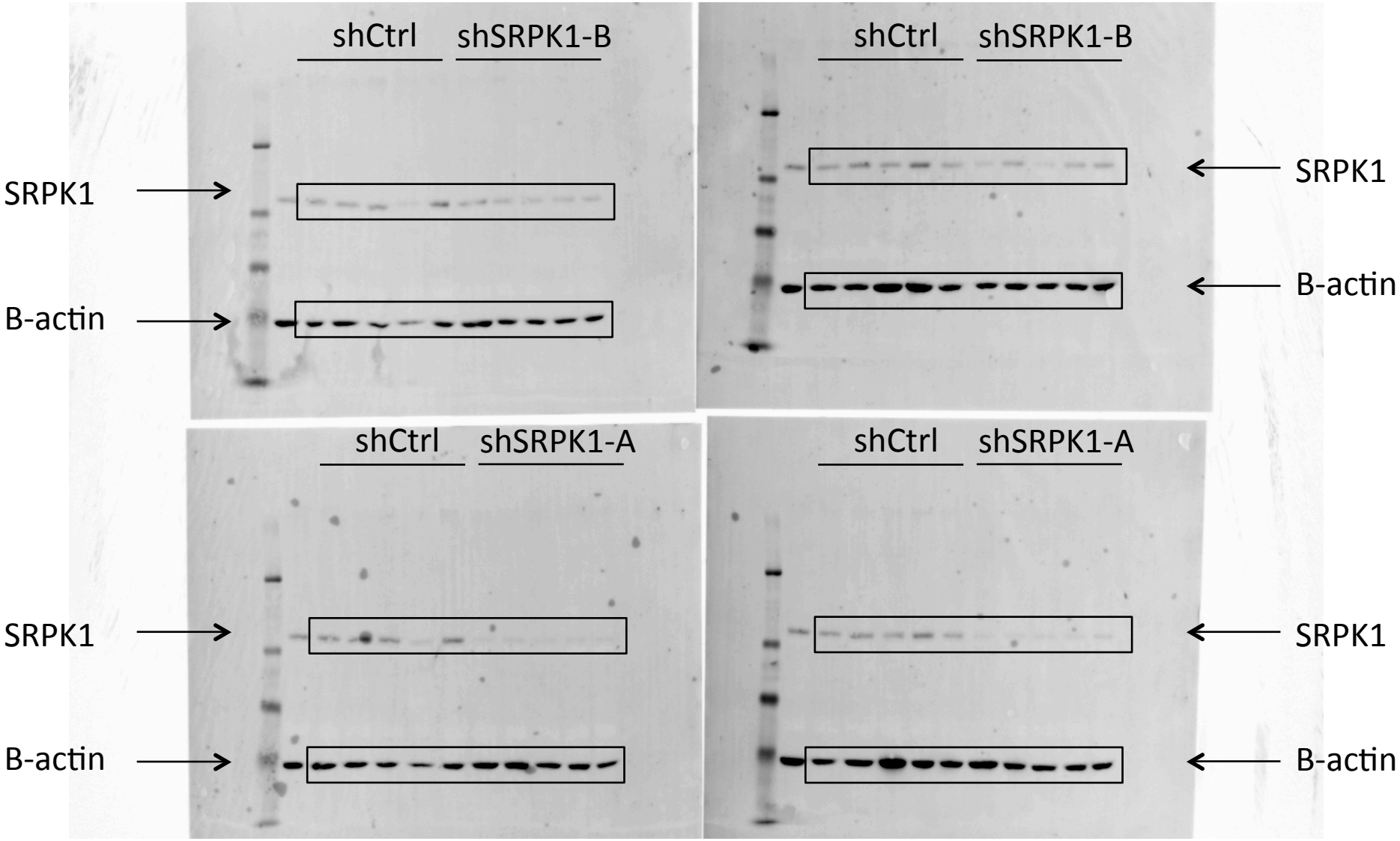
Full unedited gels for Figure 5A



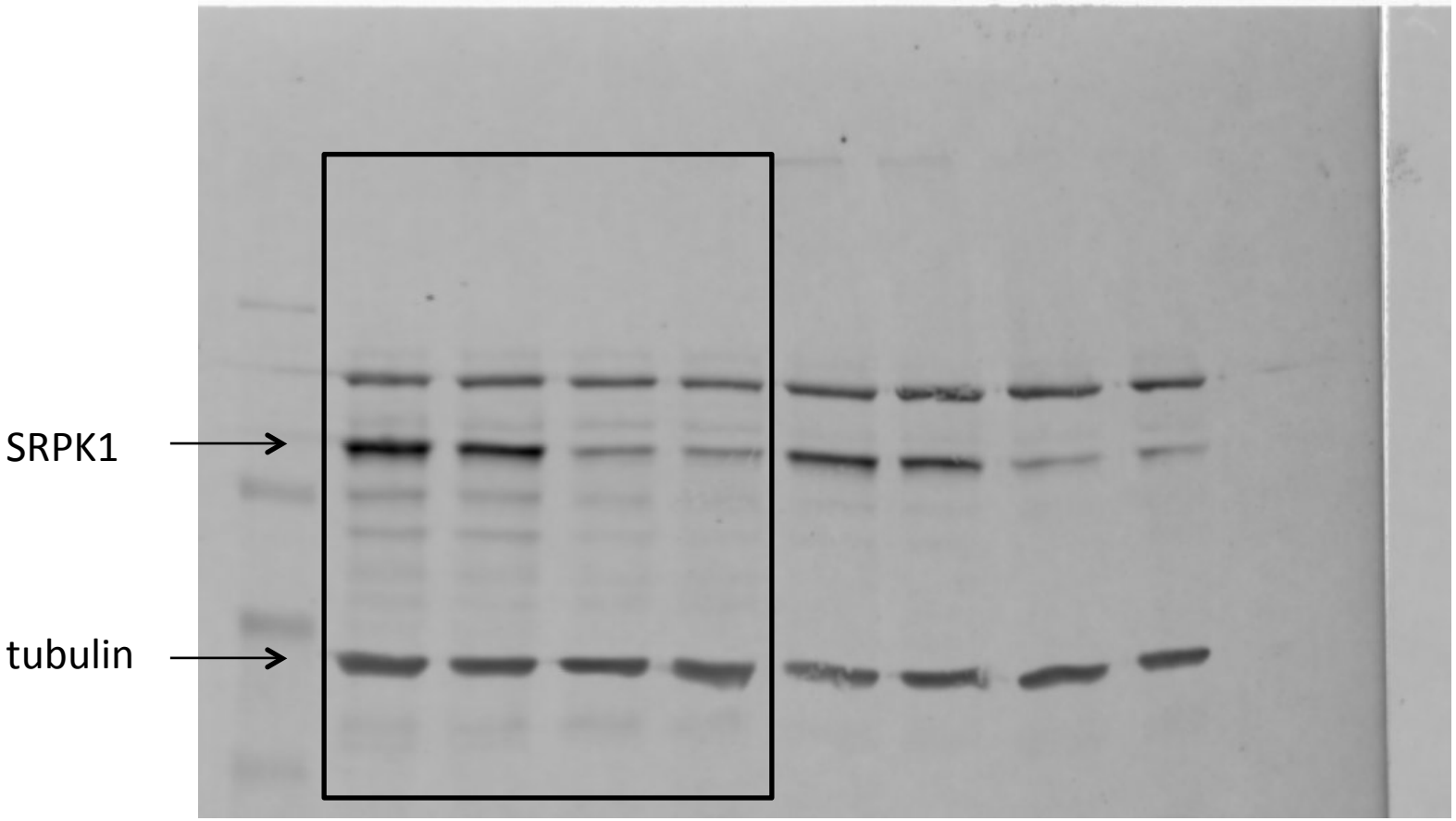
Full unedited gel for Supplemental Figure 9A



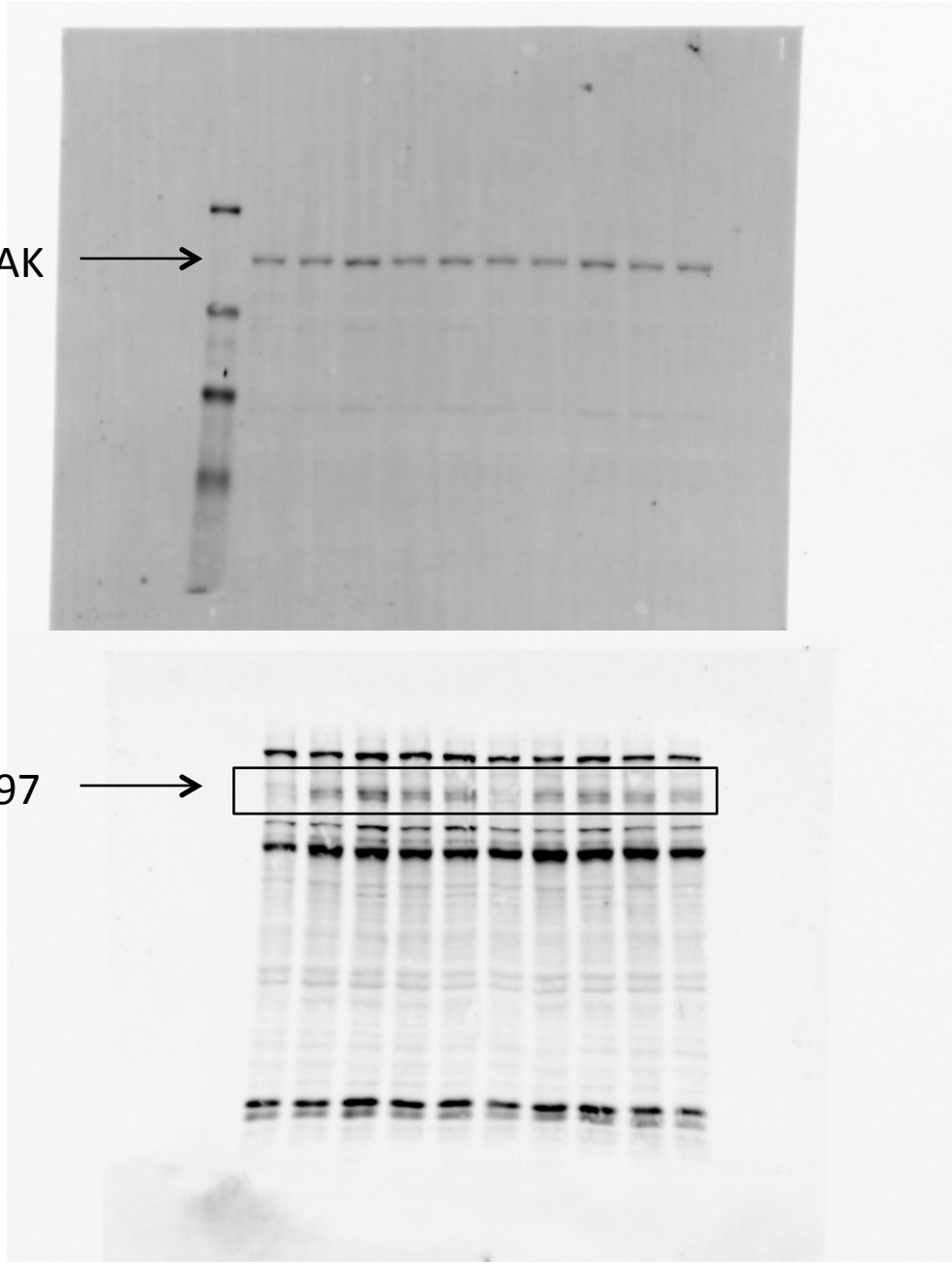
Full unedited gels for Supplemental Figure 11D



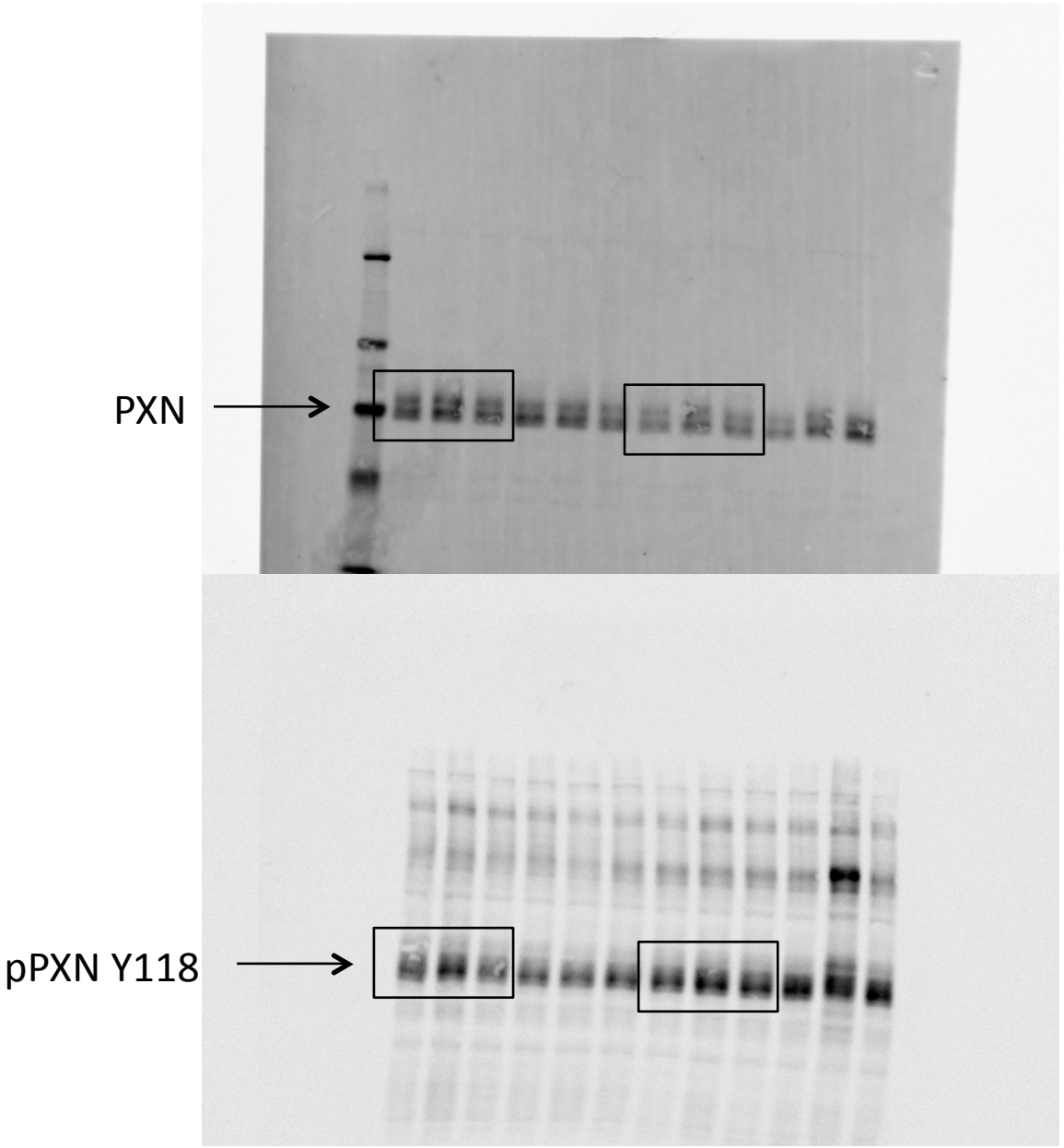
Full unedited gels for Supplemental Figure 14A



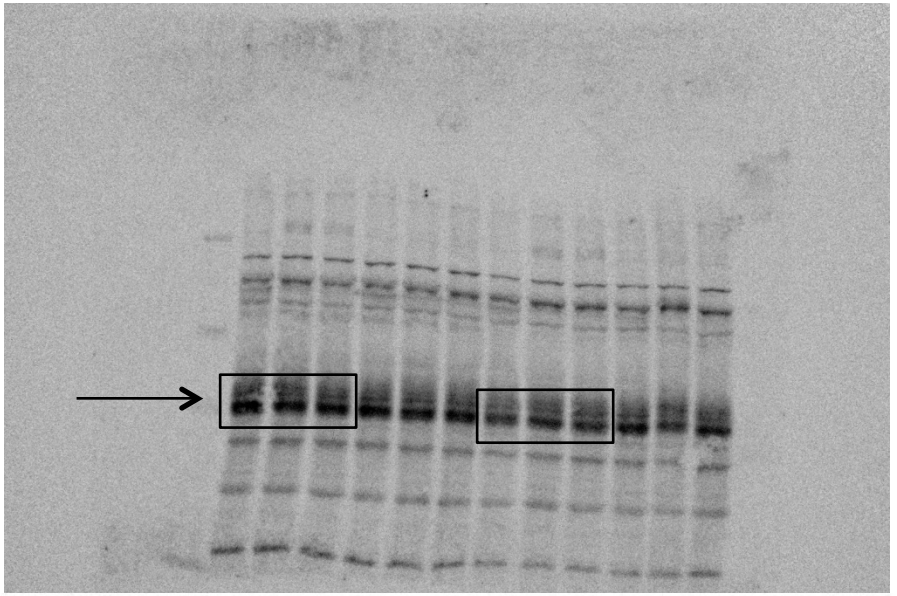
Full unedited gels for Supplemental Figure 15A

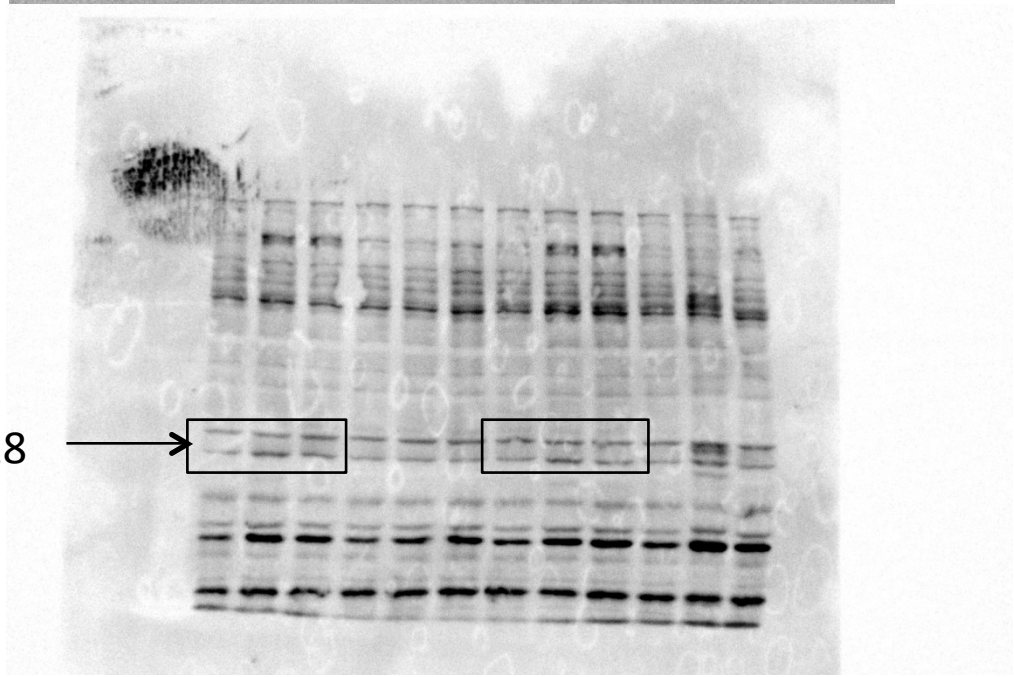


Full unedited gels for Supplemental Figure 15B



Supplemental Figure 15B





Supplemental Figure 15B