

Supplementary Figures & Tables

Supplementary Figure S1

Schematic presentation of *CSPP1* gene structure and gene products and *CSPP1* specific probes and reagents.

(A) Schematic alignment of *CSPP1* on chromosome 8q13.1-2 and reported mRNAs of splice isoforms, gene expression array probes (cDNA arrays: IMAGE.x, Agilent platform A_x), shRNA target regions (shRNA_x), and transcriptional control elements (data from UCSC GoldenPath genome browser). Regions in mRNAs comprising coding regions for a-CSPP-L (green) and a-CSPP/CSPP-L (red) epitopes are indicated. (B) Schematic alignment of CSPP and CSPP-L protein structure and annotation of functional domains. CSPP-L specific regions are indicated in grey. Predicted coiled-coil regions are depicted as bold bars.

Supplementary Figure S2

Comparison of immunohistochemical staining of a-CSPP-L and a-CSPP/CSPP-L antibodies in normal human mammary gland tissue.

Immunohistochemical staining of normal mammary gland from mastectomy tissue with CSPP-L specific (a-CSPP-L) or the common CSPP and CSPP-L targeting a-CSPP/CSPP-L antibody shows prominent staining of breast epithelial cells. Only a-CSPP/CSPP-L shows nuclear staining of lumen lining but not basal membrane lining epithelial cells.

Supplementary Figure S3

CSPP-L expression and occurrence of centrosome amplification in selected breast cancer cell lines

(A) Immunoblots of total cell lysates of breast cancer cell lines for CSPP-L and γ -tubulin: MDA-MB-231 (basal, BaB), HCC38 (basal, BaB), HCC1937 (basal, BaA), ZR-75-1 (luminal, LuA), MCF7 (luminal, LuA). Basal type cell lines show higher expression of CSPP-L (α -CSPP-L) than luminal A type cell lines relative to the loading control protein γ -tubulin (α -GTU-88). (B) Immunoblots of total cell lysates of luminal A (MCF7, ZR-75-1) and luminal B type breast cancer cell lines (UACC-812, BT-474) for CSPP-L and γ -tubulin indicate no prominent differences in CSPP-L expression. Exposure time for CSPP-L detection was increased compared to (A). (C) Quantification of centrosome amplification in the six breast cancer cell lines used in Figure 2A. (# of γ -tubulin stained

centrosomal foci per cell, at least 150 cells were scored per cell line; error bars depict standard error of the mean). Basal type breast cancer cell lines show higher frequency of centrosome amplification than luminal type cell lines.

Supplementary Figure S4

Correlation of *CSPP1* gene copy number alterations and mean *CSPP1* mRNA expression in Oslo0/Ull and Oslo1/MicMa cohorts of primary operable breast cancer.

(A) Extended analysis of PAM50 subtype specific mean *CSPP1* mRNA expression in Oslo0/Ull cohort (n=80) as shown in Fig4I, here including tabular overview of paired statistical significance analysis (left panel). PAM50 subtype specific correlation plots of *CSPP1* copy number and mean mRNA expression (right panels). Regression correlation values are indicated. (B) PAM50 subtype specific mean *CSPP1* mRNA expression in Oslo1/Ull cohort (n=115) with tabular overview of paired statistical significance analysis (left panel). PAM50 subtype specific correlation plots of *CSPP1* copy number and mean mRNA expression (right panels). Regression correlation values are indicated.

Supplementary Figure S5

Validation of differential gene expression in nuclear CSPP1 positive and negative basal-like breast carcinomas.

(A) Hierarchical cluster analysis of basal-like breast carcinomas in the Oslo0/Ull cohort used for identification of differentially expressed genes between nuclear CSPP1 positive and negative cases. (B) Identical to Fig5B,C. Hierarchical cluster analysis of all basal-like breast carcinomas in the Oslo0/Ull cohort on basis of the identified differentially expressed genes. Scoring results are indicated for cases with immunohistochemical CSPP1 staining. Cases used for SAM are indicated with # symbol (upper panel). Box plots of PAM50 breast cancer subtype centroid correlation coefficients of “nuclear CSPP1 negative” and “nuclear CSPP1 positive” cluster group biopsies (lower panel). P-values indicate statistically significant differences in median correlation coefficients (t-test or (*) Mann-Whitney). (C) Analysis of basal-like breast cancer biopsies of the Oslo1/MicMa cohort (n=26) as described in (B).

Supplementary Figure S6

PAM50 subtype specific mean mRNA expression of *CSPP1* and genes of the 8-gene signature in METABRIC Discovery and Validation cohorts

Whisker-hair plots of PAM50 subtype specific mean mRNA expression of indicated genes. *CSPP1* expression panel is identical to Fig6A. Paired statistical significance analysis for PAM50 subtype specific expression differences are listed in Supplementary table 3.

Supplementary Figure S7

PAM50 subtype specific mean mRNA expression of *CSPP1*, *ESR*, *EGFR*, *ERBB2*, *PGR*, and *MKI67* in METABRIC Discovery and Validation cohorts

Whisker-hair plots of PAM50 subtype specific mean mRNA expression of indicated genes. *CSPP1* expression panel is identical to Fig6A. Paired statistical significance analysis for PAM50 subtype specific expression differences are listed in Supplementary table 3.

Supplementary Figure S8

Effects of serum starvation and β -estradiol stimulation on the expression of CSPP1 proteins in MCF7 cells.

(A) Immunoblots of total cell lysates of MCF7 breast cancer cells grown in phenolred containing DMEM supplemented with 0.01 mg/ml human recombinant insulin and 10% fetal calf serum (FCS) and stimulated for 24h with indicated doses of β -estradiol (E2). Expression of CSPP-L (α -CSPP-L) increases relative to the loading control protein γ -tubulin in an E2 dependent manner, as monitored by increased phosphorylation of the estrogen receptor alpha (α -P-Ser104/106-ER α). (B-D) MCF7 breast cancer cells grown in phenolred free DMEM medium using different additives. Cells were grown to 60% confluence in 10% FCS supplemented DMEM. The medium was then changed to DMEM + 10% FCS for 48h (control), DMEM + 0.1% FCS (serum starved 48h), DMEM + 0.1% FCS for 24h followed by re-stimulation with DMEM + 10% FCS + 200nM E2, or DMEM + 10% FCS + 200nM E2 for 48h. CSPP1 protein expression was determined by immunoblotting for CSPP-L relative to γ -tubulin (B), immunofluorescence microscopy for CSPP-L (green), γ -tubulin (red), and DNA (blue) (C). Cell cycle phase distribution and total CSPP1 protein expression was measured by FACS analysis in parallel samples by staining for α -CSPP/CSPP-L and DNA content (Hoechst) (D). Serum starved MCF7 cells

Nuclear CSPP1 expression in breast cancer

arrested in G₀/G₁-phase (DNA histogram in (D)) and showed flattened morphology (data not shown). Re-stimulation of serum starved MCF7 cells with serum and E2 induced release from cell cycle arrest (increased S-phase fraction in DNA histogram, (D)) and cell rounding and cell-cell contact formation (data not shown). Total CSPP1 protein expression (α-CSPP/CSPP-L staining intensity in FACS analysis) was cell cycle phase independent and only slightly increased in response to E2 stimulation (dot blots, (D)). In contrast, CSPP-L expression is sensitive to serum starvation and induced upon E2 stimulation (A and B). CSPP-L partially co-localized with γ-tubulin at the centrosome and concentrated in centrosome surrounding granules reminiscent of centriolar satellites (C).

Supplementary Table 1

Cross tabulation of nuclear CSPP1 expressing cancer cell frequency and cytoplasmic CSPP1 expression.

CSPP1 expression varied both within and between tumor samples in the Oslo0/UII cohort. The frequency of tumor cells with nuclear staining were grouped into four categories (0=0-2%; 1=2-50%, 2=50%-75% and 3=>75%). Cytoplasmic staining of tumor cells was scored only by intensity (0=absent, 1=weak, 2=medium, 3=strong).

Supplementary Table 2

Clinico-pathological features and the CSPP1 expression groups in the Oslo0/UII cohort.

p-values calculated by Fisher's exact test.

Supplementary Table 3

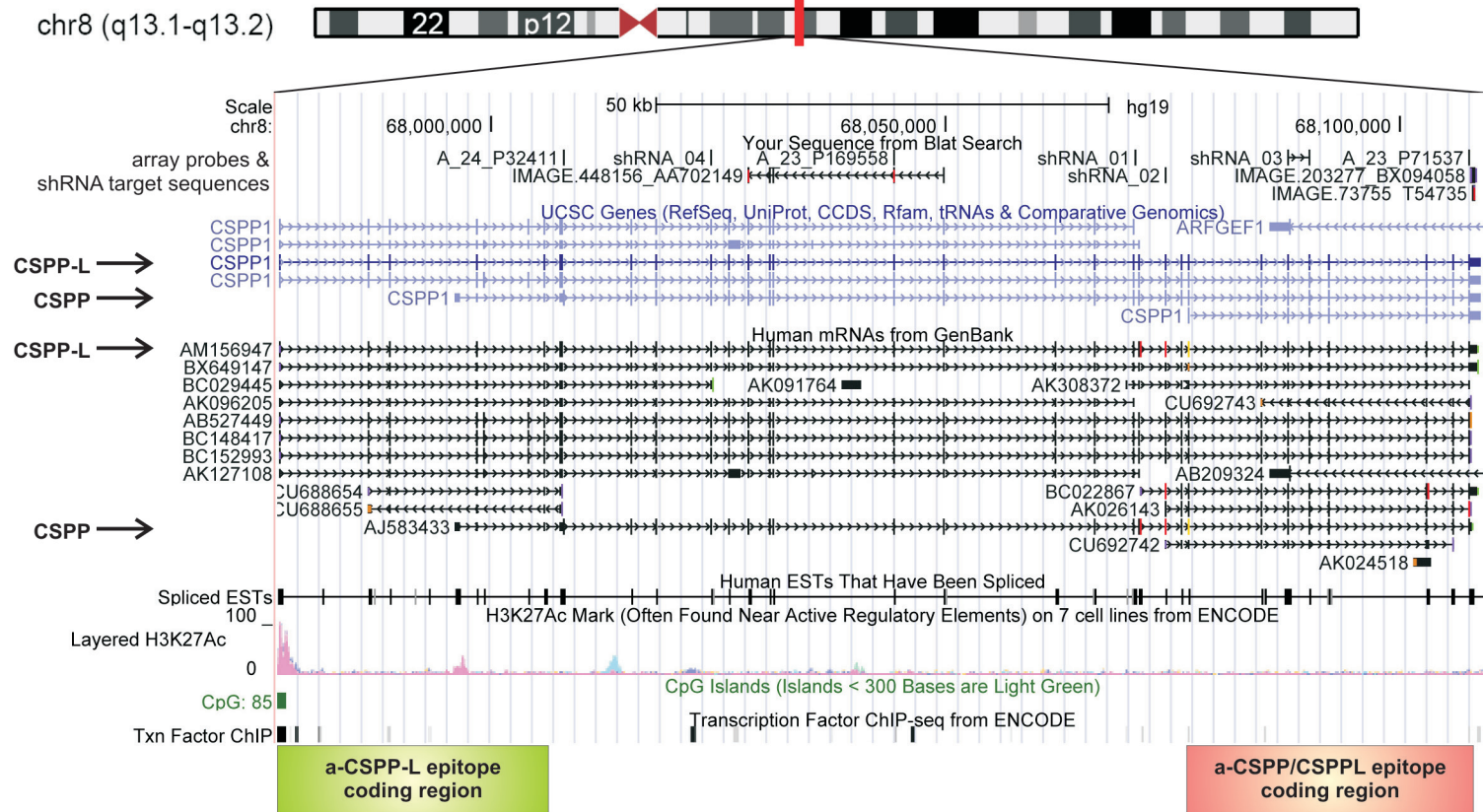
Cross tabulations of paired statistical significance analysis for subtype specific expression of selected genes in METABRIC Discovery and Validation cohorts

p-values <0.05 are considered as statistically significant and highlighted in bold

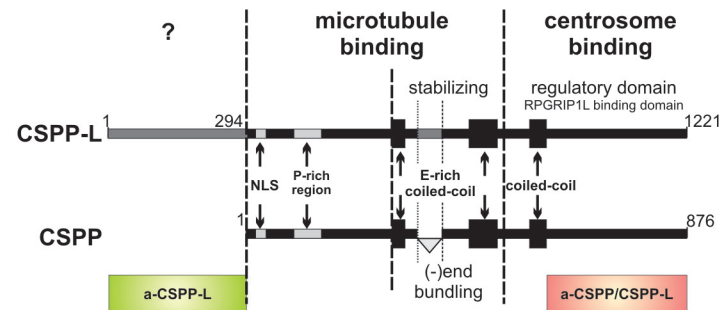
Supplementary Figure S1

A

genome, gene expression array probes, shRNA target sequences and mRNA sequences alignment



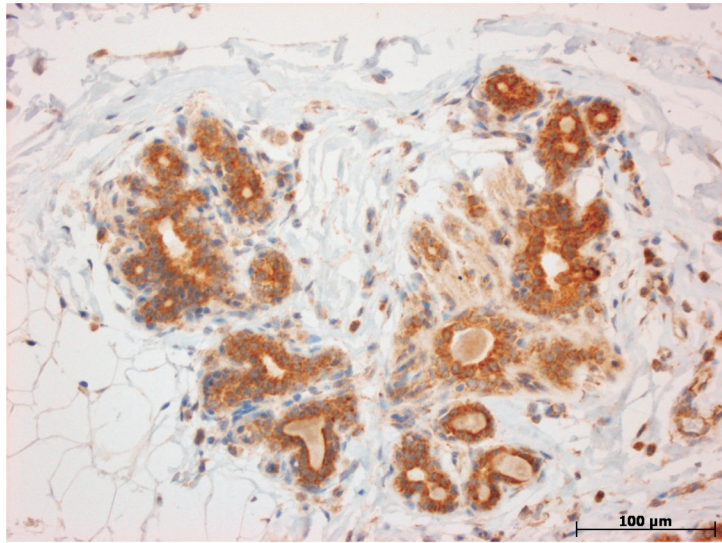
B



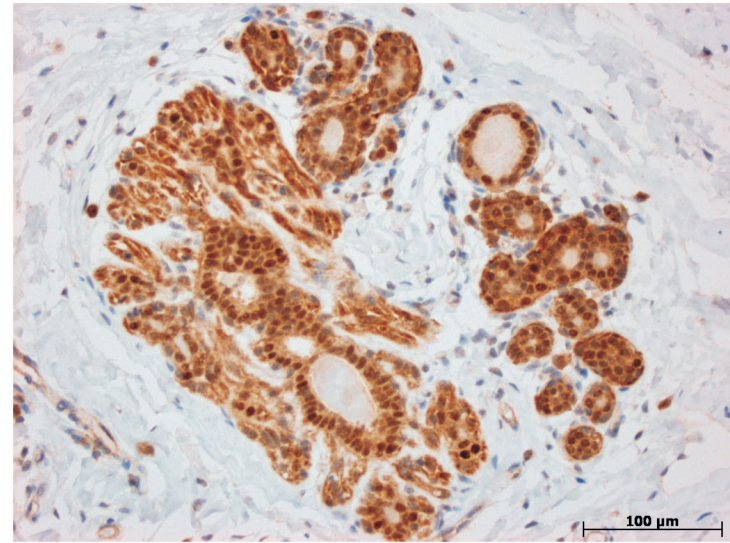
Supplementary Figure S2

a-CSPP-L and a-CSPP/CSPP-L staining pattern in paired biopsies of normal breast tissue

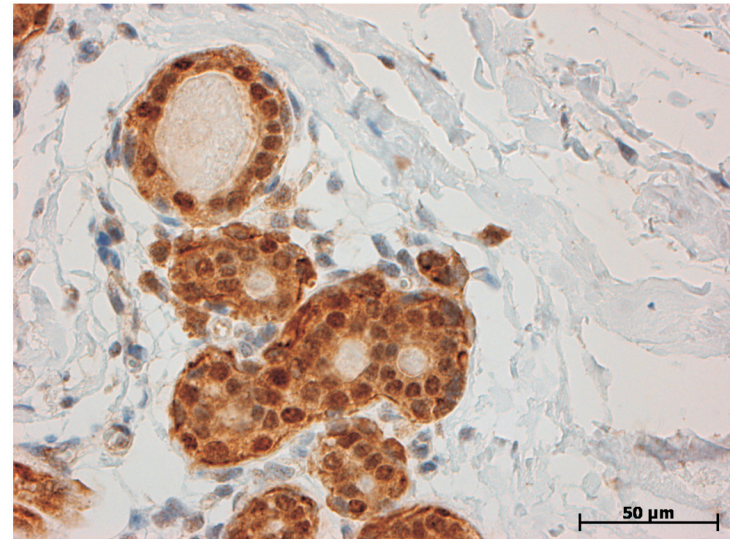
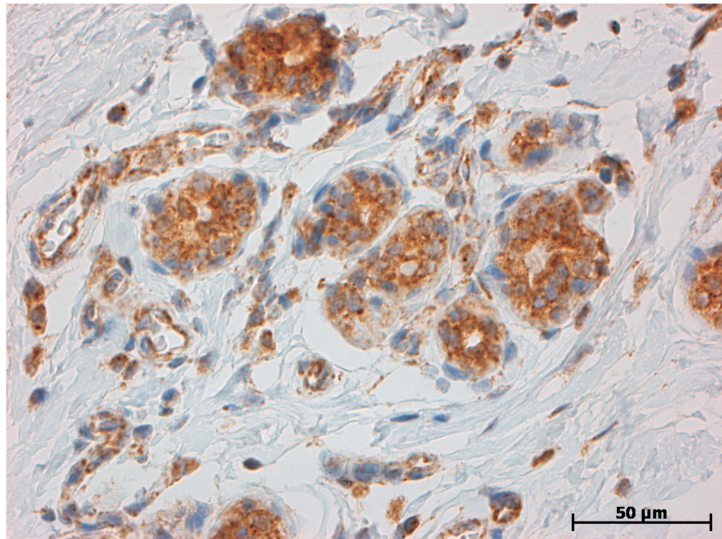
a-CSPP-L



a-CSPP/CSPP-L

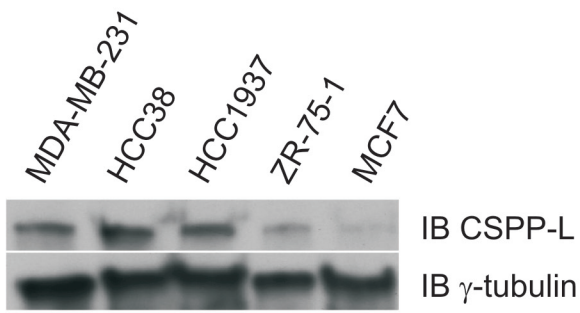


40x



Supplementary Figure S3

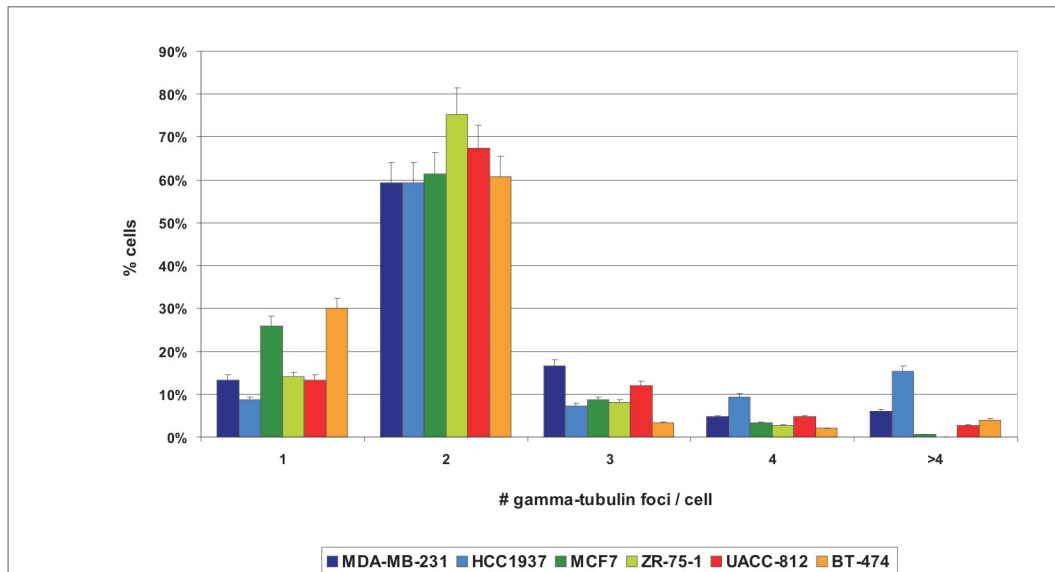
A



B



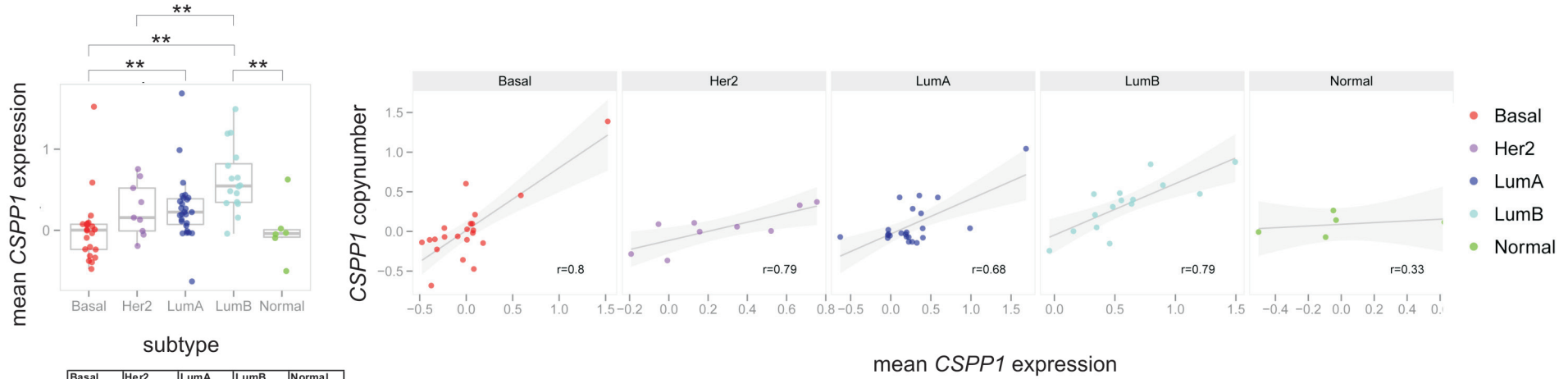
C



Supplementary Figure S4

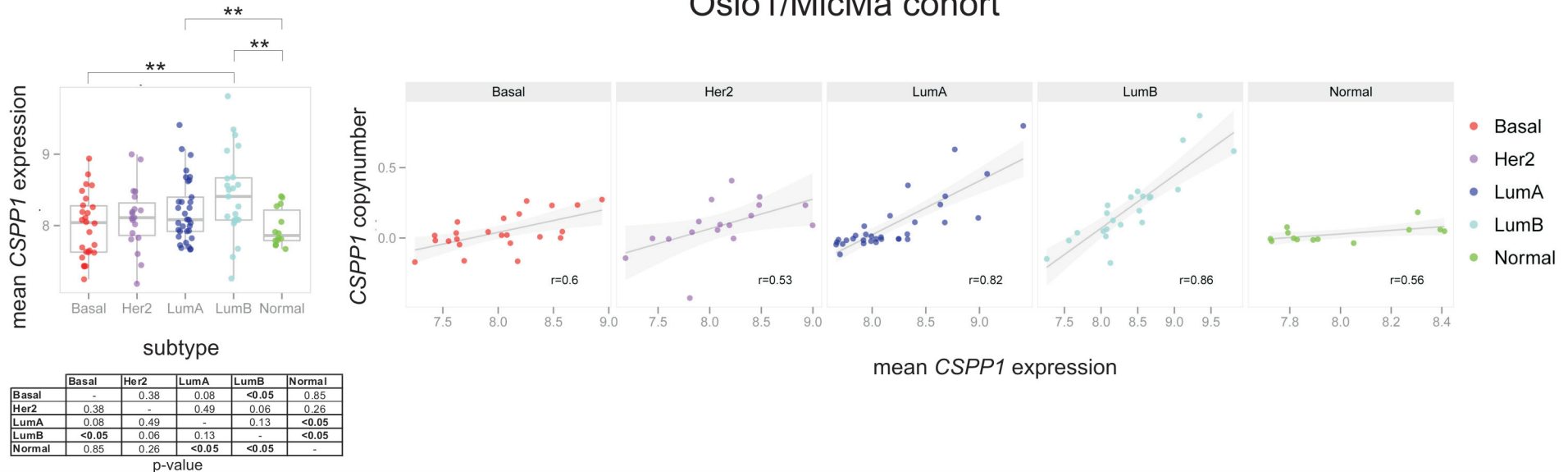
A

Oslo0/Ull cohort



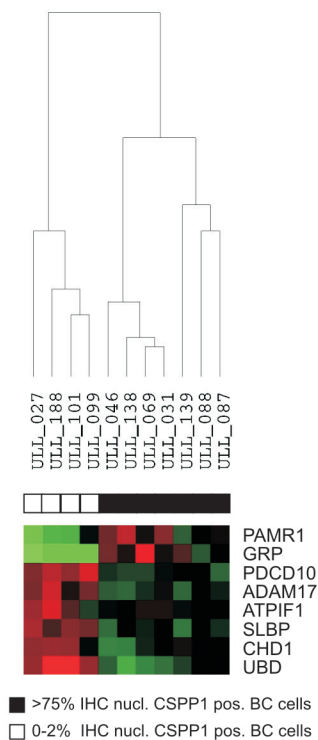
B

Oslo1/MicMa cohort

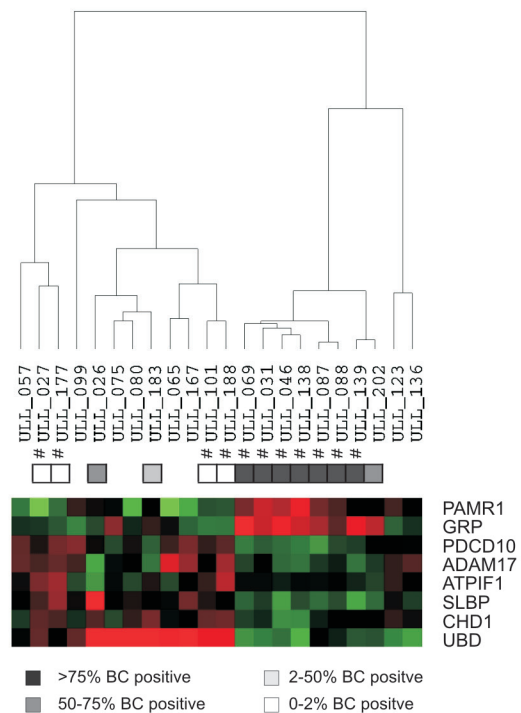


Supplementary Figure S5

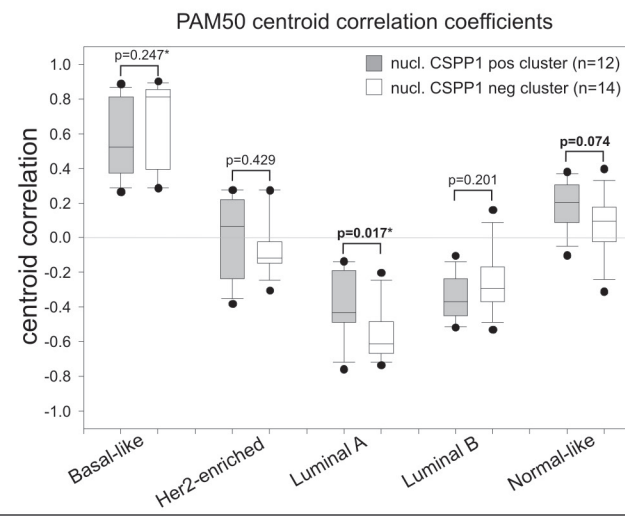
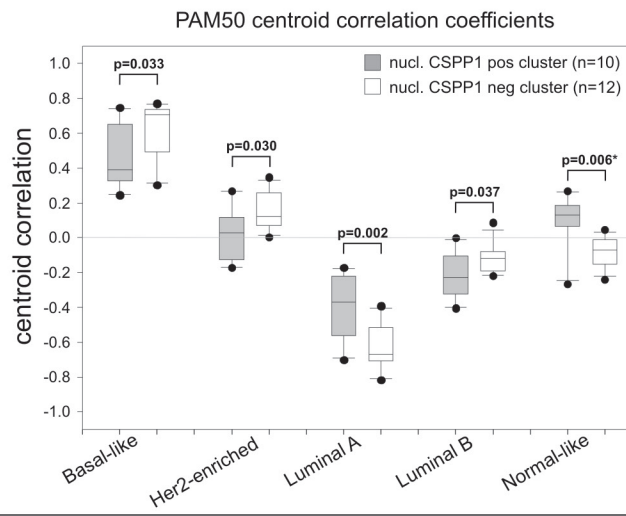
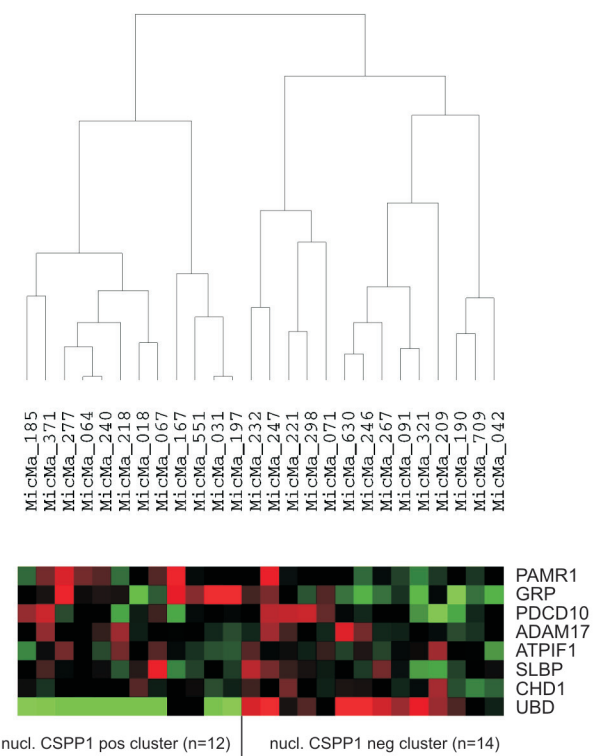
A

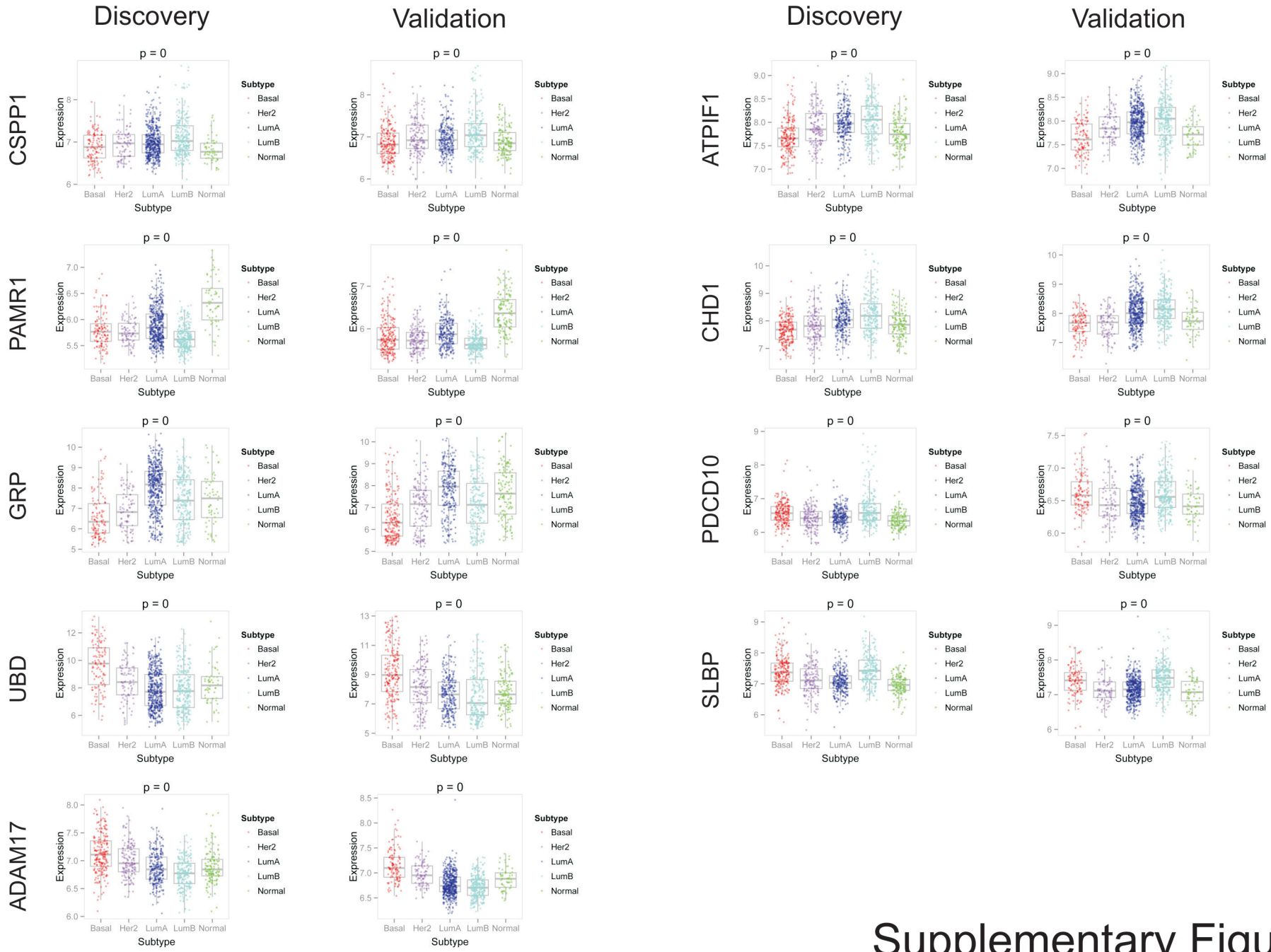


B

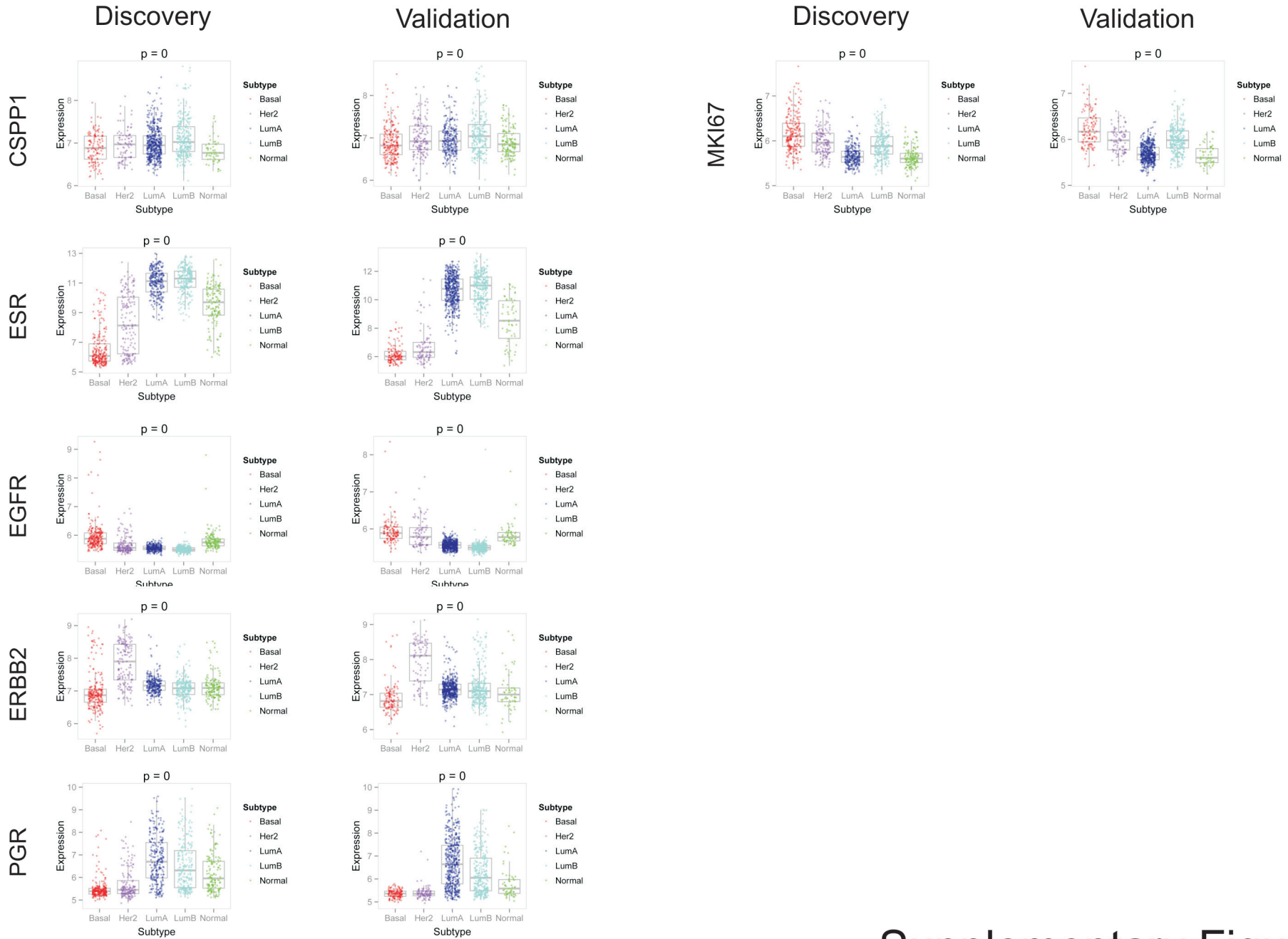


C





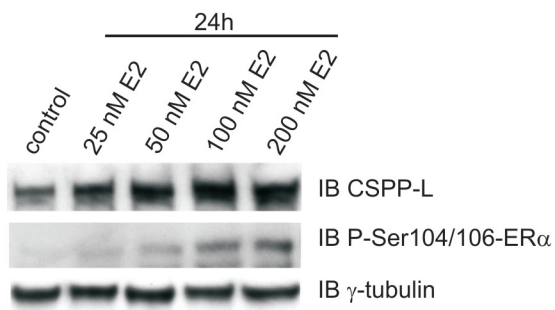
Supplementary Figure S6



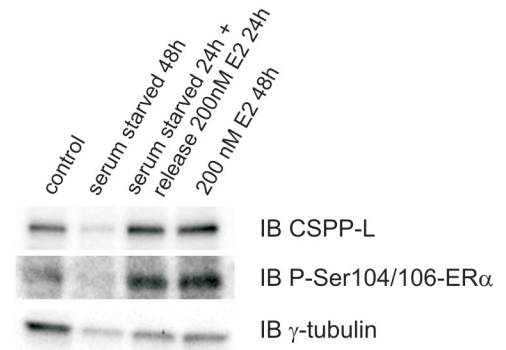
MKI67

Supplementary Figure S8

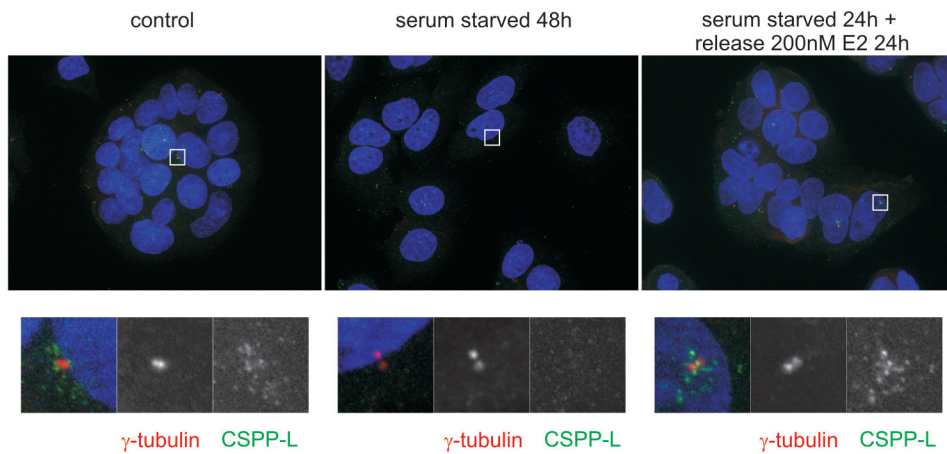
A



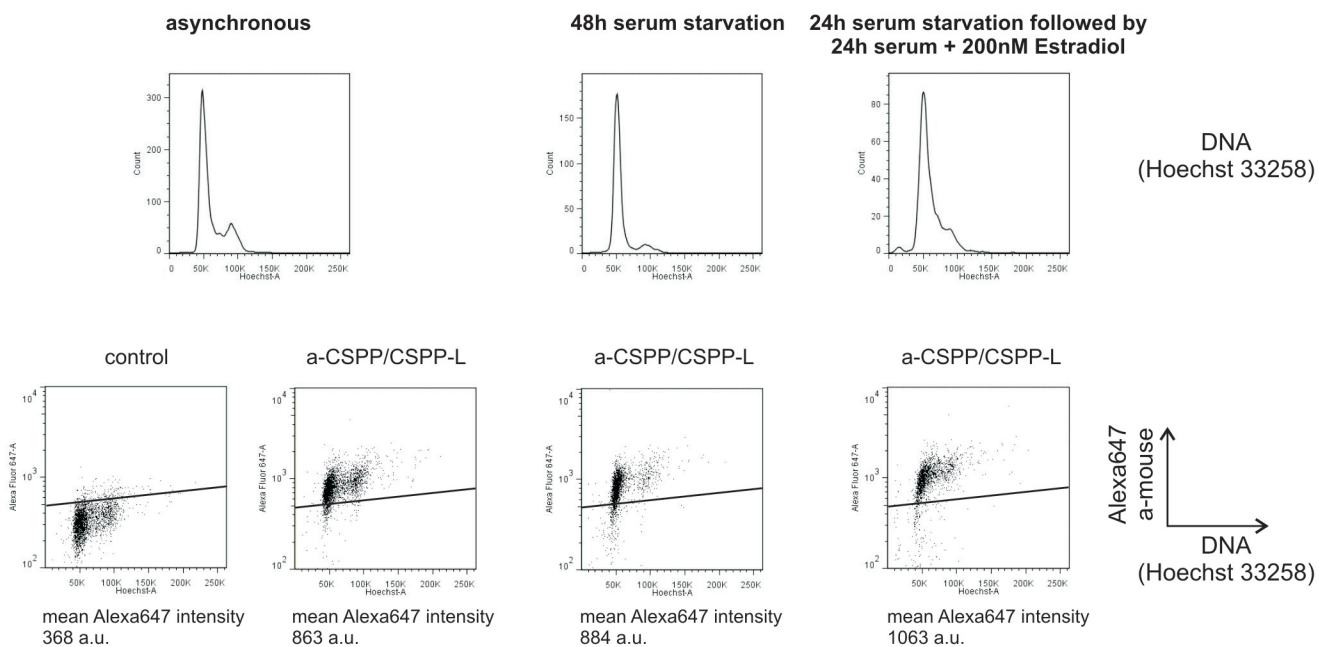
B



C



D



Supplementary table 1

Cross tabulation of nuclear CSPP1 expressing cancer cell frequency and cytoplasmic CSPP1 expression

CSPP1 nucleus (freq. Nucl. Pos.)	CSPP1 cytoplasm			
	Negative	Weak	Medium	Strong
0 (<2%)	4 (40%)	4 (40%)	1 (10%)	1 (10%)
1 (2-50%)	3 (33.3%)	2 (22.2%)	3 (33.3%)	1 (11.1%)
2 (50-75%)	4 (22.2%)	6 (33.3%)	5 (27.8%)	3 (16.7%)
3 (>75%)	10 (10.2%)	28 (28.6%)	40 (40.8%)	20 (20.4%)

p= 0.177

Supplementary table 2

Clinico-pathological features and the nuclear CSPP1 protein groups

	CSPP1 nuclear staining				p value
	0	1	2	3	
Histologi					
Ductal	9 (10.5%)	9 (10.5%)	14 (16.3%)	54 (62.8%)	<i>p</i> = 0.049
Lobular	0	0	2 (5.7%)	33 (94.3%)	
Other	1 (10%)	0	1 (10%)	8 (80%)	
Unknown	0	0	1 (25%)	3 (75%)	
Stage					
1	3 (12%)	1 (4%)	2 (8%)	19 (76%)	<i>p</i> =0.710
2	2 (3.7%)	2 (3.7%)	8 (14.8%)	42 (77.8%)	
3	2 (6.1%)	4 (12.1%)	5 (15.2%)	22 (66.7%)	
4	0	0	0	1 (100%)	
Unknown	3 (13.6%)	2 (9.1%)	3 (13.6%)	14 (63.6%)	
Grade					
1	0	0	2 (25%)	6 (75%)	<i>p</i> =0.552
2	5 (5.2%)	8 (8.3%)	12 (12.5%)	71 (74%)	
3	5 (17.2%)	1 (3.4%)	4 (13.8%)	19 (65.5%)	
Unknown	0	0	0	2 (100%)	
ER					
Negative	6 (12.8%)	1 (2.1%)	6 (12.8%)	34 (72.3%)	<i>p</i> =0.216
Positive	3 (4.1%)	6 (8.2%)	11 (15.1%)	53 (72.6%)	
PgR					
Negative	4 (8.5%)	4 (8.5%)	5 (10.6%)	34 (72.3%)	<i>p</i> =0.819
Positive	5 (6%)	5 (6%)	12 (14.5%)	61 (73.5%)	
TP53 mutation					
wt/silent	4 (4.2%)	8 (8.4%)	16 (16.8%)	67 (70.5%)	<i>p</i> =0.142
mutation	5 (16.7%)	1 (3.3%)	2 (6.7%)	22 (73.3%)	
unknown	1 (10%)	0	0	9 (90%)	
GATA3 mutation					
wt/silent	5 (10.4%)	2 (4.2%)	9 (18.8%)	32 (66.7%)	<i>p</i> =0.265
mutation	0	1 (33.3%)	0	2 (66.7%)	
unknown	4 (5.1%)	6 (7.6%)	8 (10.1%)	61 (77.2%)	
Expression group					
Luminal A	1 (5.6%)	2 (11%)	3 (16.7%)	12 (66.7%)	<i>p</i> =0.513
Luminal B	0	2 (22.2%)	1 (11.1%)	6 (66.7%)	
HER2-enriched	0	0	3 (50%)	3 (50%)	
Basal-like	4 (28.6%)	1 (7.1%)	2 (14.3%)	7 (50%)	
Normal-like	0	0	1 (20%)	4 (80%)	
Lymphnode					
Negative	6 (9.2%)	3 (4.6%)	6 (9.2%)	50 (76.9%)	<i>p</i> =0.319
Positive	2 (4%)	3 (6%)	10 (20%)	35 (70%)	
Unknown	2 (10%)	3 (15%)	2 (10%)	13 (65%)	

