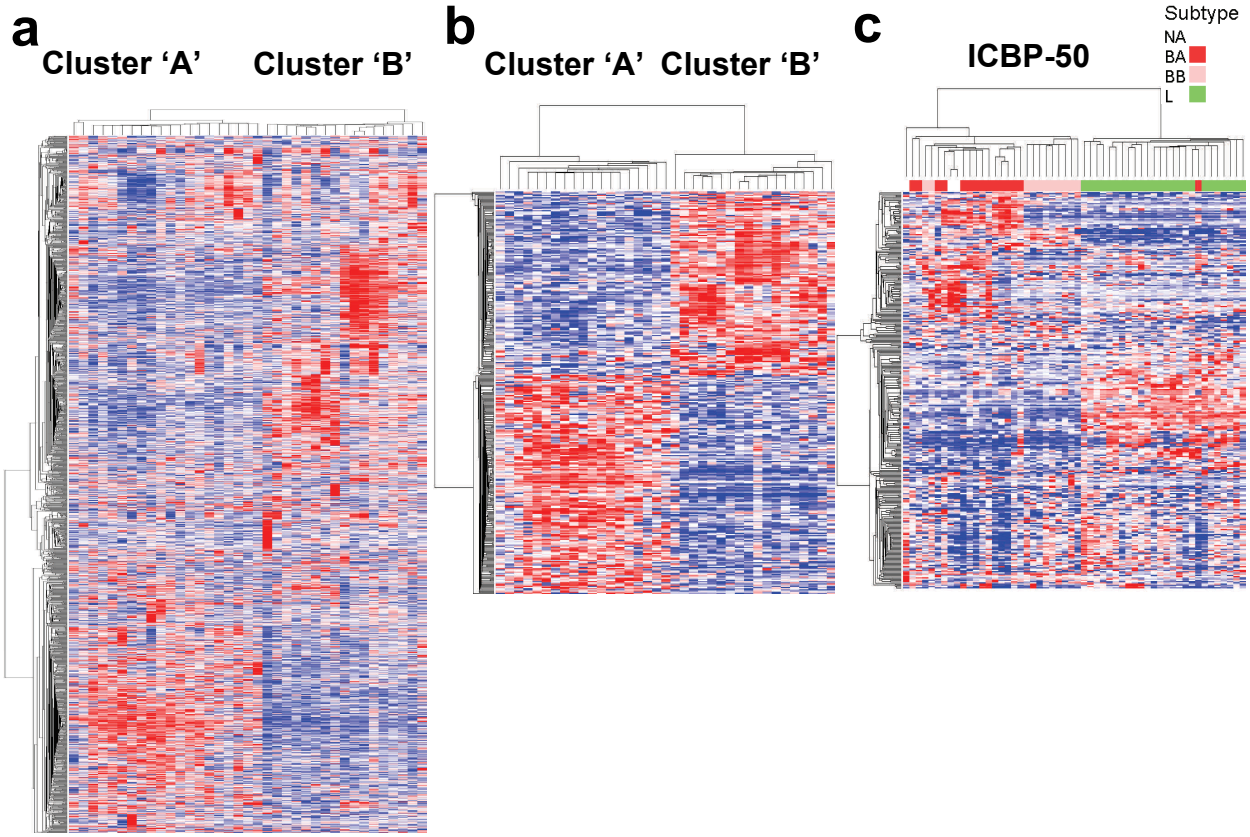
Supplementary Information for:

Digital RNA quantification in tumor cell-sparse breast cancers following neoadjuvant chemotherapy identifies DUSP4 deficiency as a mechanism of drug resistance in basal-like breast carcinoma

Justin M. Balko[1], Rebecca S. Cook[2,5], David Vaught[2], María G. Kuba[3], Todd W. Miller[2,5], Neil E. Bhola[1], Melinda E. Sanders[3,5], Nara M. Granja-Ingram[3], J. Joshua Smith[4], Ingrid M. Meszoely[4,5], Janine Salter[6,7], Mitch Dowsett[6,7], Katherine Stemke-Hale[9], Ana M. González-Angulo[8,9], Gordon B. Mills[8,9], Joseph A. Pinto[10], Henry Gómez[11], Carlos L. Arteaga[1,2,5]

Departments of Medicine[1], Cancer Biology[2], Pathology[3], and Surgery[4]; Breast Cancer Research Program[5], Vanderbilt-Ingram Comprehensive Cancer Center; Vanderbilt University, Nashville, TN; [6]Breakthrough Breast Cancer Centre, Institute of Cancer Research, [7]Academic Department of Biochemistry, Royal Marsden Hospital, London, UK; Departments of Breast Medical Oncology[8] and Systems Biology[9], University of Texas, M.D. Anderson Cancer Center, Houston, TX.; [10] Division de Investigacion, Oncosalud, Lima, Perú. [11]Instituto Nacional de Enfermedades Neoplásicas, Lima, Perú.

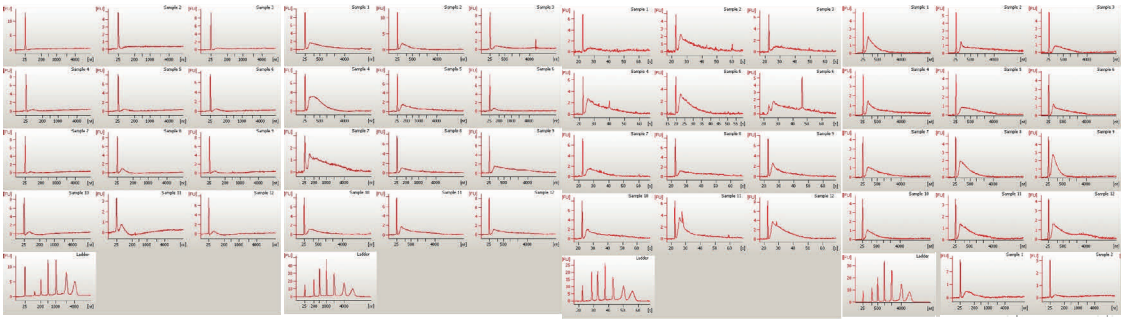
Supplementary Figure 1: Metagene and gene signature selection



Supplementary Figure 1: Metagene and gene signature selection. **a)** Hierarchical clustering of pre-treatment gene expression data from ER-negative tumors of patients who did not achieve pCR following NAC in study EORTC10994. **b)** Hierarchical clustering of the EORTC10994 data using the 354 most differentially expressed probesets (P<0.01) between the two most prominent clusters from (a) (CLUSTER signature). **c)** Hierarchical clustering of the ICBP50 panel of breast cancer cell lines annotated by molecular subtype as reported by Neve *et al.* [1] using the 354-probeset CLUSTER signature. Annotation of subtypes: NA, not available; BA, basal-like group A; BB, basal-like group B; L, luminal.

Supplementary Figure 2: Quality control and normalization of the Nanostring nCounter data

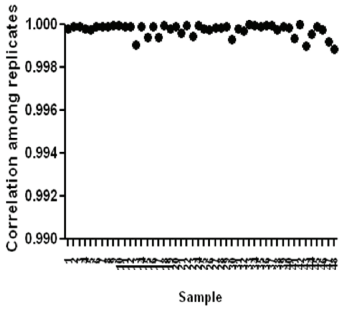
a **b** **c** **d**



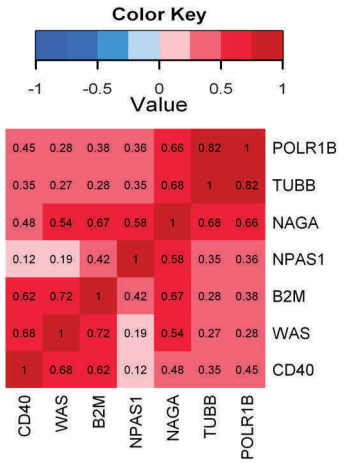
e

Panel	Sample	[Intact RNA] (ng/ul)	Panel	Sample	[Intact RNA] (ng/ul)
a	1	Blank	c	1	39
a	2	21	c	2	26
a	3	15	c	3	59
a	4	6	c	4	54
a	5	8	c	5	49
a	6	11	c	6	34
a	7	5	c	7	32
a	8	18	c	8	42
a	9	6	c	9	44
a	10	6	c	10	54
a	11	13	c	11	70
a	12	10	c	12	26
b	1	21	d	1	29
b	2	12	d	2	13
b	3	49	d	3	24
b	4	24	d	4	21
b	5	7	d	5	23
b	6	28	d	6	23
b	7	17	d	7	31
b	8	32	d	8	27
b	9	17	d	9	16
b	10	26	d	10	26
b	11	17	d	11	34
b	12	10	d	12	27
			d	1	11
			d	2	29

f

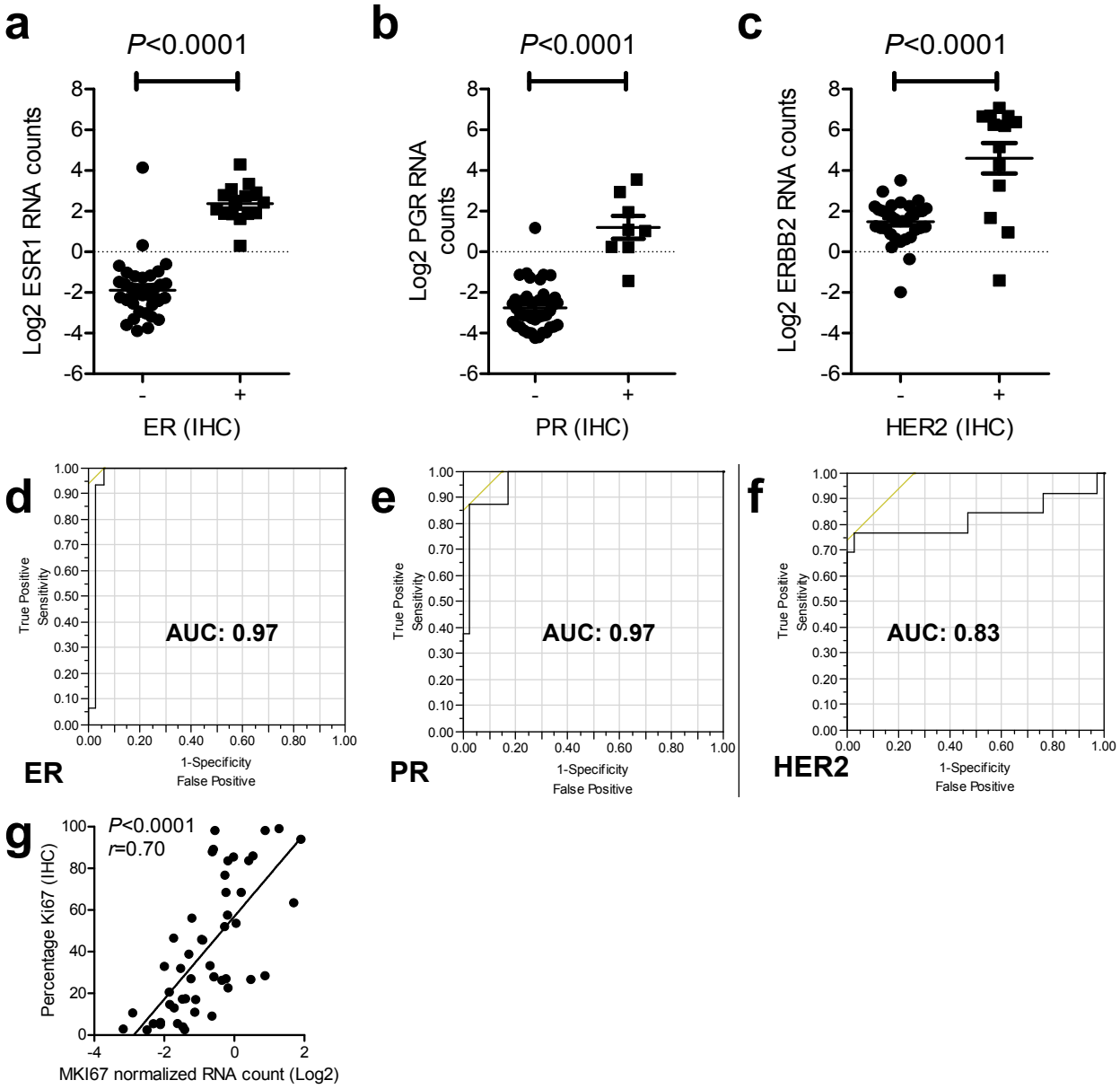


g



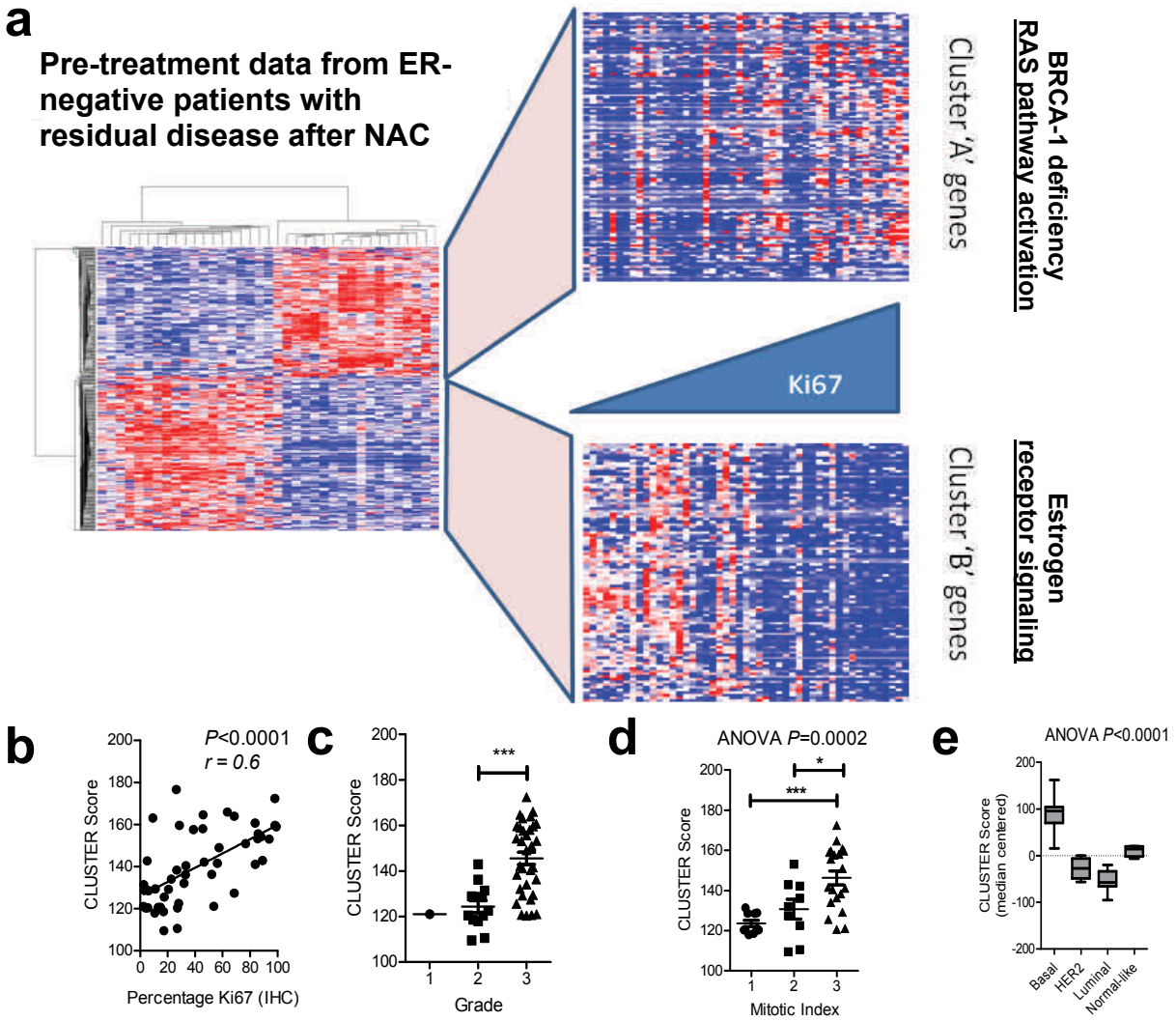
Supplementary Figure 2: Quality control and normalization of the Nanostring nCounter data. **a-d)** Individual plots of RNA fragment counts of the 49 samples. Plots were generated using an Agilent Bioanalyzer. **e)** RNA quantification of fragments >100 nucleotides by Agilent Bioanalyzer smear analysis was performed in order to estimate approximate loading conditions for Nanostring analysis. **f)** Plot of the inter-replicate Nanostring correlations from the 49 patient samples **g)** Correlation matrix for the raw RNA counts for the 7 pre-selected normalization genes across all 49 samples. Inter-gene correlations ranged from 0.12 to 0.82 for all gene-pairs.

Supplementary Figure 3: Digital quantification of gene expression in FFPE tumors generates robust measurements of univariate parameters



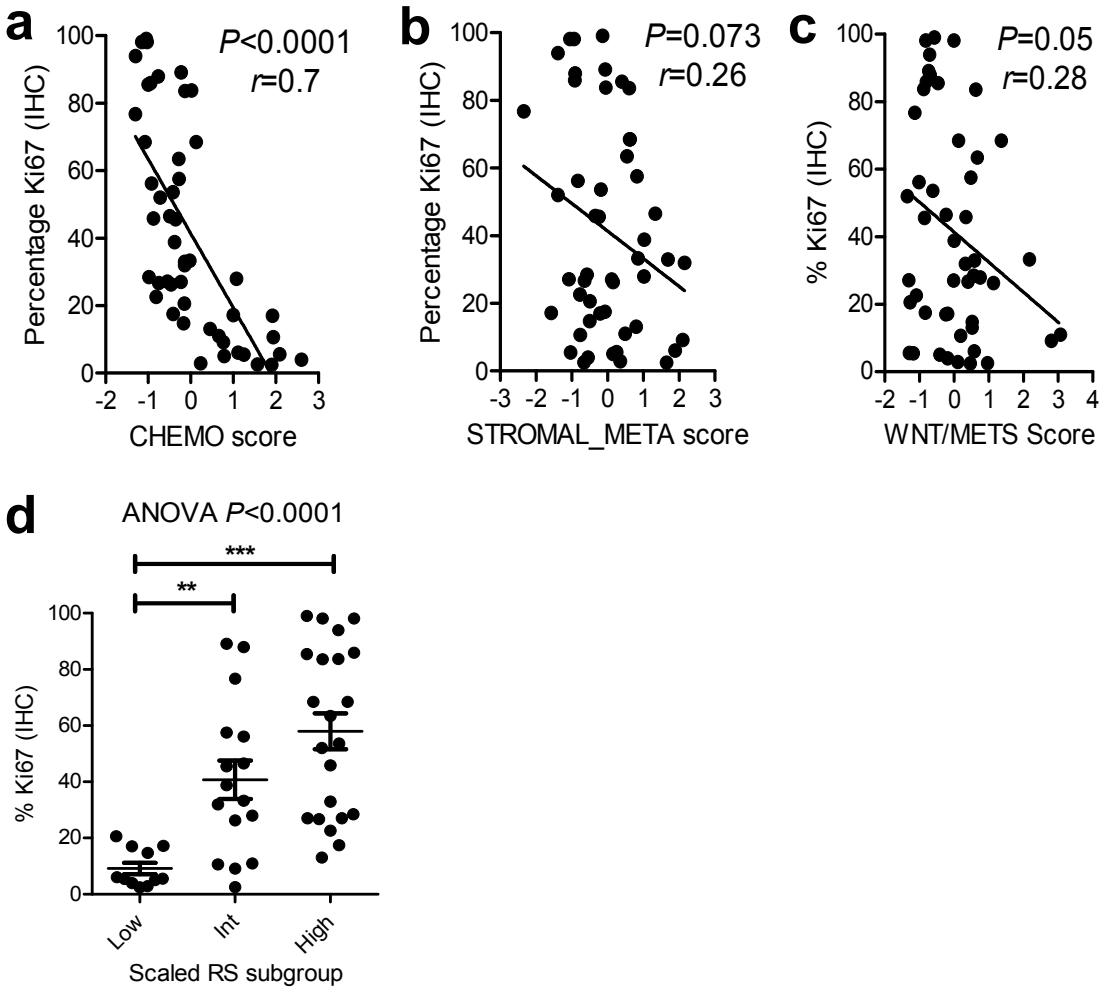
Supplementary Figure 3: Digital quantification of gene expression in FFPE tumors generates robust measurements of univariate parameters. **a-c)** Plots of log2 normalized Nanostring transcript counts according to the clinical ER, PR, and HER2 status. P-values represent the results of a two-sample t-test. **d-f)** Receiver-operator characteristic (ROC) based on a logistic regression model to predict binary (+/-) clinical receptor status using Nanostring gene expression for ER, PR, and HER2. AUC is calculated based on the area under the (1-specificity) - sensitivity curve. **g)** Correlation of normalized log2 RNA transcript counts for *MKI67* versus percentage of tumor cells staining positive for Ki67 by IHC.

Supplementary Figure 4: The CLUSTER gene signature encodes elements of proliferation, grade and prognosis.



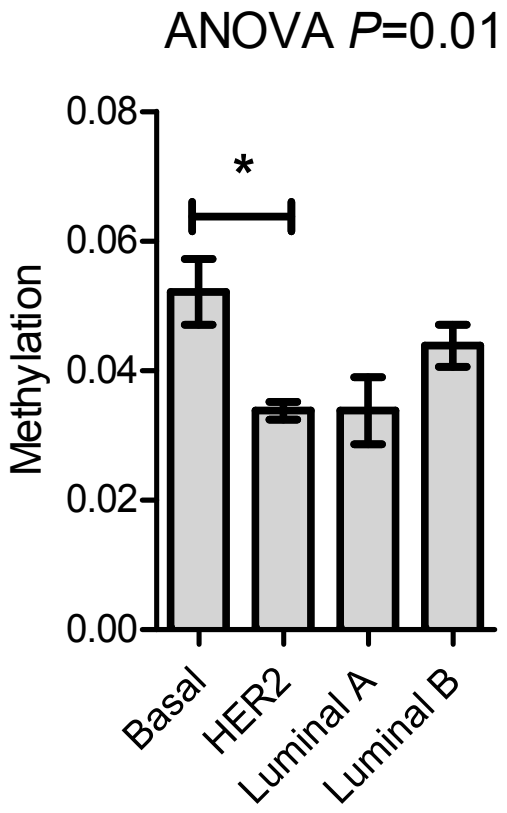
Supplementary Figure 4: The CLUSTER gene signature encodes elements of proliferation, grade and prognosis. **a)** Heatmap of the CLUSTER gene signature in 49 post-NAC residual tumors, as assayed by Nanostring. Samples are arranged according to increasing Ki67 staining by IHC. **b)** Correlation of a composite score for the CLUSTER signature versus the Ki67 score. **c)** Association of the composite score for the CLUSTER signature versus clinical grade in the post-treatment (surgical) specimen. A t-test was used to compare grade 2 and grade 3; *** $P<0.001$. **d)** Association of the composite score for the CLUSTER signature versus post-treatment (surgical) mitotic index. A Bonferonni post-hoc test was used to make multiple comparisons among the groups; * $P<0.05$, *** $P<0.001$. **e)** Association of the composite score for the CLUSTER signature versus molecular subtype as determined by hierarchical clustering (Fig. 1c)

Supplementary Figure 5: Predictive metagene signatures and associations with Ki67



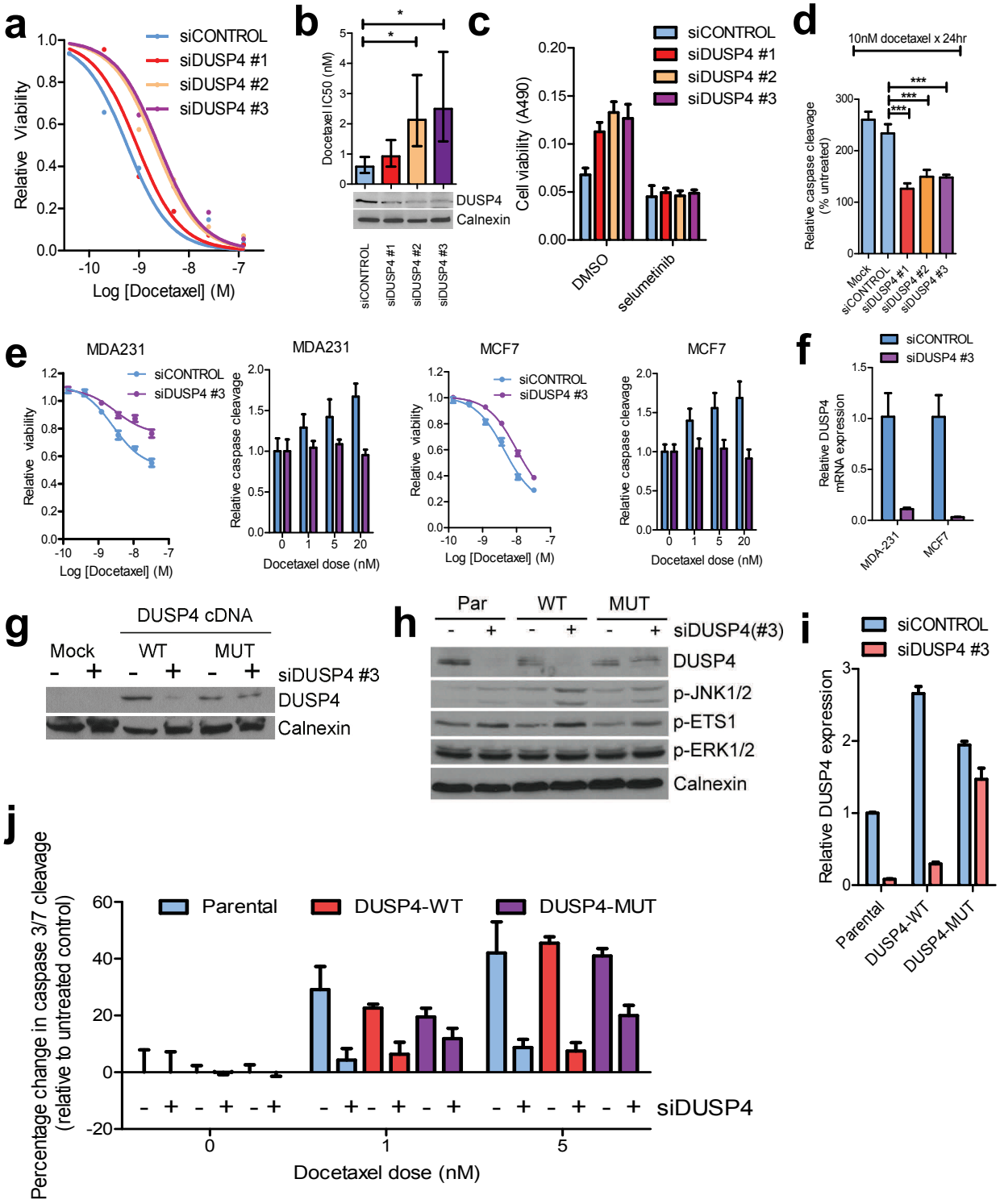
Supplementary Figure 5: Predictive metagene signatures and associations with Ki67. **a-c)** Associations of the Z-scores for the 3 previously published gene signatures/ metagenes (CHEMO, STROMAL_META, and WNT/METS) with post-NAC Ki67 score. Pearson's correlation coefficients were calculated and the association was tested by ANOVA. **d)** Association of the recreated (see supplementary methods) OncotypeDx Recurrence Score (RS) stratification groups versus percentage of tumor cells staining positive for Ki67 by IHC. A Bonferroni post-hoc test was used to make multiple comparisons among the groups; ***P*<0.01, *** *P*<0.001

Supplementary Figure 6: Independent validation of DUSP4 promoter methylation patterns across molecular subtypes of breast cancer



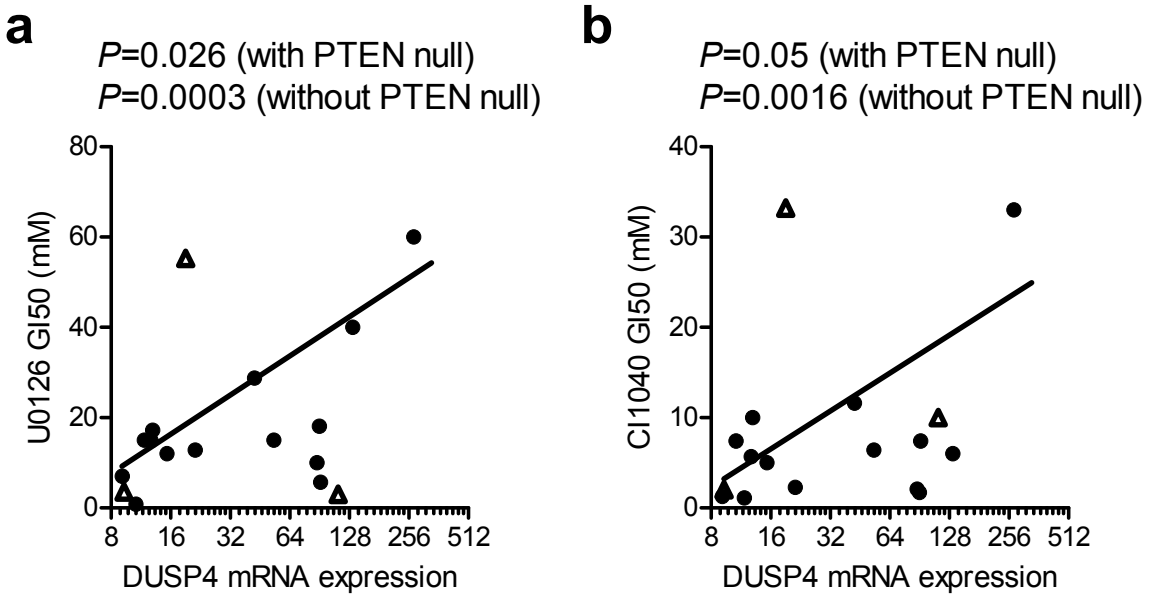
Supplementary Figure 6: Independent validation of D*USP4 promoter methylation patterns across molecular subtypes of breast cancer.* Association of molecular subtype (as annotated by the authors of the study) with DUSP4 promoter methylation data (DUSP4_P925_R) for 60 annotated breast tumors extracted from GSE22135[3]. Relative promoter methylation was tested for differences among subtypes by ANOVA with a Bonferonni post-test to compare selected columns of data (basal-like vs. all others). ** $P<0.01$; *** $P<0.001$

Supplementary Figure 7: DUSP4 depletion induces resistance to the anti-proliferative and anti-apoptotic effects of docetaxel.



Supplementary Figure 7: DUSP4 depletion induces resistance to the anti-proliferative and anti-apoptotic effects of docetaxel. MCF10A cells were transfected with non-targeting siRNA (siCONTROL) or one of three *DUSP4*-targeted siRNAs (siDUSP4 #1 - #3). After 24 h, the cells were replated for the experiments below (a-d). **a)** SRB assay of siRNA cells treated with 5-fold dilutions of docetaxel or DMSO control for 96 h. **b)** Top panel: Docetaxel IC_{50} was calculated from the sigmoidal dose response curves in (a) and tested by ANOVA with Bonferonni post-hoc t-tests (* $P<0.05$). Bars represent mean ± 95% CI. Bottom panel: Western blot of MCF10A cell lysates harvested 72 h after siRNA transfection. **c)** SRB assay of MCF10A cells transfected with siDUSP4 or siCONTROL, followed by treatment with selumetinib for 96 hr. Bars represent mean ± SD of 3 experiments. **d)** Caspase-Glo assay performed after 24 h of docetaxel treatment in siRNA transfected MCF10A cells. Bars represent mean ± SD (n=3). **e)** MDA-231 and MCF7 cells were transfected with siDUSP4 #3. After 16 h, the cells were replated on 96-well dishes. Twenty-four h later, cells were treated with docetaxel at the indicated concentrations. Twenty-four h later, the Caspase-Glo assay was performed to measure caspase 3/7 cleavage and in duplicate plates, the SRB assay was performed to measure cell viability after 48 h of docetaxel treatment. All values are normalized to siRNA-transfected cells (no docetaxel treatment) to control for differences in cell viability due to DUSP4 knockdown. **f)** qRT-PCR demonstrating mRNA knockdown of DUSP4 at 48 h after siRNA transfection. **g)** Immunoblot analysis of HEK293T cells transfected with pLX301-DUSP4-WT or mutant pLX301-DUSP4-MUT, followed by siRNA transfection (siCONTROL or siDUSP4 #3) **h)** MDA-231 cells were transduced with DUSP4-WT or DUSP4-MUT and selected with puromycin. Ninety-six h after transfection with siDUSP4 #3 or siCONTROL, cells were harvested and cell lysates prepared for immunoblot analysis. **i)** In parallel to (h), qRT-PCR for DUSP4 was performed 72 h after siRNA treatment. **j)** Forty-eight h after siRNA transfection, cells from (h) were replated on 96-well plates, allowed to adhere overnight, and treated for 24 h with 0, 1, 5, or 20 nM docetaxel. Caspase 3/7-Glo assays were performed and luminescence was measured to quantify apoptosis. Data are presented as % increase over control, bars representing mean +/- SD.

Supplementary Figure 8: DUSP4 mRNA expression correlates with sensitivity to MEK inhibitors.



Supplementary Figure 8: DUSP4 mRNA expression correlates with sensitivity to MEK inhibitors. Association of GI50 values for U0126 and CI1040 versus DUSP4 mRNA expression in BLBC cell lines. Sensitivity data were previously published [1,4] and used to determine the linear association with *DUSP4* mRNA expression. Gene expression data were used to determine *DUSP4* levels were previously published as part of the IC-BP50 panel annotation dataset[1]. P-values are given for both the full analysis (with PTEN-null cell lines) and after removing cell lines that are known to harbor a PTEN mutation (BT549, MDA468 and HCC70 [denoted with open triangles]). The trend lines represent the best-fit line for cells with wild-type PTEN only.

Supplementary Figure References:

1. Neve, R.M., *et al.* A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes. *Cancer Cell* **10**, 515-527 (2006).
2. Paik, S., *et al.* A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. *N Engl J Med* **351**, 2817-2826 (2004).
3. Bediaga, N.G., *et al.* DNA methylation epigenotypes in breast cancer molecular subtypes. *Breast Cancer Res* **12**, R77.
4. Mirzoeva, O.K., *et al.* Basal subtype and MAPK/ERK kinase (MEK)-phosphoinositide 3-kinase feedback signaling determine susceptibility of breast