SUPPLEMENTARY MATERIAL

I. Supplementary Figures:

II. Supplementary Tables:

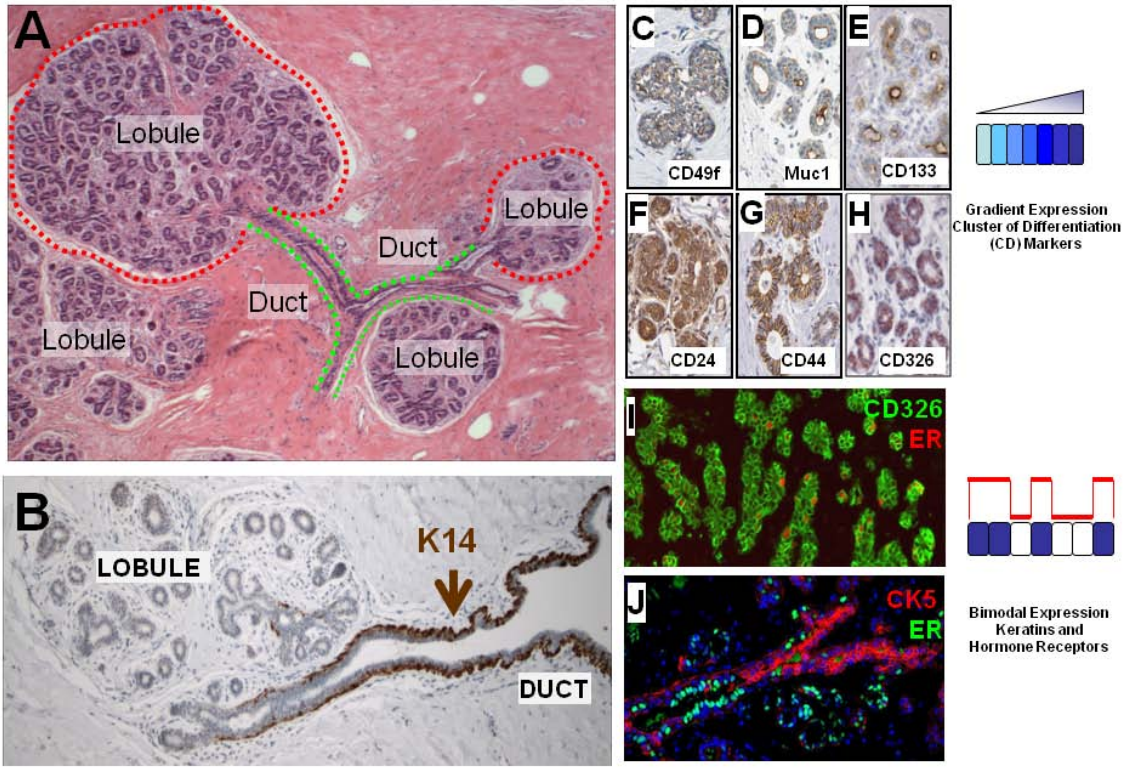
III. Supplementary Materials and Methods

IV. Supplementary References

V. Images corresponding to Table S2

SUPPLEMENTARY FIGURES

Figure S1: Anatomy of normal human breast and expression patterns of keratins, cluster differentiation markers and hormone receptors in normal human breast.



(A) H&E stained section of normal human breast epithelium, showing lobules and interlobular ducts (40x).

(B) Immunohistochemistry for K14 on normal human breast epithelium. While the interlobular ducts stain with K14 (brown) in the myoepithelial cells only, the majority of myoepithelial cells in the lobules are negative for K14 in this section (100x).

(C-H) Immunohistochemistry for cluster differentiation markers makers CD49f, Muc1, CD133, CD24, CD44 and CD326 (Ep-CAM), revealed a gradient of expression; while some cells expressed lower levels and others higher levels, there were not distinct positive and negative populations. There was no differential expression between ducts and lobules (100x).

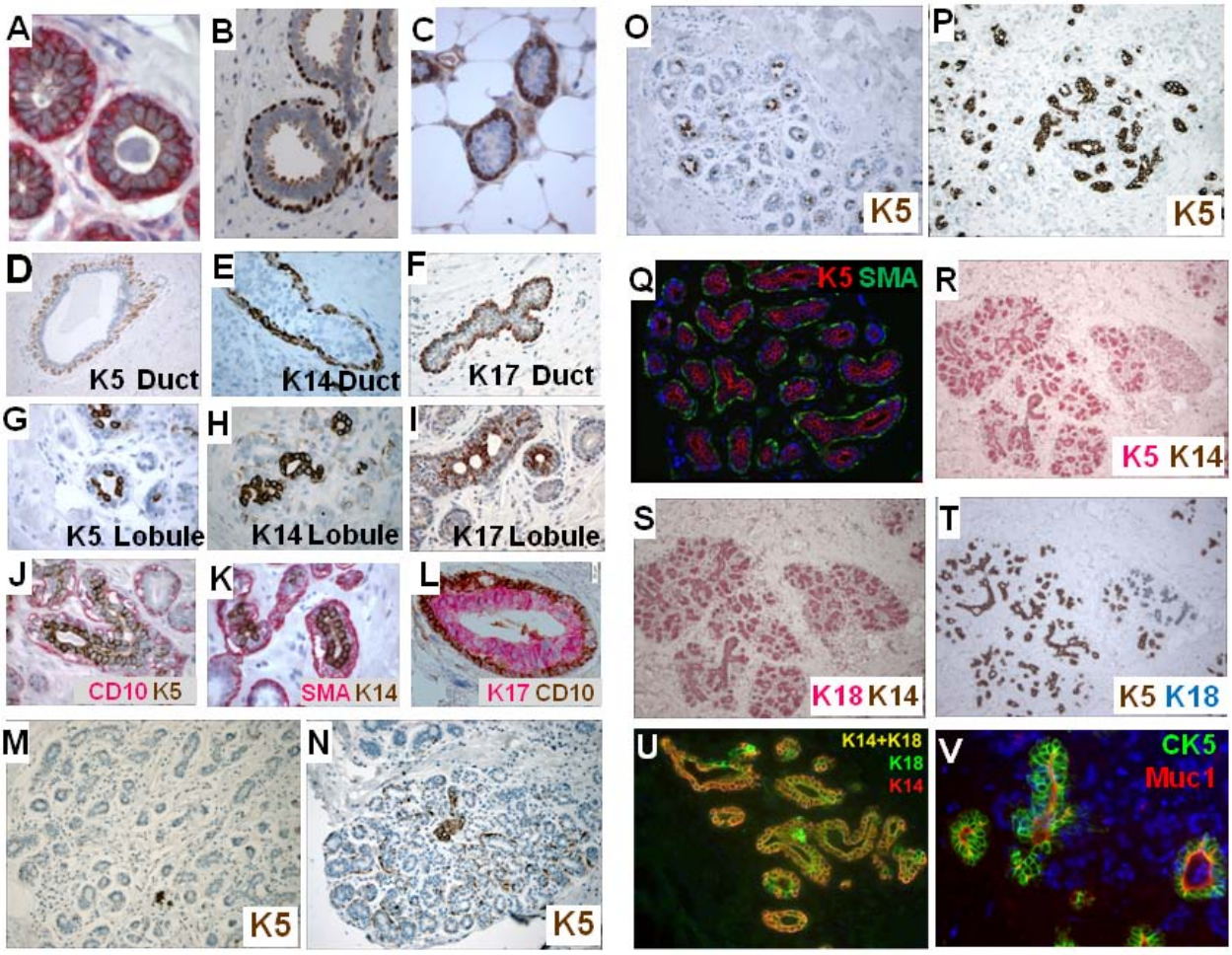
(I) Immunofluorescence staining for ER and CD326 (Ep-CAM) shows that both ER+ and ER− cells are CD326+. Thus, there is no correlation between the expression of these two markers. ER=red, CD326=green (100x).

(J) Immunofluorescence staining for ER and K5 shows that ER+ and K5+ staining are bimodal (biphasic). Unlike the CD markers there was a clearly positive and a negative cell population for each marker, and ER+ and K5+ were mutual exclusive, i.e., if a cell is K5+ it is almost always ER−, and vice versa. ER=green, K5=red (100x).

See http://sylvester.org/ince/supplemental-material for original high resolution images and additional examples.

Discussion: We evaluated the *in situ* expression of putative breast stem cell markers such as CD44, CD24, CD326 (Ep-CAM), and CD133 with specific attention to breast lobules. Many of these CD surface markers have been useful to distinguish different hematopoietic cell subtypes from each other due to their cell-type specific expression. Several of these markers have been used to enrich putative breast stem cell populations with FACS (1). Based on the hematopoietic paradigm, we expected that there would be two distinct luminal cell layer subpopulations (positive and negative) in normal breast tissue with each of these markers. However, immunohistochemical staining of normal human breast tissue sections for these putative breast stem cell markers - CD24, CD44, CD49f, CD133, CD326 (Ep-CAM) and Muc1 - did not discriminate distinct subpopulations of breast cells. Surprisingly, they were broadly expressed in the vast majority of breast epithelial cells in whole tissue sections evaluated by standard immunohistochemistry. Although there was a gradient of staining intensities, distinct positive and negative populations were difficult to discern. Also, we did not observed any differences in the expression of these markers in lobules vs. ducts. In the end, many of the classic CD markers did not prove to be useful in distinguishing between luminal breast cell types when examined *in situ* with immunostains. It is worth noting that enzymatic digestion of solid tissues for long periods with collagenase and trypsin at 37°C is necessary to generate single cell suspensions for FACS analysis. Thus, it is possible that some of the apparent CD-marker negative breast cell sub-populations previously reported using FACS analysis may have emerged due to surface protein degradation during tissue digestion (2). Another possible reason might be the differences between the sensitivity and dynamic range of FACS vs. *in situ* staining. In addition, since FACS is more quantitative than IHC, it allows isolation of subpopulation of cells using markers that are expressed in a gradient pattern.

Figure S2: Keratin 5, 14 and 17 expression in normal human mammary gland ducts versus lobules.



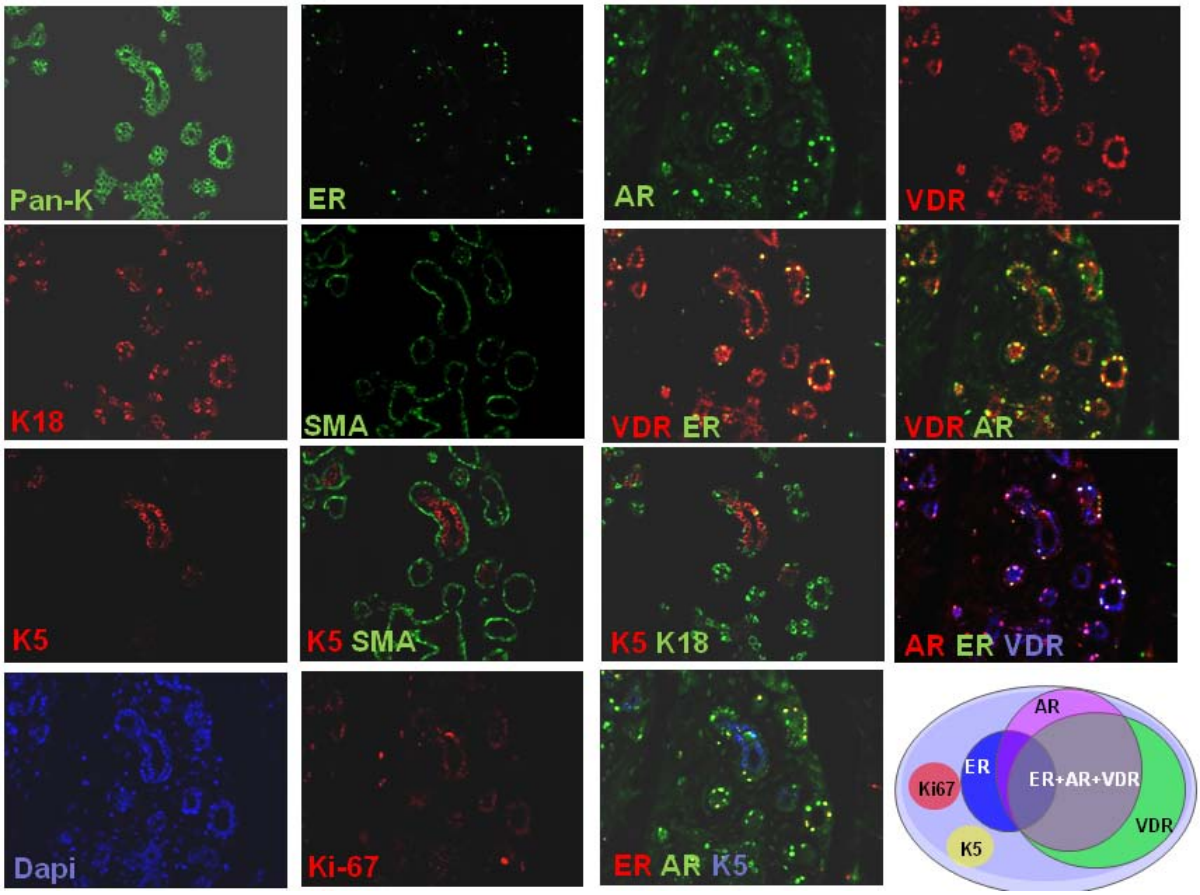
(A) Double immunostains show that all luminal cells express Claudin-4 (brown) and all myoepithelial cells express SMA (red) (600x). **(B)** All myoepithelial cells express p63 (600x).**(C)** Keratin 5 immunostains of mouse mammary glands show that K5 is expressed only in the myoepithelial layer, unlike human breast that has abundant luminal K5 expression (K5=brown,400x). **(D-F)** Keratins 5, 14 and 17 are expressed in the myoepithelial (basal) layer in the interlobular ducts of normal human breast tissue (K5, 14, 17=brown; blue counter stain, 100x). **(G-I)** Keratins 5, 14 and 17 are expressed in the luminal layer in the lobules of normal human breast tissue (K5, 14, 17=brown; blue counter stain, 200x). **(J-L)** Double immunostains with pan-myoepithelial markers CD10 and SMA show that K5/14/17 cells are located above the CD10/SMA+ myoepithelial cells next to the lumen. K5 brown (J), K14=brown (K), K17=red (L), CD10=red (J), SMA=red (K), CD10= brown (L) (J-K= 200x, L=400x). **(M-O)** The frequency of K5+ cells can be dramatically different from lobule to lobule even in the same section from the same person; in many lobules there are almost no K5+ cells (M,

100x), and in rare lobules K5 is expressed in some of the myoepithelial cells but not in luminal cells (N,100x). In many lobules nearly half of the luminal cells are K5+ (0, 100x). **(P-Q)** In some lobules almost all the luminal cells are K5+, either with single K5 stain (Brown, 100x), or with double immunostain that confirms the luminal nature of these cells (K5=red, SMA=green200x). **(R-T)** Double **i**mmunostaining of serial sections of the same lobule show that in some lobules nearly all the cells are K5 and K18 double positive. K5 red, K14 brown (Q), K18 red, K14 brown (R), K5 brown, K18 Blue (S). In these double IHC stains both markers are cytoplasmic and the brown stain is dominant. Thus absence of brown in panel R suggest that the large lobule on left is almost entirely composed of red (K5+/K14−) cells and absence of brown in panel S indicates that the same cells are K18+/K14−. In panel T, the brown K5 staining confirms that the large lobule on the left is almost entirely composed on K5+/K18+ double positive cells, with a small lobule on the right that is blue consistent with a K5−/K18+ lobule (20x). **(U)** Consistent with double IHC results in panels R-T**,** double IF stains show that in some lobules almost all luminal cells are double K18 and K14 positive. K18=green, K14=red, K14/18 double positive cells=yellow (200x). **(V)** Most of the K5 luminal cells express Muc1, a marker of differentiation. K5=green, Muc1=red (400x). See http://sylvester.org/ince/supplemental-material for original high resolution images and additional examples.

What might be the roots of the common misconception that K5/14/17 are basal keratins? The difference in keratin 5, 14 and 17 expression patterns between ducts and lobules highlights one of the shortcomings of using homogenized tissue fragments for research as opposed to *in situ* examination. One of the reasons we refer to 'omics' approaches as having **low morphologic resolution** is because in most of these studies mRNA or protein profiles derived from normal or tumor lysates are extracted from a mm^3-cm^3 fragments of tissue. Naturally, such an approach completely misses the important anatomical differences in expression of K5, K14, K17, such as differences between ducts vs. lobules. Another contributing factor leading to the wide-spread mistaken belief about K5, K14 and K17 might be over-generalizing results among species. For example, K5/14 are indeed exclusively basal (myoepithelial) keratins in mouse mammary tissue. It seems that some have extrapolated these findings to human breast, without careful examination. A third reason might be over-generalizing mRNA expression data. It has been shown that for many keratins there is only a weak correlation between mRNA and protein levels, which might have resulted in referring to some cells as basal due to keratin mRNA expression, even though these cells express very low protein levels for these keratins (3). Lastly, others before us have pointed out in several papers that K5/14 are expressed in the luminal layer (4-8). K5/14/17, however, are still referred to as basal keratins by most, perhaps partly because some papers that contain important histopathological findings are still published without expert pathologist involvement or review (9).

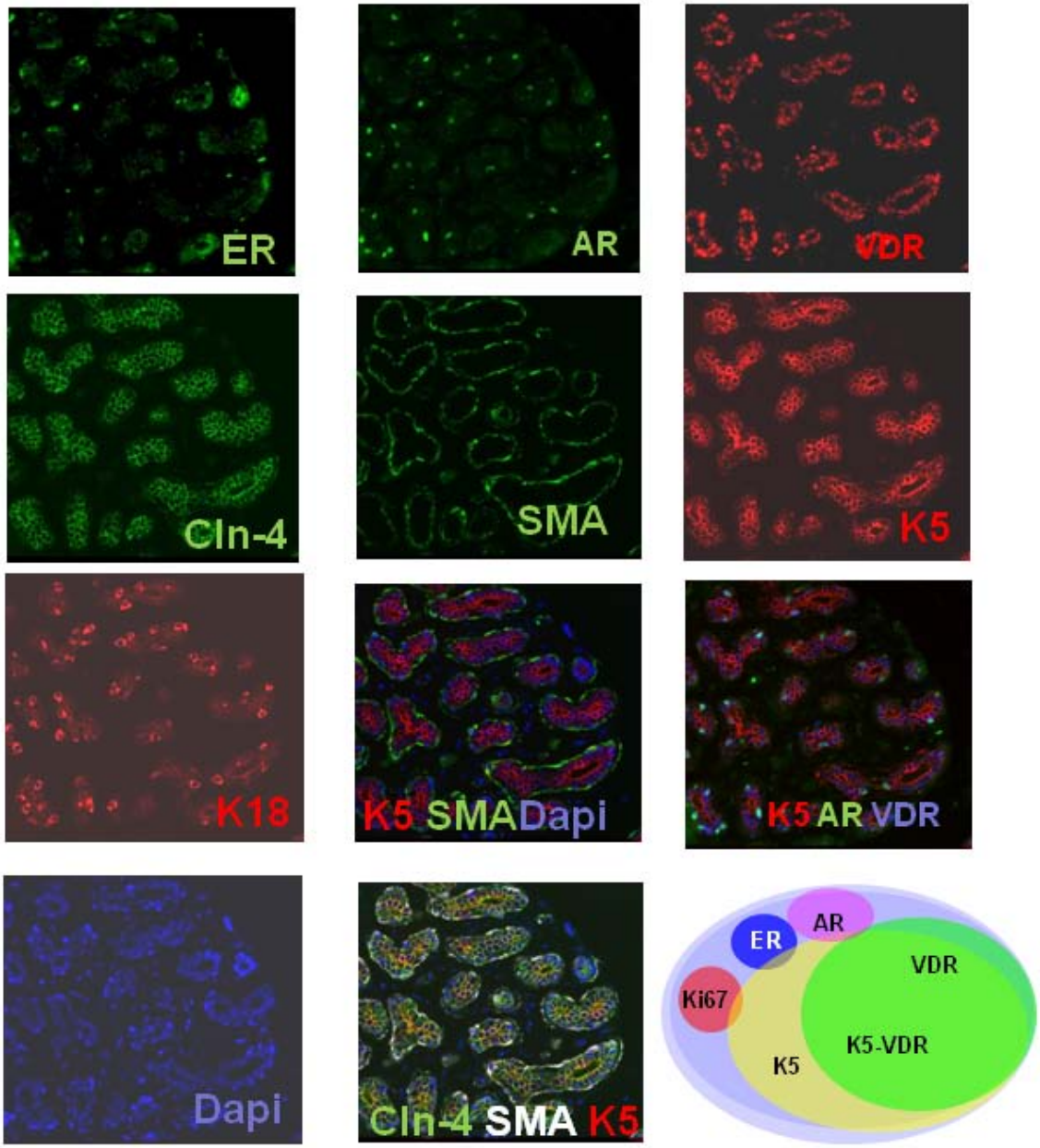
Figure S3: Multiplex immunofluorescent staining of normal human breast lobules.

S3A: Normal human breast lobule with high AR and VDR expression.



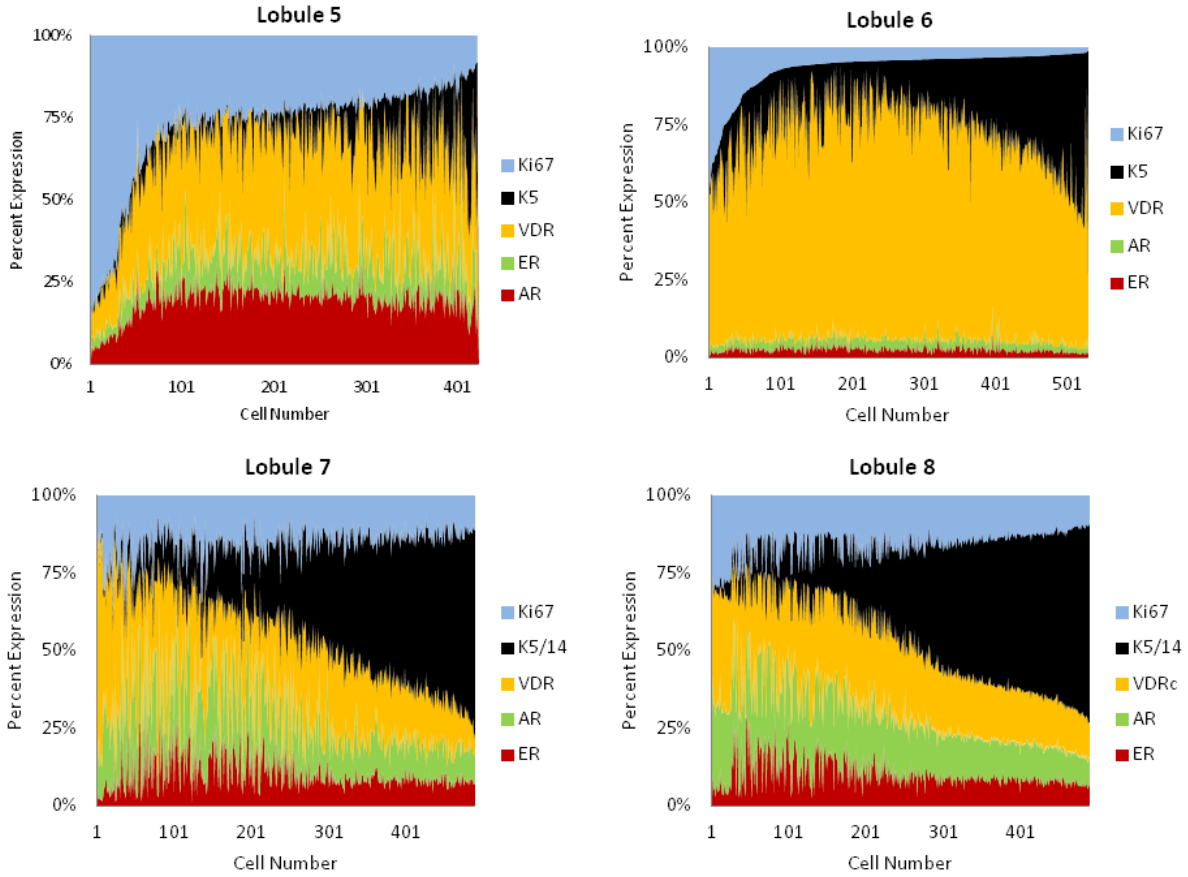
Example 2 of multiplexed immunofluorescence using the GE Healthcare multiplexed marker platform (example 1 is shown in Figure 3). The same section of normal breast epithelium was evaluated serially for markers Pan-K, K18, K5, ER, SMA, Ki-67, AR, and VDR. Merged images for selected images are shown (200x). Na-K ATPase and CD10 (see supplemental data). A Venn diagram depicting the dynamic relationship of the various markers in this case is shown. This lobule is predominantly composed of AR+ VDR+ and AR/VDR+ cells. The ER+, K5+ and Ki-67+ cells are rare. Representative images from GE Healthcare multiplexed immunofluorescence marker platform are shown (see supplemental data for the remaining images).It has been proposed that 'microdissection' of tissue may improve the quality of -omics approaches, which can be the case in tumors. However, the cell-to-cell diversity we describe here makes this all but impossible for normal tissues. Thus, datasets that use normal bulk tissue as a control completely miss the normal tissue diversity, which skews tumor vs. normal comparisons. See http://sylvester.org/ince/supplemental-material for original high resolution images and additional examples.

3B: Normal human breast lobule with high VDR/K5 expression.



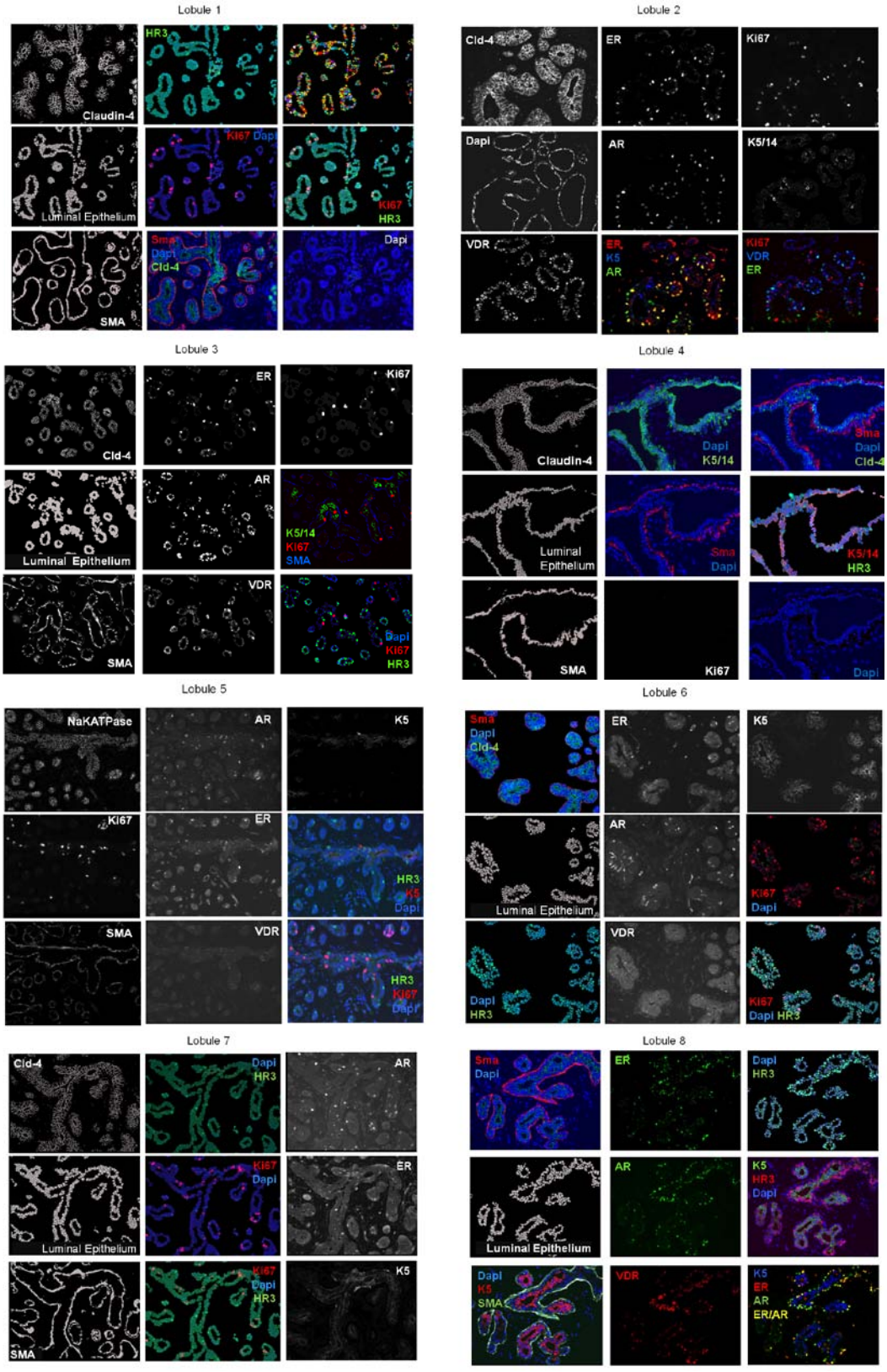
The same section of normal breast epithelium was evaluated serially for markers ER, AR, VDR, Cln-4 (Claudin-4), SMA, K5, K18. Merged images for selected images are shown (400X). A Venn diagram depicting the dynamic relationship of the various markers in this case is shown. This lobule is predominantly composed of K5+, VDR+ and K5/VDR+ cells. The ER+, AR+ and Ki-67+ cells are rare. Representative images from GE Healthcare multiplexed immunofluorescence marker platform are shown. See http://sylvester.org/ince/supplemental-material for original high resolution images and additional examples.

S3C: Multiplex image analysis of normal human breast lobules.

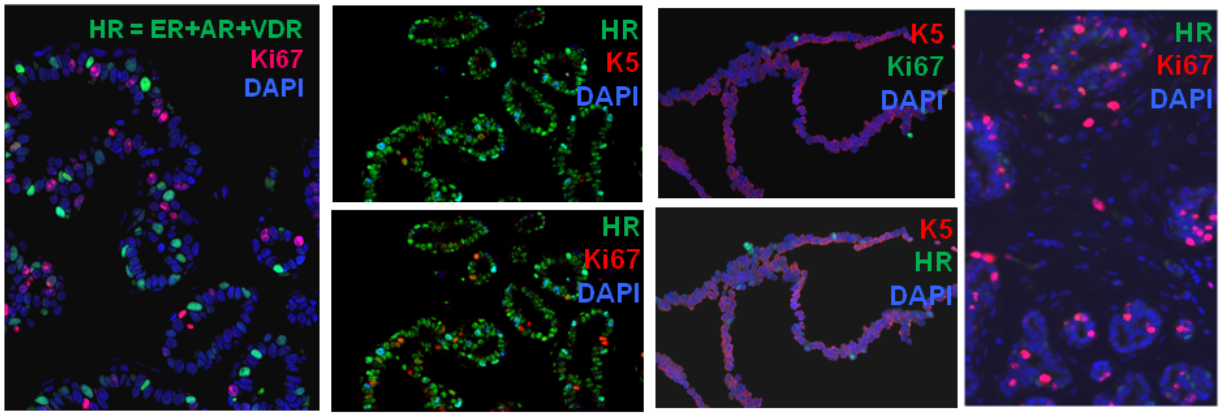


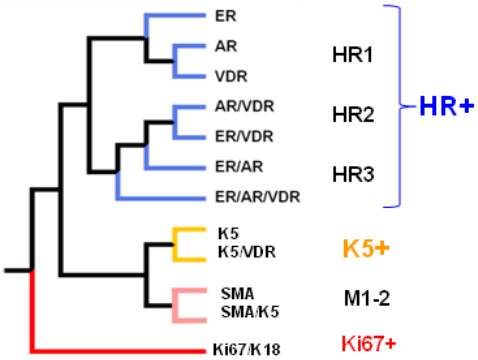
Histograms show the relative ER (red), AR (green), VDR (yellow), K5 (black) and Ki67 (blue) expression in each cell. The cells are numbered in the abscissa (x-axis), and relative contribution of each marker to the total fluorescence of each cell is expressed as a percentage of the total (ordinate, y axis). The results were sorted from low Ki67 to high Ki67 values. Immunofluorescence images show the eight lobules that were analyzed. Lobule 5 is an example of lobule with high ER/AR/VDR (HR) and low K5. The more proliferative cells with high Ki67 where >75% of the total fluorescence comes from Ki67 (~ cells 1-25) express lower levels of HR. Lobule 6 is an example of a lobule dominated by VDR+ and VDR/K5+ cells. As expected the proliferating Ki67+ cells are low in this lobule. Lobule 7 and 8 are examples of average lobules that are a mixture of all cell types. Nevertheless, cells with lower numbers express high HR and low K5, and as HR expression decreases K5 increases. Since these two lobules are dominated by HR and K5, the Ki67 high proliferative cells are rare.

S3C: Multiplex image analysis of normal human breast lobules (400X)



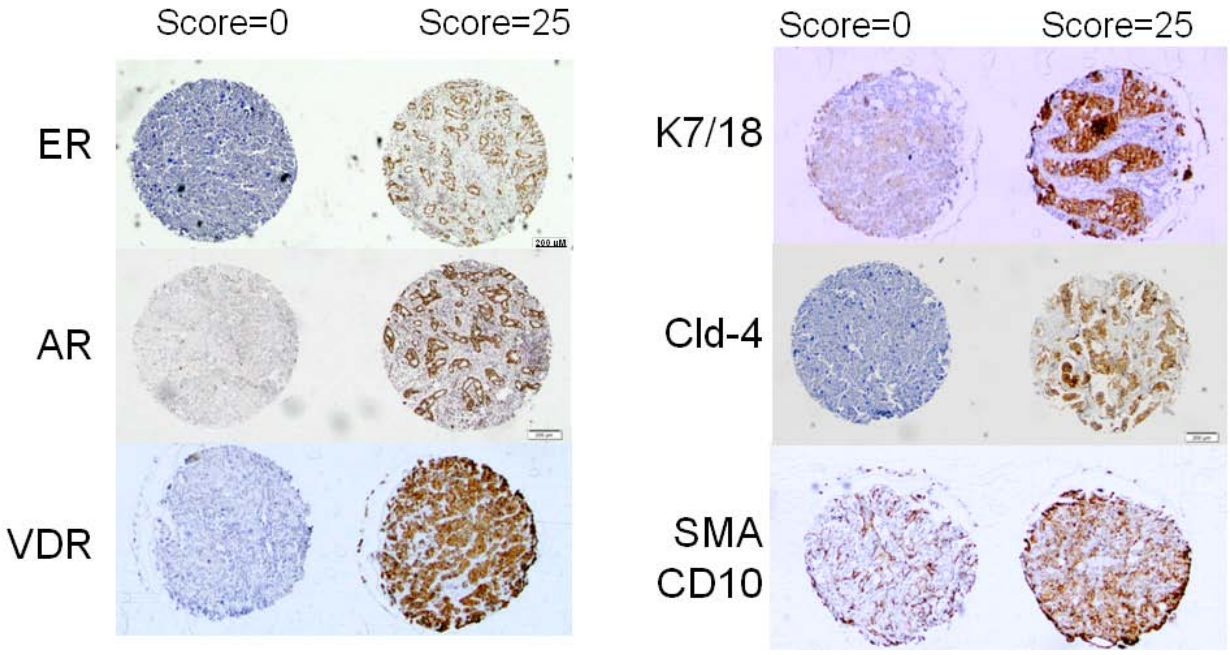
S3D: Models of normal breast cell diversity.





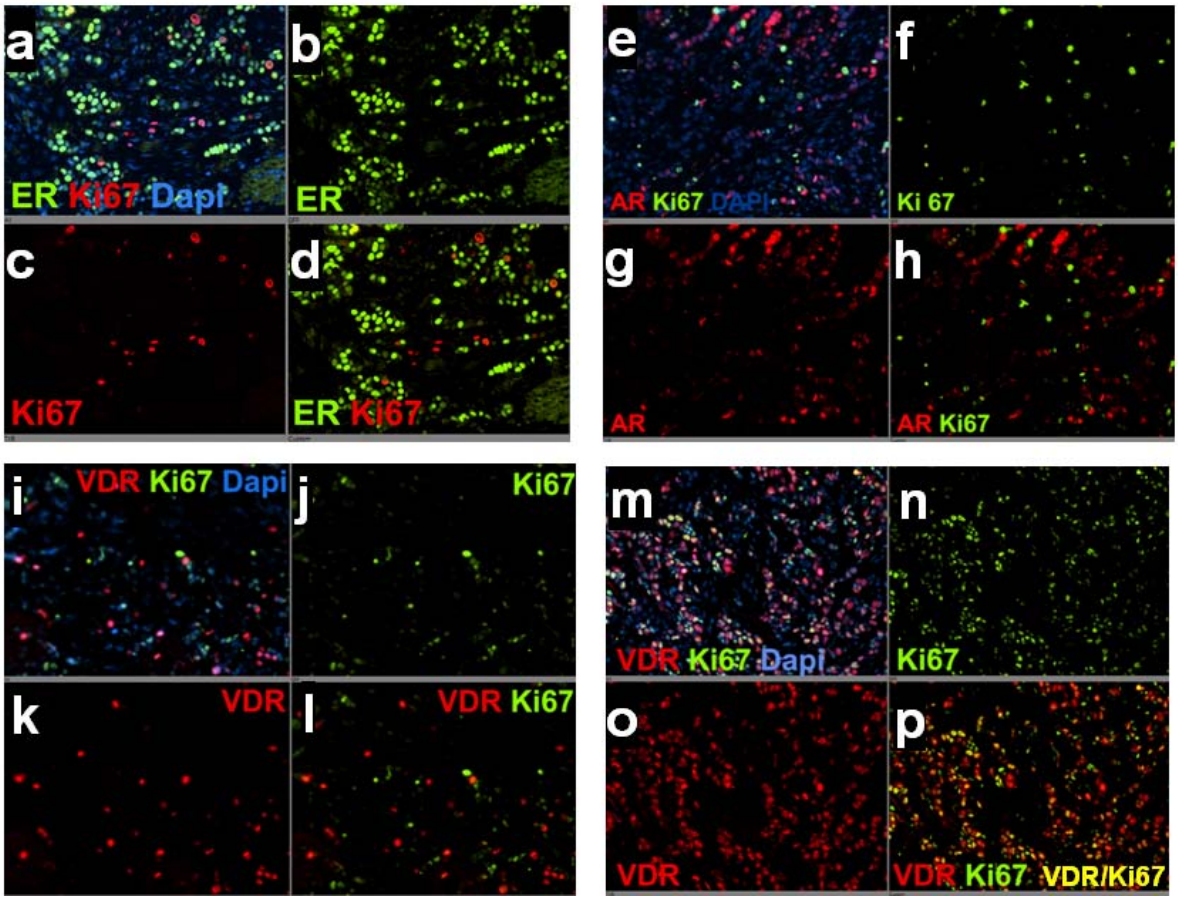
Merged images of multiplexed immunofluorescence staining on the same section are shown. **(Left Column)** typical breast lobule with moderate levels of HR expression and mutually exclusive Ki-67 positive proliferating cells (red); **(Second Column)** Breast lobule that is almost entirely composed HR+ cells (green) with infrequent K5+ (**top**) and Ki-67+ (**bottom**) cells (red); **(Third Column)** Breast lobule that is almost entirely composed K5+ cells (red), with infrequent Ki-67+ (**top**) and HR+ (**bottom**) cells (green); **(Fourth Column)** Highly proliferative breast lobule with numerous Ki-67+ cells (red), with no HR+ cells (green); **(Lower Panel)** Using a phylogenetic analysis software (Mesquite), we constructed an inferred possible differentiation lineage tree of the breast cell subtypes. According to this model it appears that all the HR+ cells have a common progenitor that is different from the common progenitor of K5+ luminal cells and myoepithelial cells (M1-2). The common progenitor of all three groups (HR+, K5+, and M1-2) appears to be a Ki-67/K18+ and HR/K5/SMA− cell consistent with a proliferating transit-amplifying progenitor cell type. See http://sylvester.org/ince/supplemental-material for original high resolution images and additional examples.

Figure S4: Expression of cell specific markers in human breast tumors.



Similar to normal tissues, ER, AR, VDR, K5, K7, K14, K18, Cld-4, SMA, CD10 were expressed in a binomial pattern in tumors. A typical example for each marker is shown in adjacent cores in the TMAS; negative cores on the left (<1% staining, and 0 intensity = score of 0) and positives core on the right (>80% staining, and 5 intensity = score of 25). These cores were adjacent to each other on the TMA (scale bar= 200 µm). One of the advantages of *in situ* staining is the ability to discriminate cell specific expression; for example SMA/CD10 were strongly expressed in the tumor cells in some cases (core on the right with score of 25), and in the stromal cells in other cases (core on the left with a score of 0). This kind of cell type specific expression information is lost in molecular analysis of tumor extracts.

Figure S5: Expression of ER, AR, VDR and Ki-67 in breast cancers.



(A-D) Majority of Ki-67+ tumor cells were mutually exclusive with ER+ tumor cells (200x).

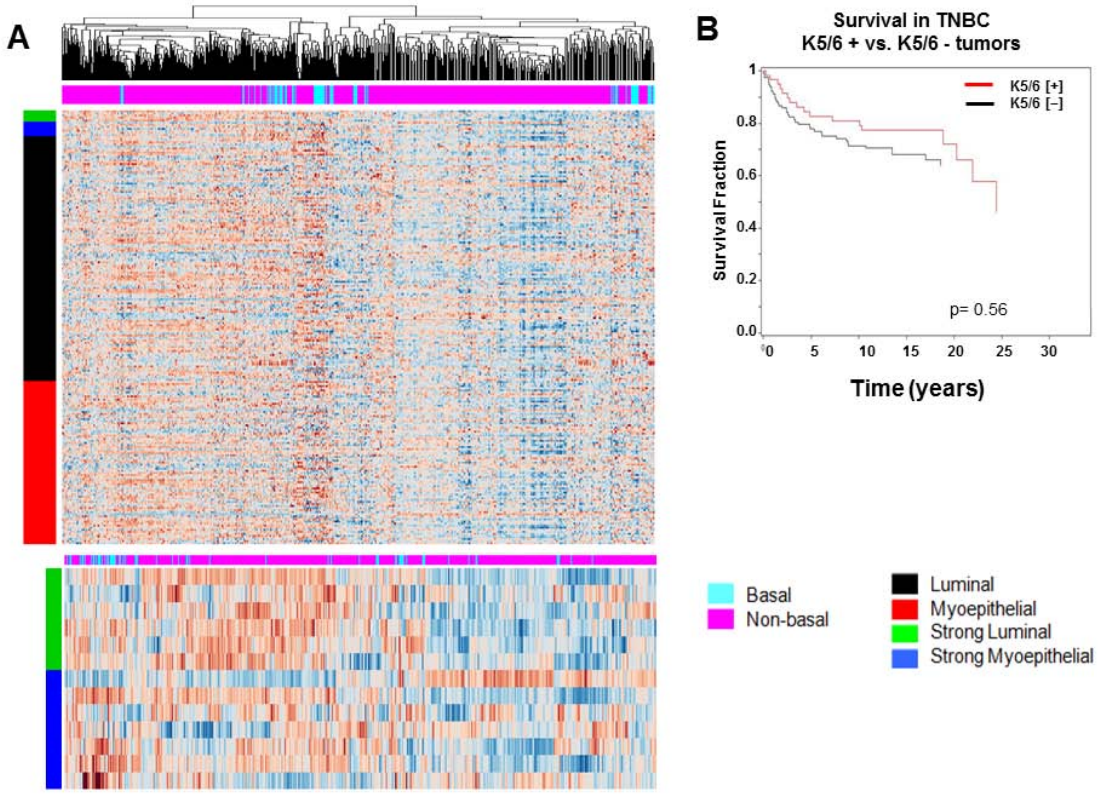
(E-H) Majority of Ki-67+ tumor cells were mutually exclusive with AR+ tumor cells (200x).

(I-L) Ki-67+ tumor cells were generally mutually exclusive with VDR+ tumor cells (200x).

(M-P) In some areas the VDR+ tumor cells were Ki-67+, indicating that in tumors the relationship between VDR expression and proliferation is different than ER and AR (200x).

See http://sylvester.org/ince/supplemental-material for original high resolution images and additional examples.

Figure S6: Expression of Keratin 5 and 14 in human breast tumors.



(A) Analysis of mRNA in normal human breast cells. The normal luminal vs. basal specific profiles were derived by combining three different studies that profiled highly purified luminal vs. myoepithelial cells (10, 11). A subset of mRNAs were differentially expressed between luminal (131 transcripts, Table S4A) and myoepithelial cells (90 transcripts, Table S4B) consistently in all three studies – these served as the basis for strong luminal (green bar) and strong myoepithelial (dark blue) consensus signatures; differential expression in two of three studies formed the basis for luminal (black) and myoepithelial (red) signatures, depicted with the colored bars on the left hand side of the heatmap (12-15). When a combined human breast mRNA expression dataset was analyzed for the expression of these luminal basal genes, there was no significant difference between basal tumors vs. luminal (non-basal) tumors , marked with colored bars above the heatmap; basal tumors (light blue) vs. luminal (non-basal) tumors (pink). Rows = genes, columns = tumor sample from each patient; over-expression = orange, under-expression = blue. See supplemental table 4 for the list of luminal and myoepithelial specific transcripts.

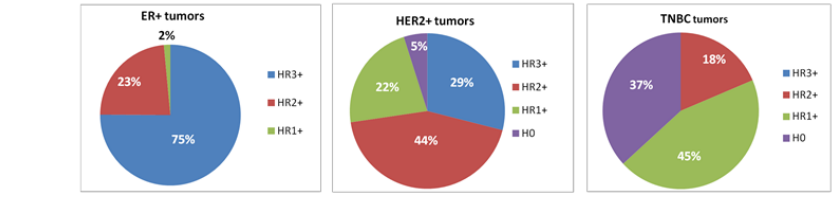
(B) Kaplan–Meier analysis of all individuals with triple negative invasive breast cancer from the Nurses' Health Study that were scored by K5/6 immunohistochemistry (n=172). There was no statistically significant survival difference between the K5/6+ (n=59) vs. K5/6− (n=113) tumors (p=0.56).

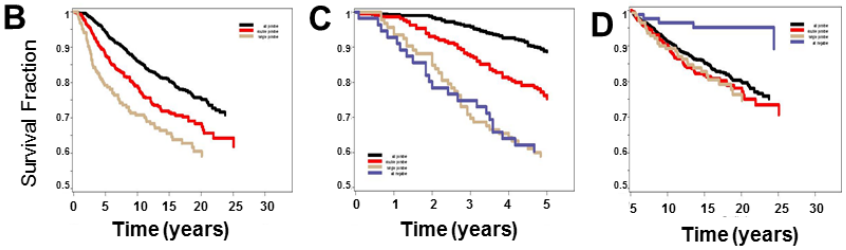
Figure S7: Reclassification of human breast tumors based HR0-3 categories

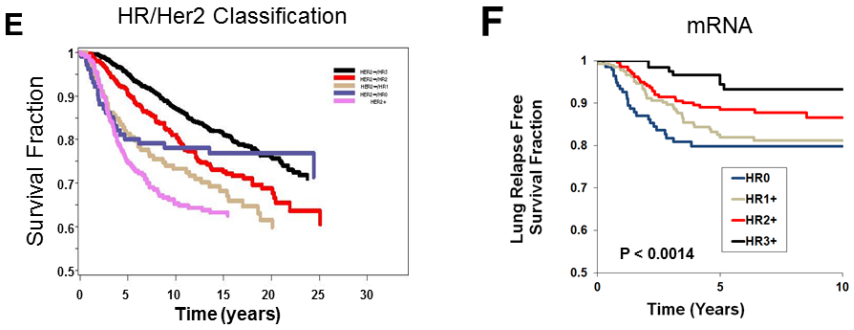
A

Classification of breast tumors based on HR0-3 categories

Characteristic	ER+ Tumors	HER2+ Tumors	TNBC	Total
HR0	0 % (0)	5.1 % (9)	36.8 % (95)	6.1% (104)
HR1	1.5 % (19)	22.0 % (39)	44.6 % (115)	10.8% (185)
HR2	23.4 % (294)	43.5 % (77)	18.6 % (48)	25.0% (429)
HR3	75.1 % (946)	29.4 % (52)	0% (0)	58.2% (998)
Total	100 % (1259)	100 % (177)	100% (258)	100% (1731)







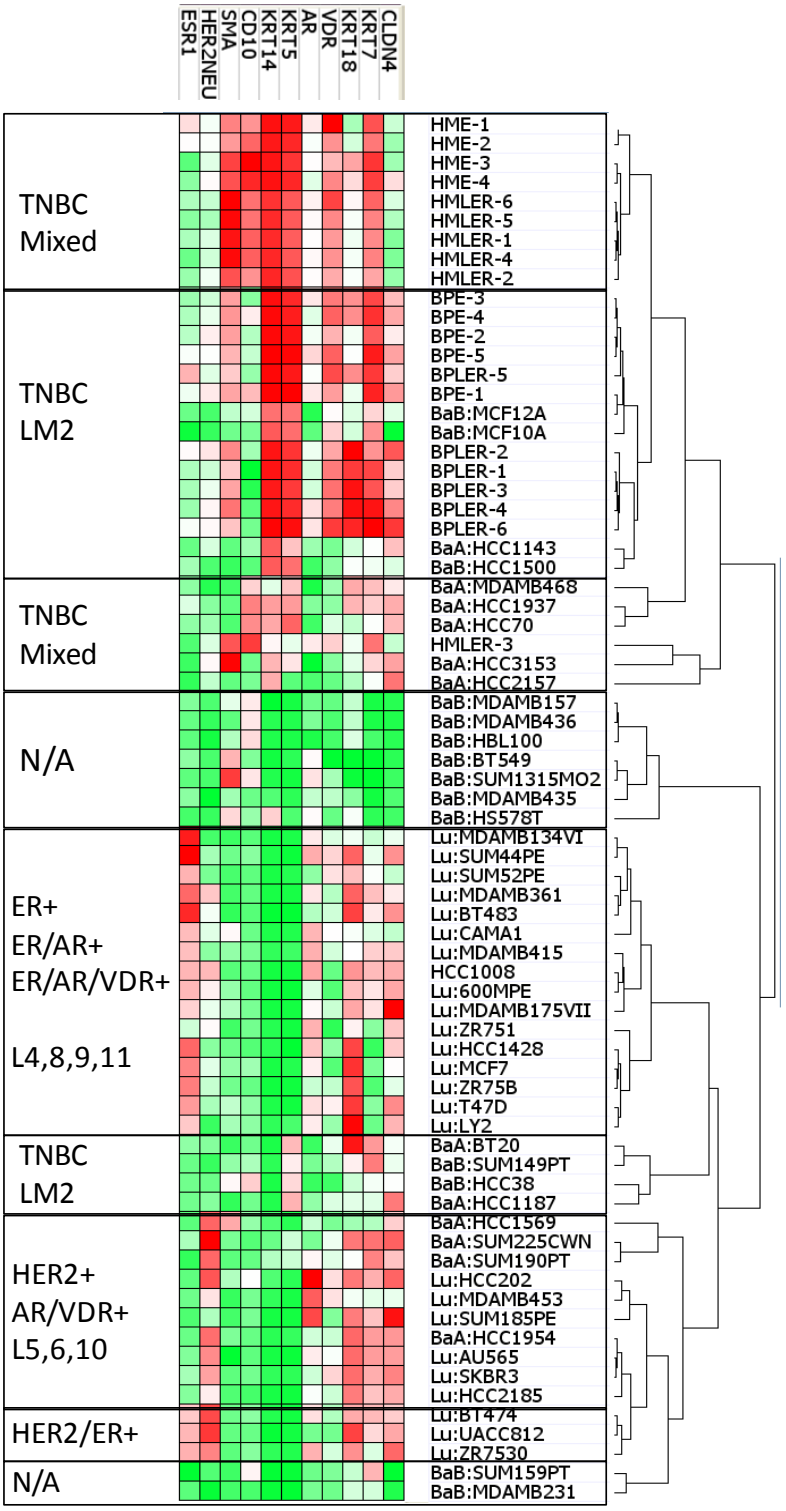
Number at risk

	0	5	10	15	20	25	25+
HR3	954	877	762	509	227	66	13
HR2	362	314	263	156	88	21	4
HR1	146	115	100	63	37	15	4
HR0	102	80	76	63	36	12	3
HER2	219	162	139	97	49	24	4

(A) Table showing the frequency of HR 0, 1, 2 and 3 tumors in the clinical categories of ER+, HER2+ and TNBC and pie charts showing the composition of breast tumors based on HR classification. **(B-D)** Kaplan–Meier analysis of all individuals with invasive breast cancer from the Nurses' Health Study that were scored by immunohistochemistry in this study according to hormone receptor categories (HR3+ = black curve, HR2+ = red curve, HR1+ = yellow curve, HR0 = blue curve). We examined survival of HR1-3 only as a continuous variable during 0-25 years (B), HR0-3 during 0-5 years (C), and HR0-3 during 5-25 years (D). **(E)** Kaplan–Meier analysis of all individuals with invasive breast cancer from the Nurses' Health Study. In this analysis the patients with HER2+ tumors were analyzed as a distinct group (HR3+ = black curve, HR2+ = red curve,

HR1+ = yellow curve, HR0 = blue curve, HER2+ = pink curve). (**F**) Kaplan-Meier analysis of lung metastasis free survival for all invasive breast cancers from an 855 patient breast tumor dataset (16). Tumors were ranked according to gene expression values for ER and scored as 'ER_High' or 'ER_Low' based on a 50% cut-off point. The same approach was used to identify 'AR_High and AR_Low' groups, as well as 'VDR_High' or 'VDR_Low'. These groups were then assembled based on HR status.

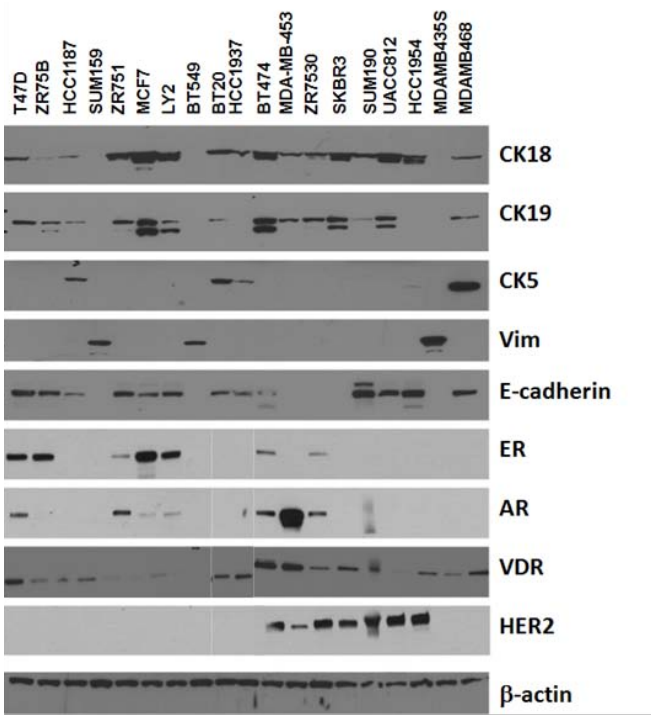
Figure S8: HR0-3 classification of human breast cancer cell lines



A) Breast cancer cell lines were clustered by mRNA expression of differentiation state transcripts (Cld-4, K7, K18, VDR, AR, K5, K14, CD10, SMA, HER2, and ER and grouped according to the normal cell phenotypes L1-11. We found that ER+ cancer cell lines are similar to normal cell types L4, L8, L9 and L11, and HER2+ cell lines are similar to L4-11 (Table S3). The majority of triple negative breast cancer (TNBC) cell lines were similar to L3 cell type (LM2) or had a mixed phenotype. Interestingly, none of the TNBC cell lines conformed to the L2-like (LM1) phenotype, even though such tumors are nearly one third of TNBC tumors. Furthermore, nine cell lines that are frequently used as models of TNBC such as; MDA-MB-231, MDA-MB-436, MDA-MB-157, MDA-MB-435, SUM-159PT, SUM1315, HBL100, BT549, and HS578T had an expression profile that was not present either in normal breast cells or in human breast cancers (i.e. negative for nearly

all of the HR and epithelial markers), classified here as no in vivo counterpart (n=9, **N/A**). Thus, while ER+ and HER2+ cell lines generally appear conform to the expected *in vivo* HR phenotypes, the TNBC cell lines do not fully recapitulate the in vivo spectrum. Clustering was performed as described in the Supplementary Materials and Methods. Red indicates a relative increase in expression, green indicates a relative reduction in expression and white indicates no change.

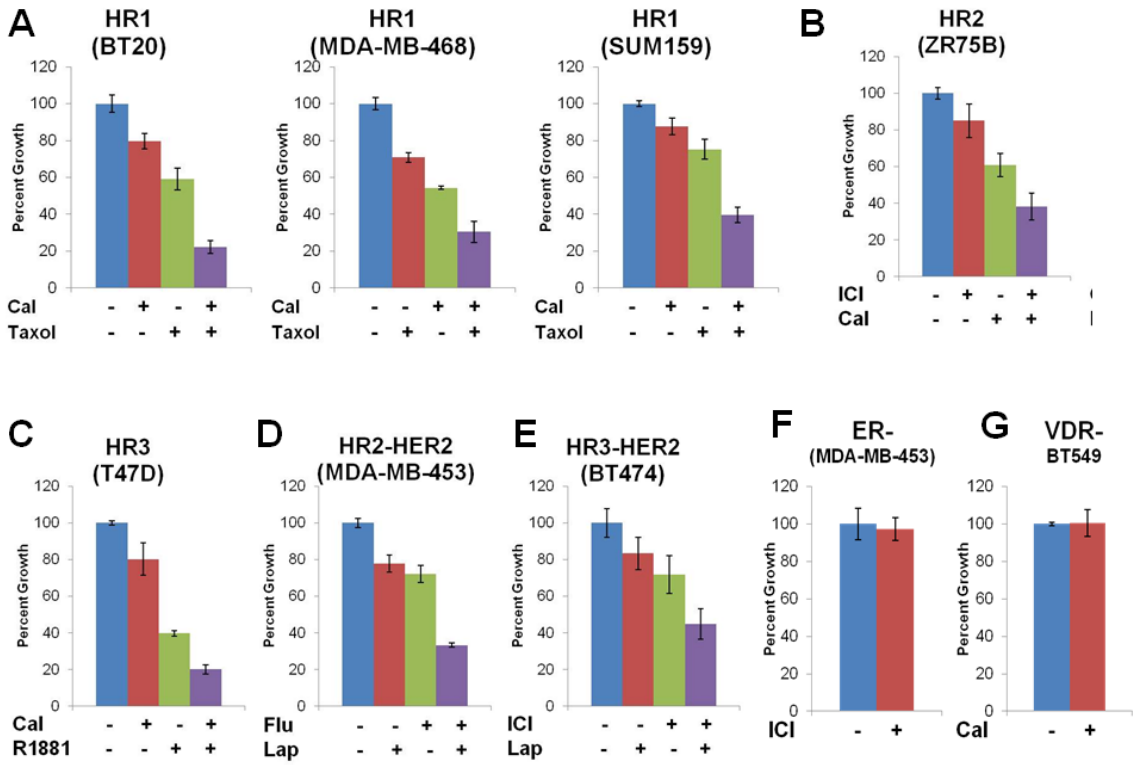
B) The HR phenotype of breast cancer cell lines confirmed with Western blots.



C) The subset of breast cancer cell lines with HR0-3 phenotypes that are selected for drug response experiments.

Phenotype	Breast Cancer Cell Line	ER	AR	VDR	HER2
HR0	BT549	-	-	-	-
HR1	BT20	-	-	+	-
HR1	MDA-MB-468	-	-	+	-
HR1	SUM159	-	-	+	-
HR2	ZR75B	+	-	+	-
HR3	T47D	+	+	+	-
HR0	UACC812	-	-	-	+
HR1	SKBR3	-	-	+	+
HR2	MDA-MB-453	-	+	+	+
HR3	BT474	+	+	+	+

Figure S9: Drug response of HR1-3 breast cancer cell lines to combined hormone treatment.



In all experiments described below low doses of individual drugs (< 50% inhibition) were used in order to demonstrate the additive effect of two drugs combined. Each experiment was repeated multiple times with similar results; representative results from one experiment are shown.

(A) VDR and Taxol combination treatment of HR1 cell lines: The ER/HER2 negative breast cancer cell lines BT20, MDA-MB-468 and SUM159 are traditionally considered TNBC models, hence not candidates for hormone treatment. However, according to our HR classification these HR1 (VDR+) tumors can be candidates for treatment with VDR agonists, which we hypothesized would inhibit their proliferation. Consistent with this we observed a reduction in proliferation when these HR1 breast cancer cells were treated with VDR agonist Calcitriol (Cal, 25nM) alone or in combination with a chemotherapeutic such as Taxol (0.5nM or 1nM). Interestingly the combined effect of these drugs was greater than either drug alone, indicating an additive effect. In the clinical setting this may allow using less toxic doses of each drug with the same efficacy as using higher doses of Taxol alone. These results indicate that combining VDR agonists with chemotherapy should be explored further for in HR1 subtype of TNBC in combination with chemotherapy.

(B) ER and VDR combination treatment of HR2 cell lines: The breast cancer cell line ZR75B is a model of ER/VDR+ HR2 tumors that could be potentially targeted with

ER/VDR combination hormone therapy. We tested this hypothesis by combining the ER-antagonist ICI 182,780 (Faslodex, ICI 5nM) with VDR agonist Calcitriol (Cal, 50nM); and found that the combined effect of these drugs was greater than using either drug alone. Representative results from multiple experiments are shown. These results indicate that combining ER antagonists with VDR agonists should be explored further in this subtype of HR2 breast cancers.

(C) AR and VDR combination treatment of HR3 cell lines: The ER+ breast cancer cell line T47D also expresses AR and VDR (HR3). In this cell line combination of AR-agonist R1881 (Methyltrienolone, 50nM) with VDR agonist Calcitriol (Cal, 50nM) inhibited proliferation more effectively than either drug alone. These experiments were carried out in phenol red free DMEM +5% charcoal stripped FBS and 17-beta estradiol (E2, 10nM) as previously described (17, 18). Representative results from multiple experiments are shown. These results indicate that combining AR and VDR agonists should be explored further in this subtype of HR3 breast cancers.

(D) AR and HER2 combination treatment of HR2/HER2 cell lines: The combination of AR-antagonist Flutamide (Flu 45 μM) and HER2 inhibitor Lapatnib (Lap, 0.5μM) additively inhibited proliferation HR2+/HER2+ breast cancer cell line MDA-MB-453. Representative results from multiple experiments are shown.

(E) ER and HER2 combination treatment of HR2/HER2 cell lines The combination ER-antagonist ICI 182,780 (Faslodex, ICI 10nM) and HER2 antagonist Lapatnib (Lap, 10nM) additively inhibited proliferation HR3+/HER2+ breast cancer cell line BT474, as previously described (19). Representative results from multiple experiments are shown. The above results (M, N) indicate that combining ER and AR antagonists with HER2 targeted therapy should be explored further for HER2+/HR+ breast cancers. Our results confirm a previous study by Emde et al. in which the BT474 breast cancer cell line was treated with ICI 182,780 and Lap (19).

(F) In ER-negative control cell line (MDA-MB-453) no inhibition of cell proliferation was seen when it was treated with ER antagonist ICI 182,780 even at a much higher concentration (Faslodex, ICI 100nM).

(G) In VDR-negative control cell line (BT549) no inhibition of cell proliferation was seen when it was treated with VDR agonist Calcitriol (Cal,) at different concentrations of 10nM, 25nM and 100nM. Control and 100nM groups were shown here.

Supplemental Table 1

	Antibodies Used to Stain Breast Sections								
	Antibody	**Company**	**Catalog#**	**Species**	**Clone**	**Ag Retrieval**	**Primary Dilution**	**Cell Compartment**	**Used in Figures**
1	**AR**	Dako	M3562	mouse	AR441	citrate PC	500	nuclear	Fig. 1-5, S2,S4
2	**CD10**	Vector	VP-C328	mouse	56C6	citrate PC	20	cytoplasm	Fig. 1-5, S2
3	**K 8/18**	Becton/Dick	349205	mouse	cam5.2	10' Protease	50	cytoplasm	Fig. 1-5, S2
4	**K14**	Serotec	MCA890	mouse	n/a	citrate PC	600	cytoplasm	Fig. 1-5, S2
5	**K17**	Millipore	MAB1677	mouse	E3	citrate PC	400	cytoplasm	Fig. S2
6	**K19**	Dako	M0888	mouse	RCK108	10' Protease	1000	cytoplasm	Fig. 1, S2
7	**K5**	Lab Vision	MS-1896	mouse	XM26	citrate pc/mv	50/30	cytoplasm	Fig. 1-5, S2
8	**K5/6**	Chemicon	MAB1620	mouse	D5/16B4	EDTA PC	400	cytoplasm	Fig. 5
9	**K7**	Dako	M7018	mouse	OV-TL 12/30	10' Protease	1000	cytoplasm	Fig. 1-5, S2,3,4
10	**ER**	NeoMarker	RM-9101-S	rabbit	SP1	citrate PC	100	nuclear	Fig. 1-5, S2,3,4
11	**Ki67**	Dako	M7240	mouse	mib1	citrate pc/mv	200	nuclear	Fig. 1-5, S2,3,4
12	**Muc-1**	Novocastra	NCL-Muc1	mouse	Ma552	citrate PC	600	cytoplasm	Fig. S3
13	**NaKATPase**	Epitomics	2047-1	rabbit	EP1845Y	citrate PC	200	membrane	Fig. 3-4
14	**p63**	Dako	M7247	mouse	4A4	citrate MV	800	nuclear	Fig. 4,5
15	**PanK**	Sigma	C1801	mouse	PCK-26	citrate MV	300	cytoplasm	Fig. S3
16	**PR**	Dako	M3569	mouse	PgR636	citrate PC	200	nuclear	Fig. 4,5
17	**SMA**	Sigma	A2547	mouse	1A4	none	20K	cytoplasm	Fig. 1-5, S2,3
18	**VDR**	Novus	NBP1-19	rabbit	n/a	citrate PC	200	nuclear	Fig. 2-4, S2,S4
19	**VDR**	Santa Cruz	sc-13133	Mouse		citrate PC	2000	cyt/nuc	Fig. 5,6
20	**Vimentin**	Abcam	ab8069	mouse	V9	citrate PC	2000	cytoplasm	Fig. 1
21	**AR**	Millipore	06-680	Rabbit	N/A	citrate PC		nuclear	
22	**K5**	Abcam	ab75869	Rabbit	EPR1600Y	citrate PC	100	cytoplasm	
23	**K7**	Abcam	Ab68459	Rabbit	EPR1619Y	citrate PC	100	cytoplasm	
24	**K14**	LabVision	LL025	Mouse	MS-620-P	citrate PC		cytoplasm	
25	**K18**	NeoMarkers	MS-142	Mouse		citrate PC	200	cytoplasm	
26	**K18**	Abcam	E431-1	Rabbit	Ab32118	citrate PC	100	cytoplasm	
27	**K19**	LabVision	MS-198-P	Mouse				cytoplasm	
28	**Ki67**	Vector Labs	VP-K451	rabbit		citrate PC	2500	nuclear	
29	**Vimentin**	Dako	M7020	mouse	Vim3B4	20' protease	400	cytoplasm	
	Direct Conjugated Antibodies								
30	**CD24-PE**	BD	555428	mouse	ML5			membrane	
31	**CD44-APC**	BD	559942	mouse	G44-26			membrane	
32	**CD133/1-PE**	Miltenyi	130-080-801	mouse	AC133			membrane	
33	**CD326-FITC**	Serotec	MCA1870FT	mouse	VU-ID9			membrane	Fig. S1
34	**CD326-Cy5.5**	BD	347199	mouse	EBA-1			membrane	
35	**K 8/18 -FITC**	Abcam	ab54728	mouse	cam5.2			cytoplasm	Fig. S2
36	**ERα-AF488**	Santa Cruz	sc-542K	rabbit	MC-20			nuclear	Fig. S1
37	**Ki-67-AF488**	Santa Cruz	sc-7846	goat	M-19			nuclear	
			Secondary Antibodies						
	Species	**Company**	**Catalog #**	**Species**	**Specificity**	**Conjugate**			
1	Rabbit	Invitrogen		goat	IgG	AlexaFluor647			
2	Rabbit	R&D	NL004	donkey	IgG	NorthLight557			
3	Mouse	Invitrogen		goat	IgG	AlexaFluor647			
4	Rabbit	GE	NA934V	donkey	IgG	HRP			
5	Mouse	GE	NXA931	Sheep	Ig	HRP			
6	Mouse	Jackson	115-115-164	Goat	IgG	R-PE			

The list of antibodies used in this study. The immunostains with twenty antibodies (1-20) are shown in figures 1-5 and S1-5. In order to exclude antibody isoform or clone specific artifacts, staining was repeated with a second antibody from a different clone or manufacturer (antibodies 21-37, data not shown). In all cases, the results with multiple antibodies were consistent.

Supplemental Table 2: Co-expression frequency of hormone receptors ER, AR and VDR, Keratins 15 and 14, and proliferation marker Ki-67

Normal human breast FFPE tissue sections were double-stained with the antibodies indicated below. Five representative sections were selected for counting the number of cells that are positive for each antibody alone and positive for both antibodies indicating co-expression. A total of 12,531 cells were counted; the corresponding 98 images showing the areas that were counted are provided in the supplemental data.

	Table S2 Image Files	Number of Cells Counted	Number of Image Files
STS2a	K5 & ER co-expression	684	4
STS2a	K14 & ER co-expression	345	5
STS2a	K17 & ER co-expression	2,284	5
STS1b	ER & Ki67 co-expression	1,026	5
STS1b	K14 & Ki67 co-expression	374	5
STS1b	K5 & Ki67 co-expression	332	9
STS1b	K17 & Ki67 co-expression	372	6
STS1c	K5 & AR co-expression	376	5
STS1c	K14 & AR co-expression	413	5
STS1c	AR & Ki67 co-expression	698	5
STS1c	ER & AR co-expression	429	5
STS1d	K5 & VDR co-expression	266	5
STS1d	AR & VDR co-expression	835	8
STS1d	ER & VDR co-expression	749	6
STS1d	VDR & Ki67 co-expression	179	5
STS1e	CD10 & Ki67 co-expression	1,084	6
STS1f	K14 & K18 co-expression	746	4
STS1f	K5 & K18 co-expression	1,339	5
Total		**12,531**	**98**

In the tables below each row is a different normal breast section stained with the indicated double immunostains (see Supplemental Data p. 30-57 for corresponding images). The total number of cells positive for each marker (column 1-2), both markers (column 3), the percentage of double-positive cells as a fraction of cells positive for each marker (column 4-5) and both markers (column 6) are depicted. Each table shows single and double positive cells as a percent of total. Only luminal cells were counted in these experiments unless otherwise indicated.

Table S2A: Co-expression frequency of ER, with K5, K14 and K17

K5 & ER co-expression				% Overlap		
Cell Counts	ER	K5	[K5+ER]	% of ER	% of K5	% of Total
Image 1	183	103	0	0%	0%	0.0%
Image 2	97	32	0	0%	0%	0.0%
Image 3	100	54	0	0%	0%	0.0%
Image 4	18	97	0	0%	0%	0.0%
Image 5						
Total	**398**	**286**	**0**	**0%**	**0%**	**0%**

K14 & ER co-expression				% Overlap		
Cell Counts	ER	K14	[K14+ER]	% of ER	% of K14	% of Total
Image 1	61	37	0	0%	0%	0.0%
Image 2	45	40	1	4%	5%	2.0%
Image 3	39	33	0	0%	0%	0.0%
Image 4	35	24	0	3%	4%	2.0%
Image 5	21	9	0	0%	0%	0.0%
Total	**201**	**143**	**1**	**0.5%**	**0.7%**	**0.3%**

K17 & ER co-expression				% Overlap		
Cell	ER	K17	[K17+ER]	% of	% of	% of
Image 1	605	12	0	0%	0%	0.0%
Image 2	629	10	0	0%	0%	0.0%
Image 3	773	12	0	0%	0%	0.0%
Image 4	197	11	0	0%	0%	0.0%
Image 5	29	6	0	0%	0%	0.0%
Total	**2233**	**51**	**0**	**0%**	**0%**	**0%**

There was no overlap between K5/ER (0.0 %, n=684), K17/ER (0.3%, n= 2286), and a very small overlap between K14/ER (0.3%, n=344). Thus, ER+ luminal cells are essentially mutually exclusive populations with K5/K14/K17 expressing luminal cells.

Table S2B: Co-expression frequency of Ki67 with ER, K5, K14 and K17

ER & Ki67 co-expression				% Overlap		
Cell Counts	ER	Ki67	[ER+ Ki67]	% of ER	% of Ki67	% of Total
Image 1	46	60	1	2%	2%	0.9%
Image 2	201	56	0	0%	0%	0.0%
Image 3	174	62	0	0%	0%	0.0%
Image 4	374	52	0	0%	0%	0.0%
Image 5						
Total	**795**	**230**	**1**	**0.1%**	**0.4%**	**0.1%**

K14 & Ki67 co-expression				% Overlap		
Cell Counts	K14	Ki67	[K14+Ki67]	% of K14	% of Ki67	% of Total
Image 1	56	120	0	0%	0%	0.0%
Image 2	28	37	2	7%	5%	3.1%
Image 3	65	23	4	6%	15%	4.5%
Image 4	12	26	0	0%	0%	0.0%
Image 5						
Total	**161**	**206**	**7**	**4.2%**	**3.3%**	**1.9%**

K5 & Ki67 co-expression				% Overlap		
Cell Counts	K5	Ki67	[K5+ Ki67]	% of K5	% of Ki67	% of Total
Image 1	46	59	1	2%	2%	1.0%
Image 2	40	61	0	0%	0%	0.0%
Image 3	19	30	0	0%	0%	0.0%
Image 4	56	20	0	0%	0%	0.0%
Image 5						
Total	**161**	**170**	**1**	**0.6%**	**0.6%**	**0.3%**

K17 & Ki67 co-expression				% Overlap		
Cell Counts	K17	Ki67	[CD10+ Ki67]	% of K17	% of Ki67	% of Total
Image 1	9	63	0	0%	0%	0.0%
Image 2	4	40	0	0%	0%	0.0%
Image 3	33	12	0	0%	0%	0.0%
Image 4	22	42	0	0%	0%	0.0%
Image 5	118	29	0	0%	0%	0.0%
Total	**186**	**186**	**0**	**0.0%**	**0.0%**	**0.0%**

The minimal overlap between Ki67/ER (0.1%, n=1026), Ki67/K5 (0.3%, n=332), and Ki67/K17(0%, n=372) suggest that K5+, K17+ or ER+ cells and proliferating (Ki67+) cells are essentially mutually exclusive populations. The K14/Ki67 overlap was slightly higher (1.9 %, n=374).

Table S2C: Co-expression frequency of AR wit K5, K14, Ki67 and ER

K5 & AR co-expression				% Overlap		
Cell Counts	AR	K5	[K5+AR]	% of AR	% of K5	% of Total
Image 1	144	46	0	0%	0%	0.0%
Image 2	82	27	0	0%	0%	0.0%
Image 3	13	15	0	0%	0%	0.0%
Image 4	22	27	0	0%	0%	0.0%
Image 5						
Total	**261**	**115**	**0**	**0%**	**0%**	**0%**

K14 & AR co-expression				% Overlap		
Cell Counts	AR	K14	[K14+ AR]	% of AR	% of K14	% of Total
Image 1	20	14	0	0%	0%	0.0%
Image 2	19	5	0	0%	0%	0.0%
Image 3	192	26	0	0%	0%	0.0%
Image 4	58	79	0	0%	0%	0.0%
Image 5						
Total	**289**	**124**	**0**	**0%**	**0%**	**0%**

AR & Ki67 co-expression				% Overlap		
Cell Counts	AR	Ki67	[AR+ Ki67]	% of AR	% of K5	% of Total
Image 1	142	47	0	0%	0%	0.0%
Image 2	102	96	1	1%	1%	0.5%
Image 3	48	6	0	0%	0%	0.0%
Image 4	33	17	0	0%	0%	0.0%
Image 5	147	59	0	0%	0%	0.0%
Total	**472**	**225**	**1**	**0.2%**	**0.4%**	**0.1%**

ER & AR co-expression				% Overlap		
Cell Counts	ER	AR	[ER+AR]	% of ER	% of AR	% of Total
Image 1	73	12	10	12%	45%	11.8%
Image 2	34	22	14	29%	39%	25.0%
Image 3	62	56	41	40%	42%	34.7%
Image 4	47	30	28	37%	48%	36.4%
Image 5						
Total	**216**	**120**	**93**	**30%**	**44%**	**27.7%**

There was very little overlap between AR/K5 (0 %, n=376), AR/Ki67 (0.1%, n=738), AR/K14 (0%, n=413). In contrast, on average 44% of the AR+ were also ER+ positive (n= 429).

Table S2D: Co-expression frequency of VDR with K15, AR, ER and Ki-67

K5 & VDR co-expression				% Overlap		
Cell Counts	VDR	K5	[K5+VDR]	% of VDR	% of K5	% of Total
Image 1	35	20	6	15%	23%	10.9%
Image 2	41	14	7	15%	33%	12.7%
Image 3	28	7	2	7%	22%	5.7%
Image 4	16	28	4	20%	13%	9.1%
Image 5	17	35	6	26%	15%	11.5%
Total	**137**	**104**	**25**	**15%**	**19%**	**10.4%**

VDR & Ki67 co-expression				% Overlap		
Cell Counts	VDR	Ki67	[VDR+ Ki67]	% of VDR	% of K5	% of Total
Image 1	43	9	0	0%	0%	0.0%
Image 2	47	2	0	0%	0%	0.0%
Image 3	10	8	0	0%	0%	0.0%
Image 4	11	21	0	0%	0%	0.0%
Image 5	24	4	0	0%	0%	0.0%
Total	**135**	**44**	**0**	**0.0%**	**0.0%**	**0.0%**

AR & VDR co-expression				% Overlap		
Cell Counts	AR	VDR	[AR+VDR]	% of AR	% of VDR	% of Total
Image 1	41	171	22	35%	11%	10.4%
Image 2	58	116	11	16%	9%	6.3%
Image 3	15	91	9	38%	9%	8.5%
Image 4	25	276	13	34%	4%	4.3%
Image 5						
Total	**139**	**654**	**42**	**23%**	**6%**	**5.3%**

ER & VDR co-expression				% Overlap		
Cell Counts	ER	VDR	[ER+VDR]	% of ER	% of VDR	% of Total
Image 1	7	166	20	74%	11%	11.6%
Image 2	65	152	24	27%	14%	11.1%
Image 3	41	87	17	29%	16%	13.3%
Image 4	70	100	20	22%	17%	11.8%
Image 5						
Total	**183**	**505**	**61**	**25%**	**11%**	**8.9%**

There was some overlap between VDR/K5 (10.4 %, n=266). On average 23% of AR+ cells were also VDR+ (n=835), and 25% of ER+ cells were VDR+ (n=749), but there was almost no overlap between VDR/Ki67 (0%, n=179).

Table S2E: Co-expression frequency of CD10 and Ki67

CD10 & Ki67 co-expression				% Overlap		
Cell Counts	CD10	Ki67	[CD10+ Ki67]	% of CD10	% of Ki67	% of Total
Image 3	36	56	4	10%	7%	4.3%
Image 2	93	115	0	0%	0%	0.0%
Image 3	55	111	0	0%	0%	0.0%
Image 4	191	209	2	1%	1%	0.5%
Image 5	90	123	3	3%	2%	1.4%
Total	**465**	**614**	**5**	**1.1%**	**0.8%**	**0.5%**

Most of the proliferating cells were in the luminal layer; only 0.5% of the CD10+ myoepithelial cells were also positive for proliferation marker Ki67 (n=1084).

Table S2F: Co-expression frequency of K14 and K5 with K18

K14 & K18 co-expression				% Overlap		
Cell Counts	K18	K14	[K18+K14]	% of K18	% of K14	% of Total
Image 1	70	14	14	17%	100%	16.7%
Image 2	71	47	47	40%	100%	39.8%
Image 3	104	94	94	47%	100%	47.5%
Image 4	109	41	41	27%	100%	27.3%
Image 5						
Total	**354**	**196**	**196**	**36%**	**100%**	**35.6%**

K5 & K18 co-expression				% Overlap		
Cell Counts	K18	K5	[K18+ K5]	% of K18	% of K5	% of Total
Image 1	217	65	65	23%	100%	23.0%
Image 2	120	38	38	24%	100%	24.1%
Image 3	215	60	60	22%	100%	21.8%
Image 4	114	33	33	22%	100%	22.4%
Image 5	213	101	101	32%	100%	32.2%
Total	**879**	**297**	**163**	**16%**	**55%**	**13.9%**

On average 36% of K18+ cells were also K14+ (n=746), and 16% of K18+ cells were also K5+ (n=1339). In contrast, 100% of the luminal K5+ or K14+ cells were also K18+.

Supplemental Table 3
Table demonstrating normal cell counterparts corresponding to breast tumor phenotypes.

	Normal Cellular States							Breast Tumors		
	Luminal Cellular States	ER	AR	VDR	K5 K14 K17	K7 K18 Cld-4	SMA CD10 p63	ER+	HER2+	TNBC
HR0	*L1-2*					+			+	+
	L3				+	+				+
HR1	*L4*	+				+		+	+	
	L5		+			+			+	+
	L6			+		+			+	+
	L7			+	+	+			+	+
HR2	*L8*	+	+			+		+	+	
	L9	+		+		+		+	+	
	L10		+	+		+			+	+
HR3	*L11*	+	+	+		+		+	+	

Supplemental Table 4

Normal human breast luminal and myoepithelial cells were previously purified and profiled for mRNA expression by three different groups (10, 20, 21). Grigoriadis *et al*. used double immuno-magnetic sorting methods (ESA/CD10) and identified 907 luminal and 955 myoepithelial transcripts in cells directly isolated from freshly dissociated normal breast tissue. Using the same approach and markers, Jones *et al* identified 132 myoepithelial and 77 luminal specific genes. Raouf *et al*., used CD10/Thy1 (myoepithelial) and CD133/Muc1 (luminal) to isolate cells (10, 20, 21). We compared the cell type specific mRNAs identified in each study and identified 131 Luminal specific mRNAs that were cell type specific in all three data sets(A) and 90 myoepithelial specific mRNAs that were detected all three of these studies (B). This consensus signature was used to examine the expression of these genes in basal-like and non-basal-like human breast tumors in SF6 (12-14, 22).

Supplemental Table 4A

Consensus Normal Luminal mRNA Signature

	hgnc_symb	Affymetrix		hgnc_symbol	Affymetrix		hgnc_symbol	Affymetrix		hgnc_symb	Affymetrix
1	***PRSS8***	***202525_at***	34	NTN4	223315_at	67	ARHGDIB	1555812_a_at	100	RNF141	226106_at
2	***SLC9A3R1***	***201349_at***	35	ENC1	201340_s_at	68	GSTTP1	203815_at	101	HEBP2	203430_at
3	***SLC44A2***	***224609_at***	36	C4orf19	219450_at	69	PGRMC2	213227_at	102	BAIAP2L1	227371_at
4	***RARRES3***	***204070_at***	37	TNFRSF21	214581_x_at	70	CGNL1	225817_at	103	VASN	225867_at
5	***LCN2***	***212531_at***	38	ISG15	205483_s_at	71	SEC23A	204344_s_at	104	LYPLA1	203007_x_at
6	***TNFAIP2***	***202510_s_at***	39	LY6D	206276_at	72		40837_at	105	KRT81	213711_at
7	CX3CL1	823_at	40	ALDH1A3	203180_at	73	KRT23	218963_s_at	106	REEP6	226597_at
8	CX3CL1	203687_at	41	GATSL3	233528_s_at	74	C9orf16	204480_s_at	107	KLF6	208961_s_at
9	IL32	203828_s_at	42	C1orf198	223063_at	75	CTGF	209101_at	108	SCCPDH	201825_s_at
10	SEMA3B	203071_at	43	ELF3	229842_at	76	DEFB1	210397_at	109	NEK6	223158_s_at
11	ERBB3	202454_s_at	44	MYEF2	222771_s_at	77	PVRL2	232079_s_at	110	CLIC3	219529_at
12	ERBB3	1563253_s_at	45	WWC1	216074_x_at	78	QPCT	205174_s_at	111	PIM1	209193_at
13	ERBB3	226213_at	46	PSEN1	238816_at	79	HADH	211569_s_at	112	NFKBIA	201502_s_at
14	FSTL3	203592_s_at	47	RAB11FIP1	225177_at	80	TMC4	226403_at	113	CMPK1	222448_s_at
15	MGLL	211026_s_at	48	CLDN3	203953_s_at	81	NT5DC2	218051_s_at	114	RTP4	219684_at
16	GADD45B	209305_s_at	49	PTGES	207388_s_at	82	APLP2	208702_x_at	115	NPAS2	205460_at
17	CHKA	204266_s_at	50	ZFYVE21	219929_s_at	83	MUC20	226622_at	116	DYNLT1	201999_s_at
18	SLC44A2	225175_s_at	51	PRSS27	232074_at	84	TAF10	200055_at	117		229648_at
19	LGALS3	208949_s_at	52	TM7SF2	210130_s_at	85	TFCP2L1	227642_at	118	SLC25A37	222528_s_at
20	PPAP2C	209529_at	53	NR2F2	215073_s_at	86	RTCD1	203594_at	119	FAM84B	225864_at
21	ATP1B1	201242_s_at	54	RPS27L	218007_s_at	87	NME4	212739_s_at	120	DDAH1	209094_at
22	MFI2	223723_at	55	MGAT4A	231283_at	88	GRAMD3	218706_s_at	121	SUMO3	200740_s_at
23	KLK6	204733_at	56	CLDN4	201428_at	89	WSB2	201760_s_at	122	VTCN1	219768_at
24	BACE2	217867_x_at	57	NEBL	207279_s_at	90	FAM62B	1555829_at	123	RASAL1	219752_at
25	TNFAIP2	202509_s_at	58	IFT172	226324_s_at	91	CGN	223232_s_at	124	C5orf32	224707_at
26	CALML5	220414_at	59	ALDH1B1	209645_s_at	92	ERRFI1	224657_at	125	PROM2	1562378_s_at
27	CDA	205627_at	60	PACSIN2	1554691_a_at	93	STEAP4	225987_at	126	TMEM45A	219410_at
28	EMP1	234233_s_at	61	MCFD2	212246_at	94	GDE1	202593_s_at	127	TSPAN14	221002_s_at
29	GLIPR2	225604_s_at	62	CNN3	201445_at	95	RRAS	212647_at	128	IRAK2	231779_at
30	EZR	208621_s_at	63	PCBD1	203557_s_at	96	RIN3	60471_at	129	MLLT4	224685_at
31	GARNL4	213280_at	64	LXN	218729_at	97	CDCP1	1554110_at	130	VKORC1L1	224881_at
32	GPRC5A	235563_at	65	RARRES1	206391_at	98	SLC12A7	218066_at	131	ZNF84	228630_at
33	TSPAN13	217979_at	66	CSGALNACT1	219049_at	99	CD99	201028_s_at			

Supplemental Table 4B

Consensus Normal Myoepithelial mRNA Signature

hgnc_symbol	Affymetrix ID
TIMP3	201147_s_at
TIMP1	201666_at
CAV1	212097_at
TRIM29	202504_at
MMP3	205828_at
DST	204455_at
HTRA1	201185_at
MYLK	202555_s_at
TP63	211194_s_at
TP63	209863_s_at
TP63	211195_s_at
TP63	211834_s_at
TP63	207382_at
PTHLH	210355_at
ITGA6	201656_at
TIMP3	201149_s_at
TIMP3	201148_s_at
TIMP3	201150_s_at
JAG1	216268_s_at
SLC1A5	208916_at
CAV1	203065_s_at
COL1A1	202310_s_at
CTSC	225647_s_at
CTSC	225646_at
SPARC	212667_at
TRIM29	211002_s_at
TRIM29	211001_at
DST	216918_s_at
SPRR1A	213796_at
SFN	209260_at
SFN	33322_i_at
SFN	33323_r_at
COL3A1	215076_s_at

	hgnc_symbol	Affymetrix ID
34	ACTG2	202274_at
35	SERPINF1	202283_at
36	CXCL14	222484_s_at
37	GJA1	201667_at
38	CCDC8	223496_s_at
39	VSNL1	203798_s_at
40	MMP1	204475_at
41	IGFBP2	202718_at
42	CTHRC1	225681_at
43	IRX4	220225_at
44	CCDC3	223316_at
45	LIMA1	217892_s_at
46	TCF4	212386_at
47	SLC7A5,SLC7A5P1	201195_s_at
48	S100A2	204268_at
49	TMTC1	224397_s_at
50	SLC38A1	224579_at
51	TPM2	204083_s_at
52	JAM3	212813_at
53	SPARCL1	200795_at
54	SGCE	204688_at
55	GPX2	202831_at
56	ANO1	218804_at
57	CCND2	200953_s_at
58	SRPX	204955_at
59	CNN1	203951_at
60	BOC	225990_at
61	ANTXR1	224694_at
62	KRT14	209351_at
63	THY1	213869_x_at
64	NNMT	202237_at
65	COL9A3	204724_s_at
66	AEBP1	201792_at

	hgnc_symbol	Affymetrix ID
67	FAM176A	227828_s_at
68	DUSP2	204794_at
69	FEZ1	203562_at
70	MRC2	37408_at
71	RCBTB2	230292_at
72	USP31	1558117_s_at
73	CROT	204573_at
74	SQSTM1	201471_s_at
75	COL6A1	213428_s_at
76	PPP1R14A	227006_at
77	TBC1D2	222173_s_at
78	ARMCX1	218694_at
79	SSBP2	203787_at
80	VIM	201426_s_at
81		226148_at
82	EPAS1	200878_at
83	POLE3	208828_at
84	AKAP1	201674_s_at
85	NQO2	203814_s_at
86	DUSP11	202703_at
87	PPPDE1	212371_at
88	MCAM	211340_s_at
89	CTNNAL1	202468_s_at
90	SERPINH1	207714_s_at

Supplemental Table 5

Means and frequencies of participants' characteristics by cross-classified ER/AR/VDR status (N=1731), Nurses' Health Study (1976-1996)

Characteristic	HR3	HR2	HR1	HR0
N (%)	1006 (58.1)	429 (24.8)	185 (10.7)	111 (6.4)
Age at diagnosis, mean (N), yr	57.6 (1006)	57.3 (429)	55.6 (185)	54.7 (111)
Menopausal status at diagnosis, N*(%)				
Premenopausal	224 (22.6)	94 (22.4)	35 (19.7)	28 (25.7)
Postmenopausal	767 (77.4)	326 (77.6)	143 (80.3)	81 (74.3)
ER status, N* (%)				
Positive	1006 (100.0)	328 (76.5)	22 (11.9)	0 (0.0)
Negative	0 (0.0)	101 (23.5)	163 (88.1)	111 (100.0)
HER2 status, N* (%)				
Positive	52 (5.2)	77 (18.2)	39 (21.2)	9 (8.2)
Negative	946 (94.8)	347 (81.8)	145 (78.8)	101 (91.8)
Nodal involvement, N (%)				
None	686 (68.2)	279 (65.0)	99 (53.5)	69 (62.2)
1 - 3	194 (19.3)	83 (19.4)	50 (27.0)	25 (22.5)
4 - 9	84 (8.4)	36 (8.4)	21 (11.4)	12 (10.8)
≥10	42 (4.2)	31 (7.2)	15 (8.1)	5 (4.5)
Tumor size (cm), N (%)				
≤2	692 (68.8)	260 (60.6)	100 (54.1)	58 (52.3)
> 2	314 (31.2)	169 (39.4)	85 (46.0)	53 (47.8)
Histological grade, N* (%)				
I (low)	252 (25.3)	62 (14.7)	10 (5.5)	2 (1.8)
II (intermediate)	632 (63.4)	234 (55.6)	82 (45.3)	15 (13.5)
III (high)	113 (11.3)	125 (29.7)	89 (49.2)	94 (84.7)
Stage†, N (%)				
I	529 (52.6)	205 (47.8)	66 (35.7)	44 (39.6)
II	326 (32.4)	149 (34.7)	76 (41.1)	47 (42.3)
III	151 (15.0)	75 (17.5)	43 (23.2)	43 (23.2)
Chemotherapy, N* (%)				
Yes	260 (34.4)	146 (47.0)	81 (64.8)	51 (65.4)
No	495 (65.6)	165 (53.1)	44 (35.2)	27 (34.6)
Hormone treatment, N* (%)				
Yes	556 (73.3)	211 (69.0)	47 (38.5)	28 (36.8)
No	203 (26.8)	95 (31.1)	75 (61.5)	48 (63.2)
Radiation treatment, N* (%)				
Yes	337 (44.2)	128 (41.4)	60 (48.4)	34 (44.2)
No	426 (55.8)	181 (58.6)	64 (51.6)	43 (55.8)

*N doesn't add to total because of missing information.

†Stage I=tumor size<=2cm and no nodal involvement;
II=tumor size<=2cm & 1-3 nodes or 2-4cm & 0-3 nodes or 4+cm & 0 nodes;
III=tumor size<=2cm & 4+ nodes or 2-4cm & 4+ nodes or >4cm & 1+ nodes.

Supplemental Table 6
Multivariate analysis of breast cancer-specific mortality by HR status
(HR0, HR1, HR2 and HR3); Hazard ratio and 95%CI for breast cancer specific mortality in the Nurses' Health Study by time since diagnosis

	Overall	<5 years		5+ years	
	HR (95%CI)[2]	HR (95%CI)[1]	HR (95%CI)[2]	HR (95%CI)[1]	HR (95%CI)[2]
HR3	1.0 (REF)	1.0 (REF)	1.0 (REF)	1.0 (REF)	1.0 (REF)
HR2	2.9 (1.60-5.21)	2.26 (1.55-3.28)	1.69 (1.14-2.50)	1.20 (0.99-1.58)	1.24 (0.93-1.66)
HR1	5.3 (2.77-9.97)	3.74 (2.47-5.66)	2.44 (1.56-3.84)	1.14 (0.77-1.70)	1.25 (0.81-1.93)
HR0	6.9 (3.37-14.39)	3.57 (2.17-5.85)	2.71 (1.56-4.71)	0.28 (0.12-0.69)	0.34 (0.13-0.86)

1-age adjusted
2-adjusted for age, year of diagnosis, HER2 status, disease stage, grade, radiation treatment, chemotherapy and hormonal treatment

SUPPLEMENTARY MATERIALS AND METHODS

Author contributions

TAI (hypothesis, concepts and project supervision), TAI and SS (experimental design), TAI, SS, AE and GM (multiplex marker analysis), SS, RH and RMT (epidemiological and statistical analysis), SS, CH, MS and AC (bioinformatic analysis), TAI, BW and AT (cell line analysis), TW (immunostains), SS, DK, and TAI (histopathologic analysis), AR and SJS (tissue microarrays), SS and TAI wrote the paper.

Tissue samples

Paraffin blocks from surgical resection specimens of normal breast tissue and of breast tumors were obtained from the archives of Brigham and Women's Hospital (BWH) in accordance with the regulations for excess tissue use stipulated by the BWH institutional review board. The study was conducted according to the principles outlined in the Declaration of Helsinki. A tissue microarray (TMA) of triple negative tumors (HTMA114) was previously described (15). TMAs BRC1501 and BRC1502 were purchased from Pantomics (Richmond, Ca). HER2 positive tumors were defined by IHC (expression >6 on scoring scale described below). A TMA of normal breast tissue (mean age 34.6 years old, range 18-56 years old) (23) and TMAs of samples from the Nurses' Health Study were previously described (24). We thank the participants and staff of the NHS cohort for their valuable contributions. We thank the following state cancer registries for their help: AL, AZ, AR, CA, CO, CT, DE, FL, GA, ID, IL, IN, IA, KY, LA, ME, MD, MA, MI, NE, NH, NJ, NY, NC, ND, OH, OK, OR, PA, RI, SC, TN, TX, VA, WA, and WY. We thank Terri Woo for expert assistance with IHC.

Immunohistochemistry and immunofluorescence of tissues

Deparaffinized sections were blocked with 3% H2O2, antigen retrieval was performed using a pressure cooker with Dako citrate buffer (pH 6.0) at 120 ºC +/- 2 ºC, 15 +/- 5 PSI, slides were blocked with 3% serum and incubated with primary antibody (indicated dilutions in Table X) at room temperature for 40 minutes. Primary antibody application was followed by 30 minute incubation with Dako Labeled Polymer-HRP as a secondary antibody, and visualized with 3, 3' - diaminobenzidine (DAB) as a chromogen (Dako Envision+ System). Mayer-hematoxylin was the counterstain. Immunostained sections were reviewed by light microscopy and scored visually with a value assigned to each individual core.

Immunofluorescence was performed using similar conditions but with fluorescence labeled secondary antibodies conjugated with fluorescein isothiocyanate (FITC), Texas Red, Cy3, Cy5, or AlexaFluor dyes, and reviewed by standard fluorescence microscope.

In Figure 4, immunostained sections were scored independently by three pathologists (TAI, DK and SS) using light microscopy and a 0 to 25 scale. The percent of tumor cells staining was quantified as (0) = 0%, (1+) =1-20%, (2+) = 21-40%, (3+) = 41-60%, (4+) = 61-80%, and (5+) = 81-100%. The intensity of staining was quantified 0 to 5. The total score was calculated by multiplying the percent score with the intensity score. See supplemental table 3 for scores.

In figure 5 (Nurses' Health Study) there were four cores for each of the 1,731 patients (6,924 cores in total). These TMAs were stained with ER, PR, HER2, VDR, AR, [K8/18/Cld-4], K5/6, and [SMA/p63/CD10] antibodies and scored semi-quantitatively. Given the enormous number of cores to be scored (1,731 x 4 x 8 = 55,392) and the bimodal expression pattern observed in pilot studies (Figure 4) we proceeded with a binomial scoring system with a 1% cut off point in this study. Scoring was 0 for no staining (<1%), 1 for positive staining. The pathologists were blinded to the scores given by the other pathologist and to survival outcomes. Scoring averages were determined per case using values assigned to all interpretable cores from the two pathology readings. If diagnostic tissue was absent or if the staining was not interpretable for all three cores, the case status was recorded as missing. ER, PR and HER2 status of each case was as previously described (25).

Analysis of multiplex immunofluorescence

It has been very difficult to immunostain FFPE tissue sections with more than 3 different antibodies simultaneously for several reasons that involves limitations due to cross channel fluorescence of conjugates, incompatibility of primary antibody hybridization conditions and species limitation of secondary antibodies (26-28).

(1) While the excitation and emission wavelengths for fluorescence conjugates such as fluorescein isothiocyanate (FITC), rhodamine, Texas Red, Cy3, Cy5, and AlexaFluor dyes are theoretically non-overlapping, bleed through between channels is not uncommon. Thus simultaneous multiplexing requires use fluorescence conjugates with very different excitation and emission wavelengths, which limits the number of fluorescent probes that can be simultaneously used to 3 or 4, for all practical purposes.

(2) Primary hybridization of tissues with single primary antibodies vs. with 3-4 different antibodies does not always produce the same results. This is because in some cases the antigen retrieval conditions are different for each antibody or optimum hybridization conditions may require different temperature or time. Thus one may have to use conditions that would work for all the antibodies simultaneously, which is many times is suboptimal for some of the antibodies if not all (29).

(3) Most secondary antibody selectivity is species based, which also limits the number of secondary antibodies that can be multiplexed simultaneously, since most well characterized primary antibodies are produced in mouse, rabbit, rat or chicken. Thus, for all intents and purposes these put an upper limit to simultaneous multiplexing.

To solve these problems we used a sequential multiplex immunofluorescence method developed at GE Global Research Center (Niskayuna, NY) where two antibodies are hybridized to the tissue section at a time and images are collected with two laser exciting the tissue in two channels. Thus, this approach allows using the optimum antigen retrieval and hybridization conditions for each antibody and there is ostensibly no cross channel bleed-through of fluorescence (30). Briefly, IF-based sequential multiplexing method requires direct conjugation of primary antibodies with either Cy3 or Cy5 dye. Tissue sections were sequentially stained with two primary antibodies labeled with Cy3 and Cy5 for each round of staining, images were then acquired and then the fluorescent labels were inactivated. Another round of staining was then performed with another two Cy3 and Cy5 labeled antibodies. This process was repeated until all antibodies of interest were stained and images acquired. For this study, total of 12 antibodies were multiplexed on the same tissue section. For more information about this multiplex staining system see (31) and *http://www.pathinformatics.pitt.edu/sites/default/files/pathinfo/content/Gerdes-Rittscher-PathInfo_HIMA-2011.pdf or http://www.gehealthcare.com/euen/oncology/esmo/pdf/How-MT-PI-can-work-together.pdf*

Image analysis was performed with MetaMorph 7.7 (Molecular Devices, Downingtown, PA). The results of the analysis of luminal epithelial cells are reported. The DAPI image was used to generate a nuclear mask image. A DAPI intensity threshold was set to identify all cell nuclei, and cut/join tools were used to outline each nucleus. Next, keratin negative areas were masked out to eliminate signals from stromal cells and the SMA image was used to mask myoepithelial cells. Integrated Morphometry Analysis / Measure Objects (IMA/MO) was used with minimum object area of 70 pixels to eliminate objects too small to be complete cell nuclei. The resulting Measured Objects image was binarized to generate a nuclear mask. The nuclear mask image was binarized, then dilated (neighborhood 2, repeat count 5, dilate without closing) to generate a binary cell mask image. The binary cell mask image was then applied to all eleven image planes using Arithmetic: source channels image stack and binary cell mask. The resulting image had the starting intensity values for every pixel the mask image was ‘on’, zero value for every pixel the mask image was ‘off’.

To ensure that every object appeared in every channel, we converted the MetaMorph binary cell mask image to a conventional 16-bit image with values 1 (where binary was ‘on’) and 0 (where binary was ‘off’). This image was then added to each of the image

channels, which resulted in equal minimum possible intensity value for each channel-object in each plane, and all image planes had identical objects. Next the data was exported to Microsoft Excel with threshold to outputs for object area and intensity (1 to 65,535). To facilitate analysis of double (HR2) and triple (HR3) hormone receptor positive cells, we used MetaMorph to generate images featuring HR2 and HR3 cells. Starting with the eleven plane image stacks of separated cell objects, we duplicated the stack, used Keep Planes command to keep the image plane channels of interest – ER, AR, VDR (nuclear) – and used Stack Arithmetic Average, to generate an image, "HR2" or "HR3", with bright nuclear signal if any of the ER, AR or VDR (nuclear) signals were high. For display purposes, we also found useful to perform Color Combine on some image triplets, ex. Red=ER, Green=AR, Blue=DAPI.

Nurses' Health Study: design and population

The NHS is a prospective cohort study initiated in 1976: 121,700 female US-registered nurses between the ages of 30 and 55 years completed a questionnaire covering factors relevant to women's health with biennial follow-up questionnaires used to update exposure information and ascertain nonfatal incident diseases. Information about the cohort, selection criteria for outcome analysis, covariates evaluated, and the statistical analysis methods are as previously described (24).

Analysis of mRNA expression profiles

Kaplan-Meier analysis of relapse free survival for all invasive breast cancers from the 855 patient breast tumor dataset (16) (https://genome.unc.edu/) was performed by first independently rank ordering all 855 tumors according to gene expression values for ER, AR, and VDR. Next, tumors were classified as 'Positive/High' or 'Negative/Low' for each gene, independent of each other using a 50th percentile cut off. The groups were then assembled based on positive or negative call for each gene; all negative=HR0, one positive=HR1+, two positive=HR2+, all three positive=HR3+.

Characterization of breast cancer cell lines

All the ATCC cells lines we used were authenticated using short tandem repeat (STR) analysis of specific loci and the cell line identity was confirmed. All cell lines were cultured according to ATCC's guidelines. BT20, HCC1187, MDAMB468 and SUM159 cells were trypsinized with 0.05% trypsin/EDTA and plated at appropriate densities (BT20: 6,000, HCC1187: 10,000, MDAMB468: 4,000 and SUM159: 1,000 cells/well) in 96 well plates in DMEM+10%FBS. The next day Calcitriol and Taxol were added at concentrations of 25nM and 1.5nM respectively (Day0). For ZR75B, 4000 cells/well were plated in 96 well plates in DMEM+10%FBS. 50nM Calcitriol and 5nM ICI180,782 were added the next day with fresh medium (Day0). Media was refreshed every two days and cell proliferation assay was performed on day 6 for above cell lines. The T47D cells were trypsinized with phenol red free Triple Express (Invitrogen) and washed once

with PBS and 10,000 cells/well were plated in phenol red free DMEM+ 5% charcoal stripped FBS (CSFBS) in 96 well plates and cultured for 3 days. On the fourth day, 10nM 17 beta estradiol (E2) was added into all the wells except non-drug treatment group, in addition to 10nM ICI180,782, 10nM R1881 and 50nM Calcitriol in different groups with fresh phenol red free DMEM+5% CSFBS (Day0). Cells were cultured in the same medium for 6 days before counting. All the proliferation assays were carried out using cell titer blue reagent (Promega) and fluorescence was measured using Bio-Tek spectrophotometer at 530/25nm (excitation) and 590/35nm (emission). Vehicle only group was used as control in all the experiments except for T47D, where E2 was used as control.

SUPPLEMENTAL REFERENCES

1. Nakshatri, H., Srour, E.F., and Badve, S. 2009. Breast cancer stem cells and intrinsic subtypes: controversies rage on. *Curr Stem Cell Res Ther* 4:50-60.
2. Panchision, D.M., Chen, H.L., Pistollato, F., Papini, D., Ni, H.T., and Hawley, T.S. 2007. Optimized flow cytometric analysis of central nervous system tissue reveals novel functional relationships among cells expressing CD133, CD15, and CD24. *Stem Cells* 25:1560-1570.
3. Kordek, R., Potemski, P., Kusinska, R., Pluciennik, E., and Bednarek, A. 2010. Basal keratin expression in breast cancer by quantification of mRNA and by immunohistochemistry. *J Exp Clin Cancer Res* 29:39.
4. Molyneux, G., and Smalley, M.J. 2011. The cell of origin of BRCA1 mutation-associated breast cancer: a cautionary tale of gene expression profiling. *J Mammary Gland Biol Neoplasia* 16:51-55.
5. Molyneux, G., Geyer, F.C., Magnay, F.A., McCarthy, A., Kendrick, H., Natrajan, R., Mackay, A., Grigoriadis, A., Tutt, A., Ashworth, A., et al. 2010. BRCA1 basal-like breast cancers originate from luminal epithelial progenitors and not from basal stem cells. *Cell Stem Cell* 7:403-417.
6. Lim, E., Vaillant, F., Wu, D., Forrest, N.C., Pal, B., Hart, A.H., Asselin-Labat, M.L., Gyorki, D.E., Ward, T., Partanen, A., et al. 2009. Aberrant luminal progenitors as the candidate target population for basal tumor development in BRCA1 mutation carriers. *Nat Med* 15:907-913.
7. Gusterson, B.A., Ross, D.T., Heath, V.J., and Stein, T. 2005. Basal cytokeratins and their relationship to the cellular origin and functional classification of breast cancer. *Breast Cancer Res* 7:143-148.
8. Gusterson, B. 2009. Do 'basal-like' breast cancers really exist? *Nat Rev Cancer* 9:128-134.
9. Ince, T.A., Ward, J.M., Valli, V.E., Sgroi, D., Nikitin, A.Y., Loda, M., Griffey, S.M., Crum, C.P., Crawford, J.M., Bronson, R.T., et al. 2008. Do-it-yourself (DIY) pathology. *Nat Biotechnol* 26:978-979; discussion 979.
10. Grigoriadis, A., Mackay, A., Reis-Filho, J.S., Steele, D., Iseli, C., Stevenson, B.J., Jongeneel, C.V., Valgeirsson, H., Fenwick, K., Iravani, M., et al. 2006. Establishment of the epithelial-specific transcriptome of normal and malignant human breast cells based on MPSS and array expression data. *Breast Cancer Res* 8:R56.
11. Jones, C., Mackay, A., Grigoriadis, A., Cossu, A., Reis-Filho, J.S., Fulford, L., Dexter, T., Davies, S., Bulmer, K., Ford, E., et al. 2004. Expression profiling of purified normal human luminal and myoepithelial breast cells: identification of novel prognostic markers for breast cancer. *Cancer Res* 64:3037-3045.
12. Desmedt, C., Piette, F., Loi, S., Wang, Y., Lallemand, F., Haibe-Kains, B., Viale, G., Delorenzi, M., Zhang, Y., d'Assignies, M.S., et al. 2007. Strong time dependence of the 76-gene prognostic signature for node-negative breast cancer patients in the TRANSBIG multicenter independent validation series. *Clin Cancer Res* 13:3207-3214.
13. Ivshina, A.V., George, J., Senko, O., Mow, B., Putti, T.C., Smeds, J., Lindahl, T., Pawitan, Y., Hall, P., Nordgren, H., et al. 2006. Genetic reclassification of

histologic grade delineates new clinical subtypes of breast cancer. *Cancer Res* 66:10292-10301.

14. Loi, S., Haibe-Kains, B., Desmedt, C., Wirapati, P., Lallemand, F., Tutt, A.M., Gillet, C., Ellis, P., Ryder, K., Reid, J.F., et al. 2008. Predicting prognosis using molecular profiling in estrogen receptor-positive breast cancer treated with tamoxifen. *BMC Genomics* 9:239.
15. Lu, X., Wang, Z.C., Iglehart, J.D., Zhang, X., and Richardson, A.L. 2008. Predicting features of breast cancer with gene expression patterns. *Breast Cancer Res Treat* 108:191-201.
16. Harrell, J.C., Prat, A., Parker, J.S., Fan, C., He, X., Carey, L., Anders, C., Ewend, M., and Perou, C.M. 2012. Genomic analysis identifies unique signatures predictive of brain, lung, and liver relapse. *Breast Cancer Res Treat* 132:523-535.
17. Birrell, S.N., Bentel, J.M., Hickey, T.E., Ricciardelli, C., Weger, M.A., Horsfall, D.J., and Tilley, W.D. 1995. Androgens induce divergent proliferative responses in human breast cancer cell lines. *J Steroid Biochem Mol Biol* 52:459-467.
18. Naderi, A., and Hughes-Davies, L. 2008. A functionally significant cross-talk between androgen receptor and ErbB2 pathways in estrogen receptor negative breast cancer. *Neoplasia* 10:542-548.
19. Emde, A., Mahlknecht, G., Maslak, K., Ribba, B., Sela, M., Possinger, K., and Yarden, Y. 2011. Simultaneous Inhibition of Estrogen Receptor and the HER2 Pathway in Breast Cancer: Effects of HER2 Abundance. *Transl Oncol* 4:293-300.
20. Lakhani, S.R., Reis-Filho, J.S., Fulford, L., Penault-Llorca, F., van der Vijver, M., Parry, S., Bishop, T., Benitez, J., Rivas, C., Bignon, Y.J., et al. 2005. Prediction of BRCA1 status in patients with breast cancer using estrogen receptor and basal phenotype. *Clin Cancer Res* 11:5175-5180.
21. Raouf, A., Zhao, Y., To, K., Stingl, J., Delaney, A., Barbara, M., Iscove, N., Jones, S., McKinney, S., Emerman, J., et al. 2008. Transcriptome analysis of the normal human mammary cell commitment and differentiation process. *Cell Stem Cell* 3:109-118.
22. Richardson, A.L., Wang, Z.C., De Nicolo, A., Lu, X., Brown, M., Miron, A., Liao, X., Iglehart, J.D., Livingston, D.M., and Ganesan, S. 2006. X chromosomal abnormalities in basal-like human breast cancer. *Cancer Cell* 9:121-132.
23. Garbe, J.C., Pepin, F., Pelissier, F.A., Sputova, K., Fridriksdottir, A.J., Guo, D.E., Villadsen, R., Park, M., Petersen, O.W., Borowsky, A.D., et al. 2012. Accumulation of multipotent progenitors with a basal differentiation bias during aging of human mammary epithelia. *Cancer Res* 72:3687-3701.
24. Santagata, S., Hu, R., Lin, N.U., Mendillo, M.L., Collins, L.C., Hankinson, S.E., Schnitt, S.J., Whitesell, L., Tamimi, R.M., Lindquist, S., et al. 2011. High levels of nuclear heat-shock factor 1 (HSF1) are associated with poor prognosis in breast cancer. *Proc Natl Acad Sci U S A* 108:18378-18383.
25. Dawood, S., Hu, R., Homes, M.D., Collins, L.C., Schnitt, S.J., Connolly, J., Colditz, G.A., and Tamimi, R.M. 2011. Defining breast cancer prognosis based on molecular phenotypes: results from a large cohort study. *Breast Cancer Res Treat* 126:185-192.

26. Gerner, M.Y., Kastenmuller, W., Ifrim, I., Kabat, J., and Germain, R.N. 2012. Histo-cytometry: a method for highly multiplex quantitative tissue imaging analysis applied to dendritic cell subset microanatomy in lymph nodes. *Immunity* 37:364-376.
27. Robertson, D., Savage, K., Reis-Filho, J.S., and Isacke, C.M. 2008. Multiple immunofluorescence labelling of formalin-fixed paraffin-embedded (FFPE) tissue. *BMC Cell Biol* 9:13.
28. Tsurui, H., Nishimura, H., Hattori, S., Hirose, S., Okumura, K., and Shirai, T. 2000. Seven-color fluorescence imaging of tissue samples based on Fourier spectroscopy and singular value decomposition. *J Histochem Cytochem* 48:653-662.
29. Shi, S.R., Shi, Y., and Taylor, C.R. 2011. Antigen retrieval immunohistochemistry: review and future prospects in research and diagnosis over two decades. *J Histochem Cytochem* 59:13-32.
30. Glass, G., Papin, J.A., and Mandell, J.W. 2009. SIMPLE: a sequential immunoperoxidase labeling and erasing method. *J Histochem Cytochem* 57:899-905.
31. Gerdes M.J., Sevinsky C.J., Sood A., Adak S., Bello M.O., Bordwell A., Can A., Corwin A., Dinn S., Filkins R.J., Hollman D., et al., Highly multiplexed single-cell analysis of formalin-fixed, paraffin-embedded cancer tissue. 2013. *Proc Natl Acad Sci U S A*. 10(29):11982-7.

Table S2A K5+ER

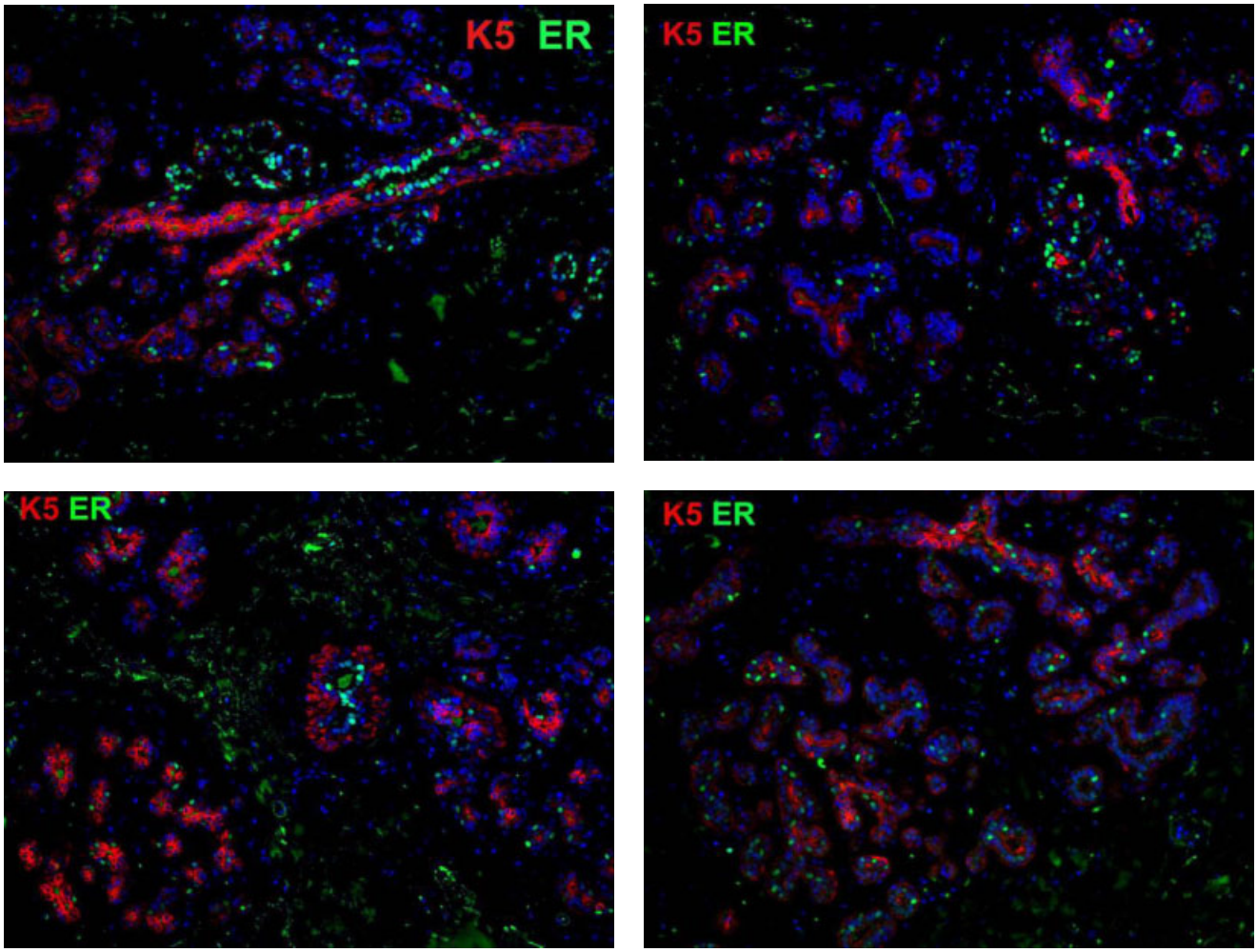
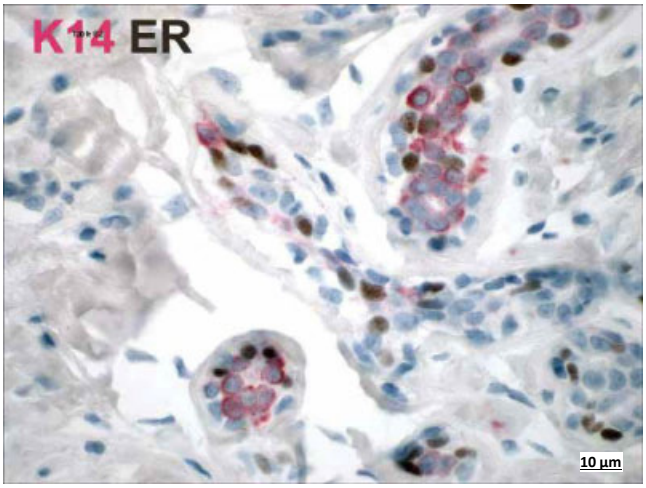
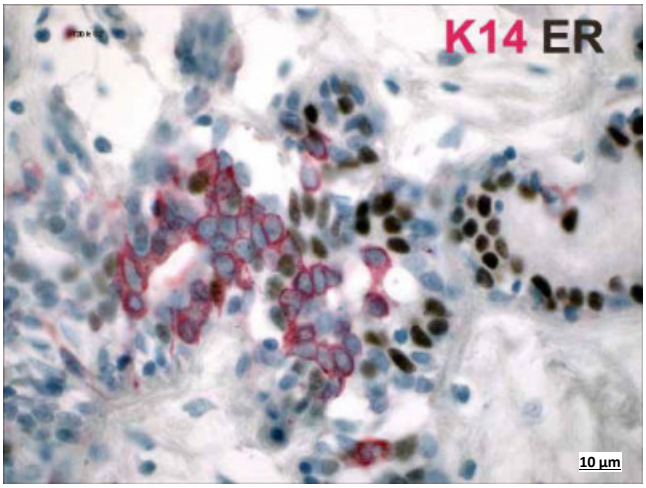
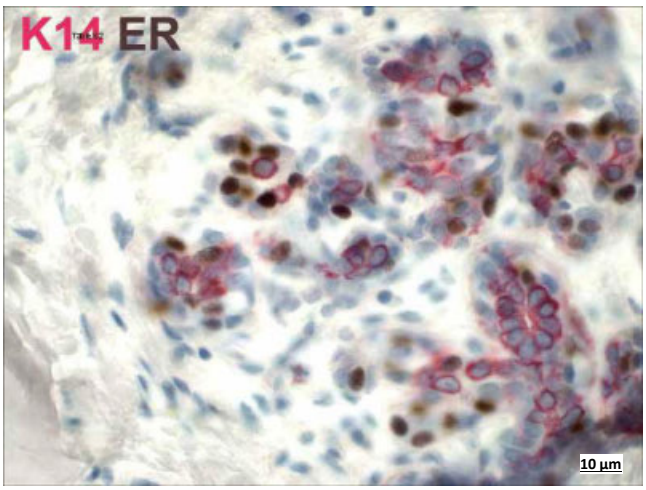
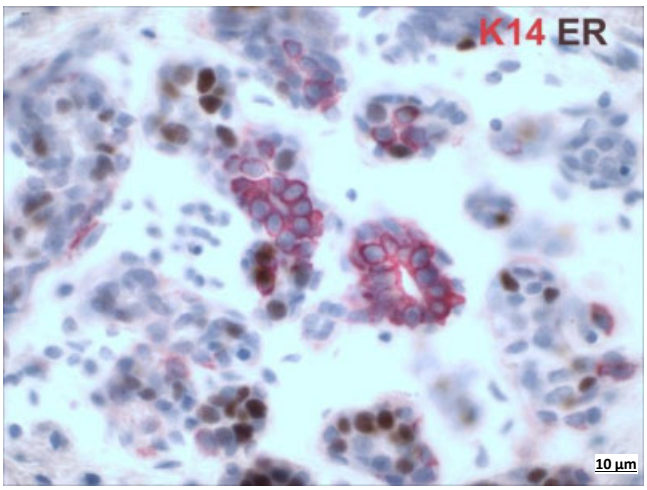


Table S2A K14+ER









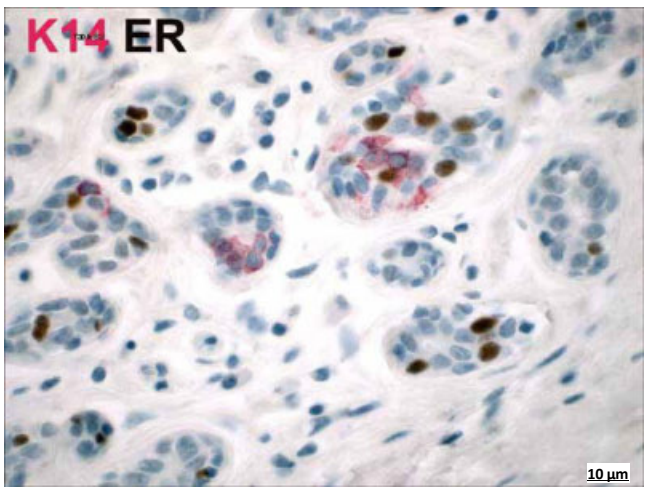
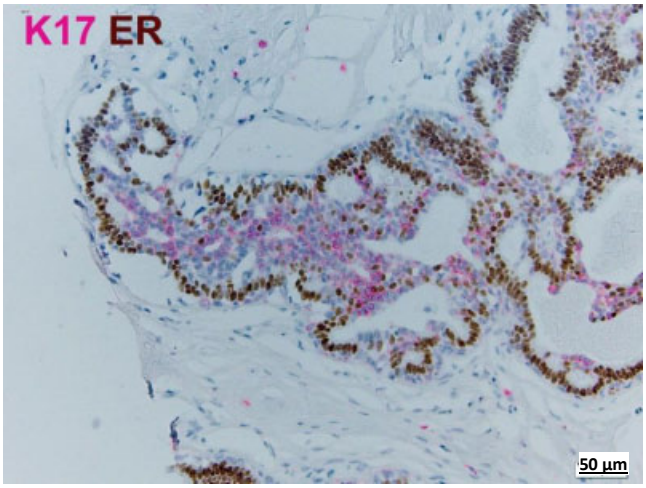
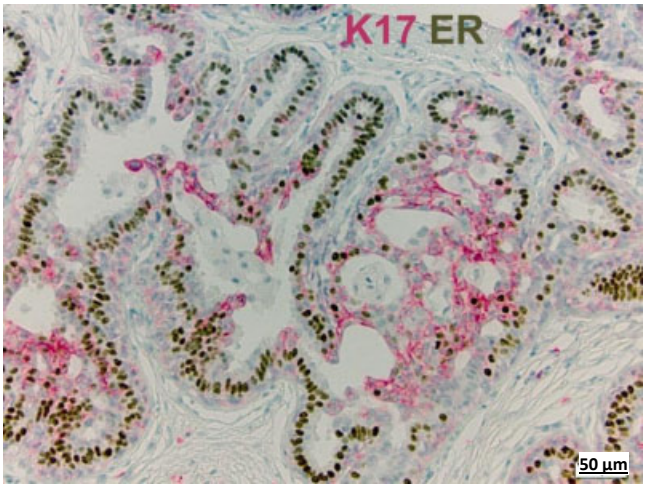
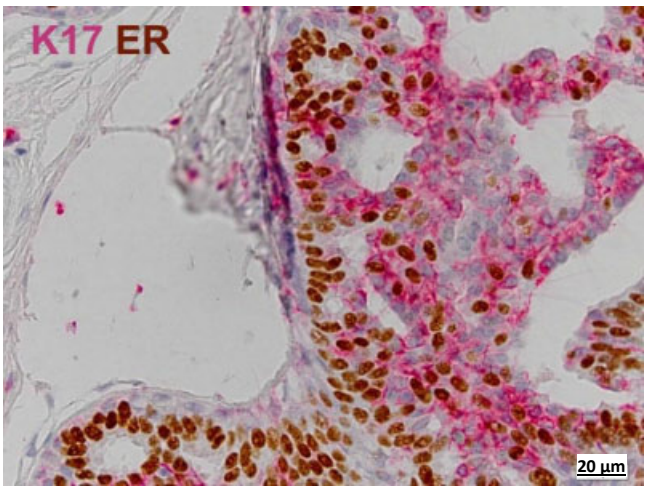
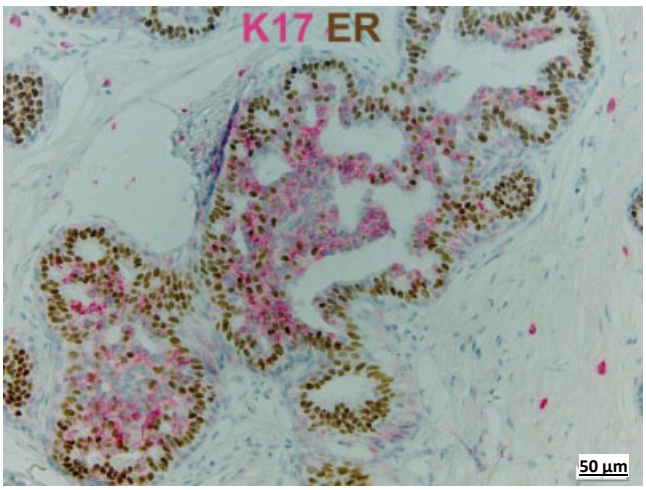


Table S2A K17+ER









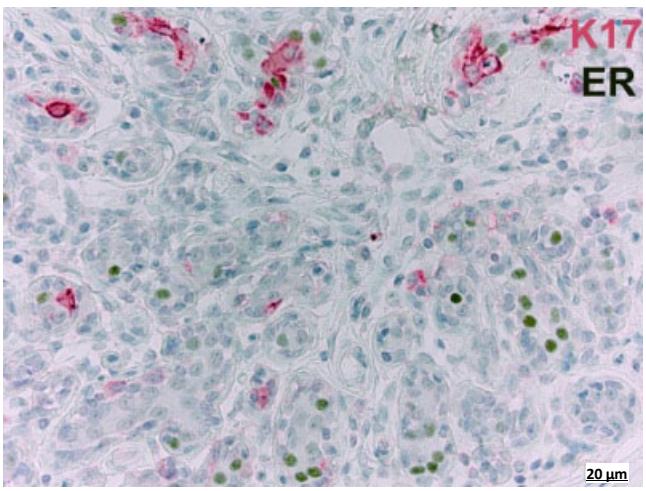
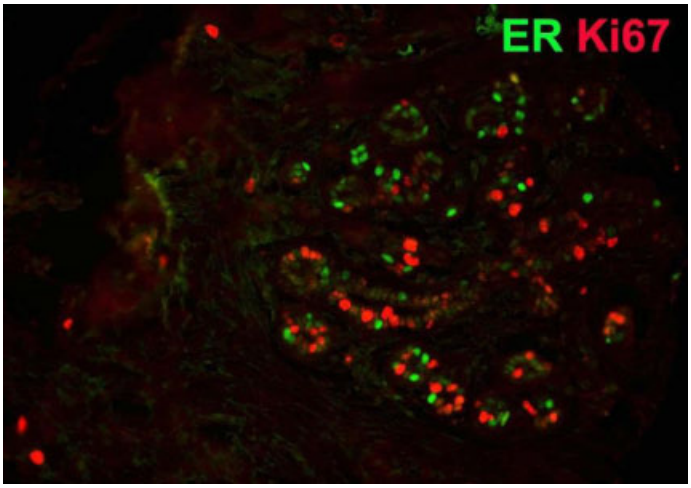
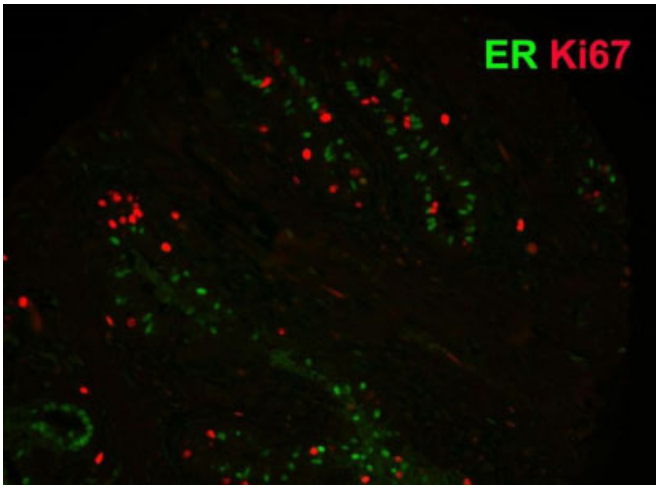
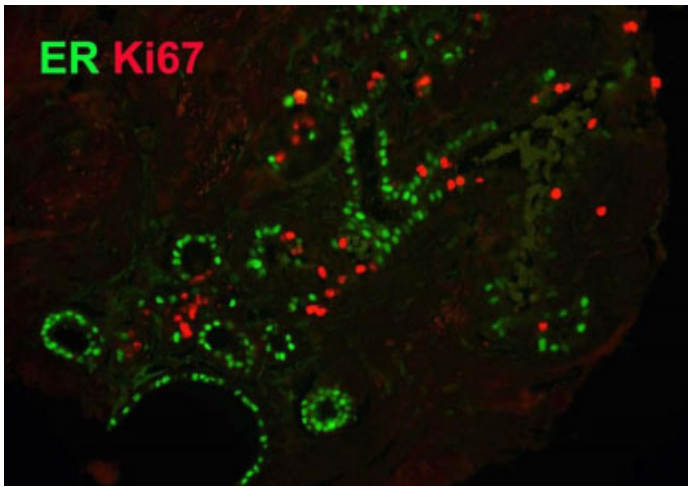
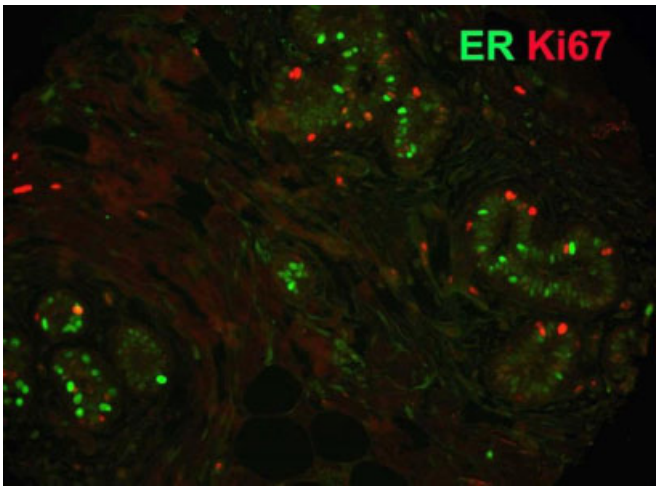


Table S2B ER+Ki67









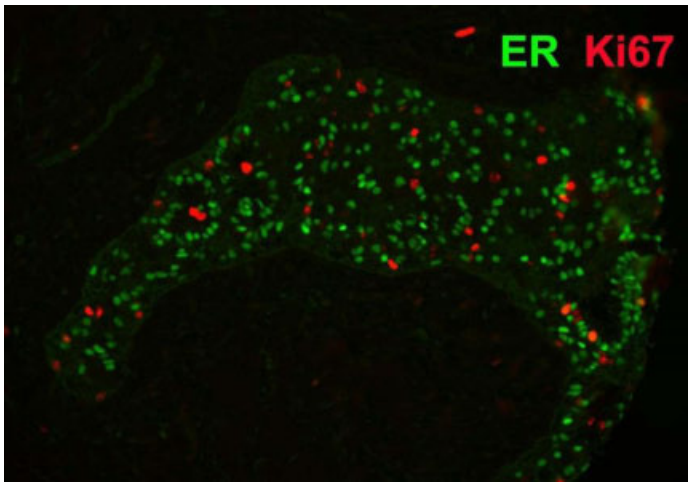
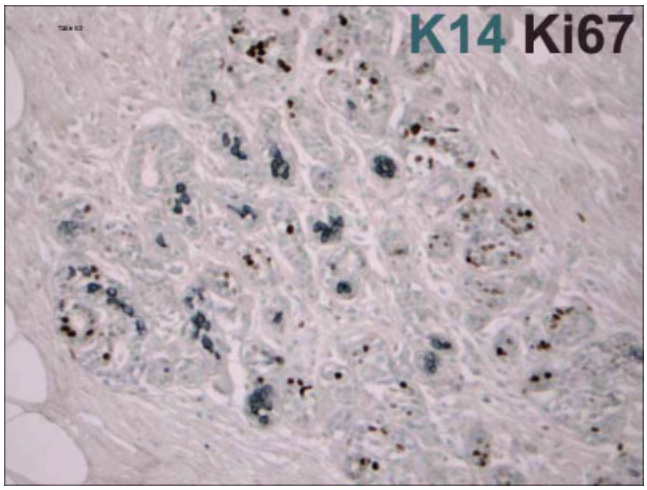
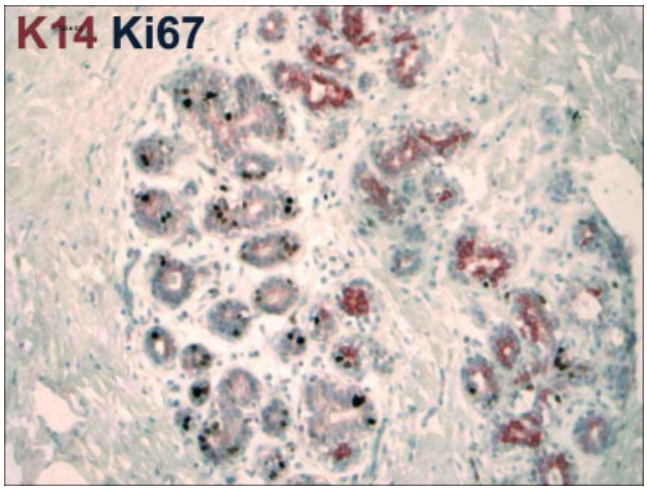
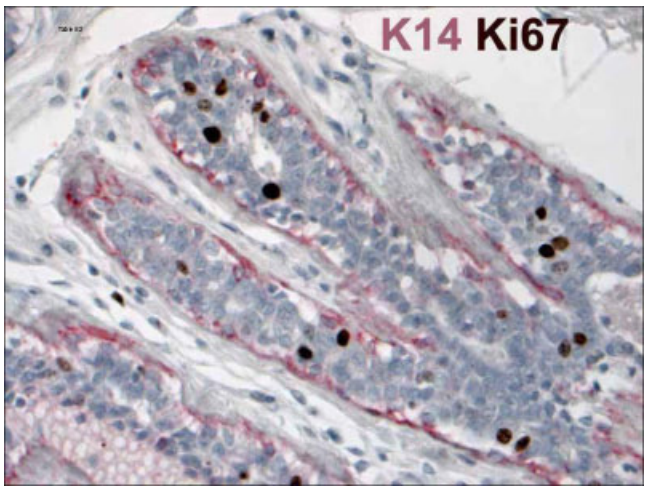
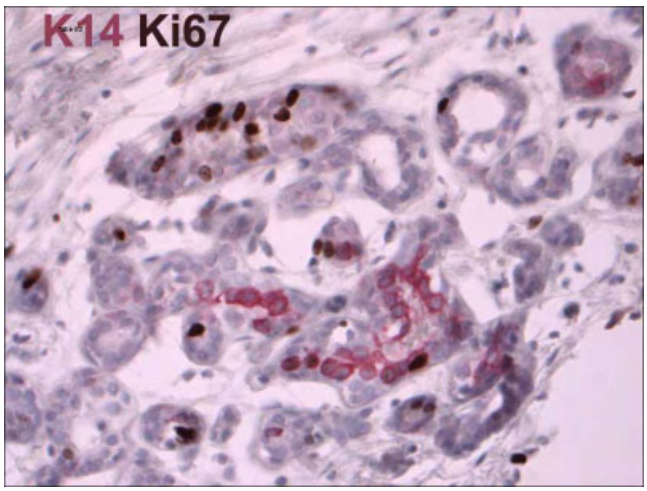


Table S2B K14+Ki67









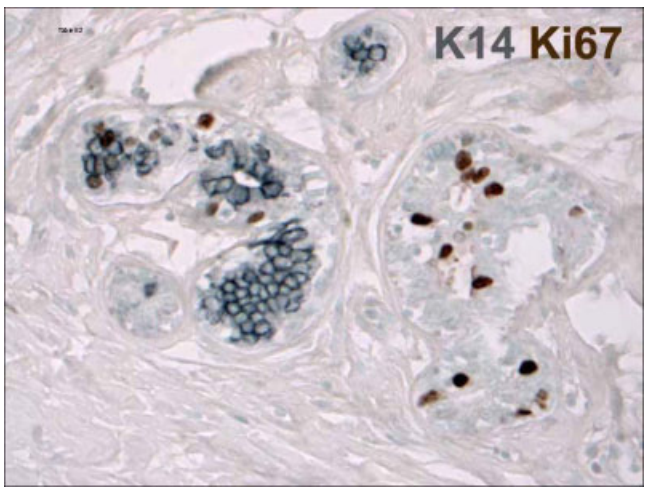
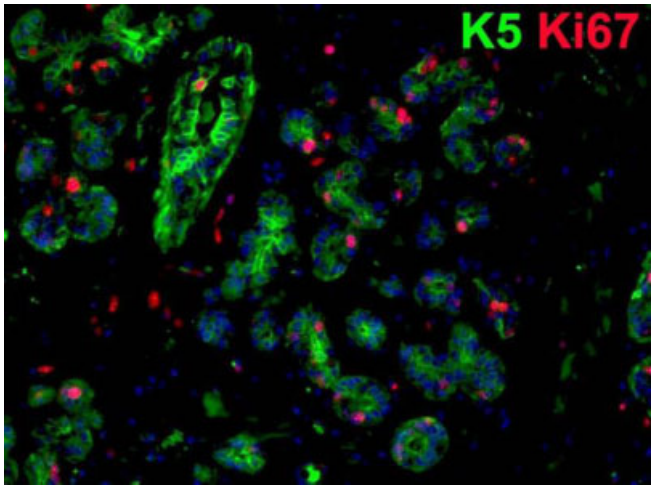
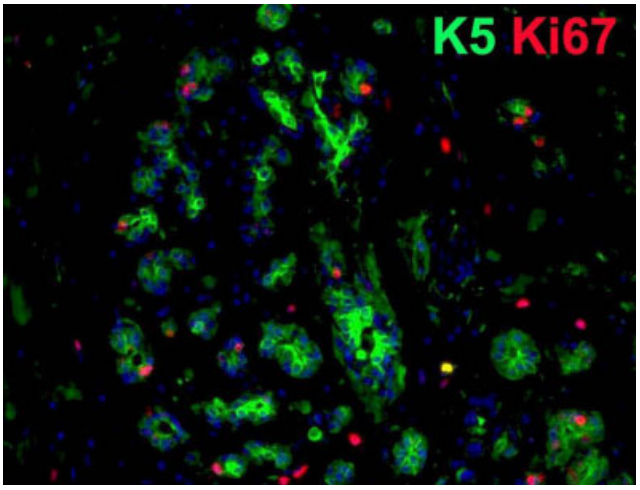
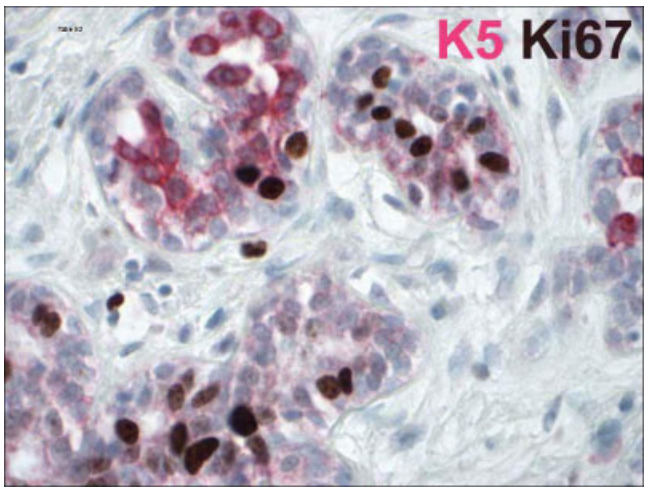
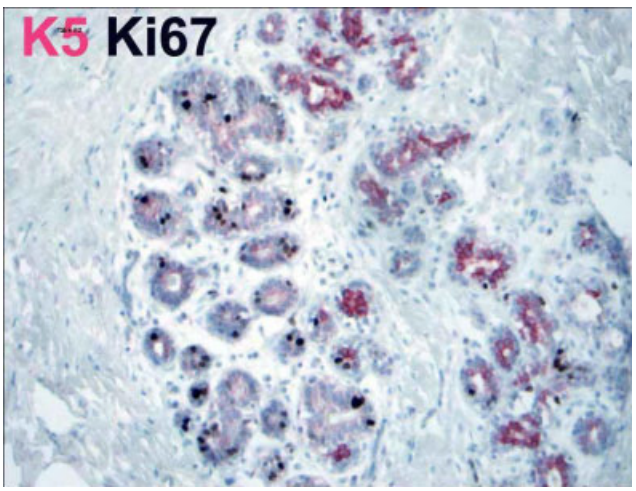


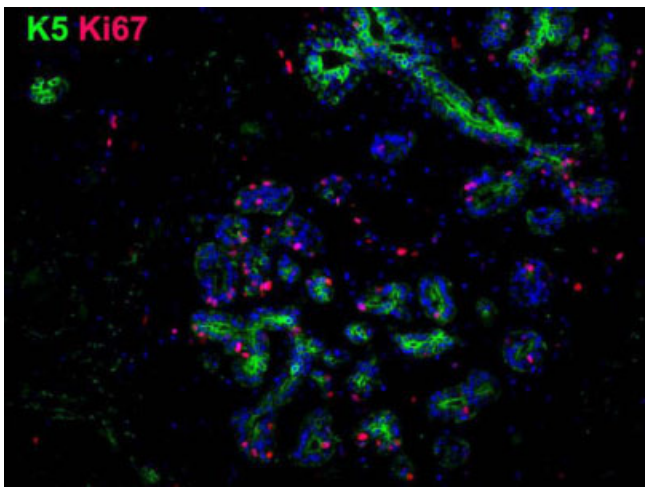
Table S2B K5+Ki67











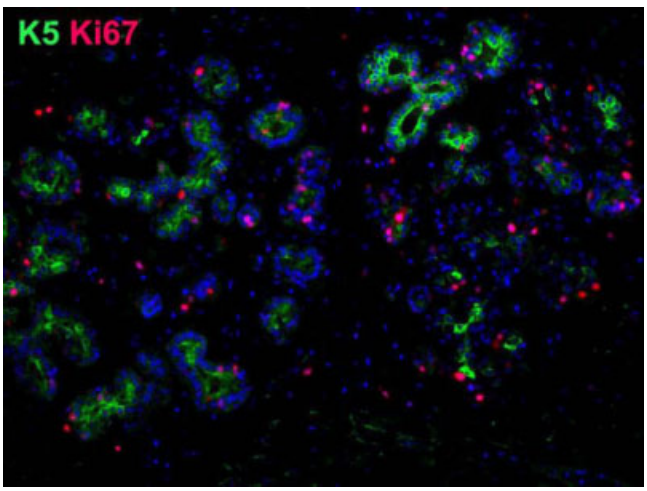
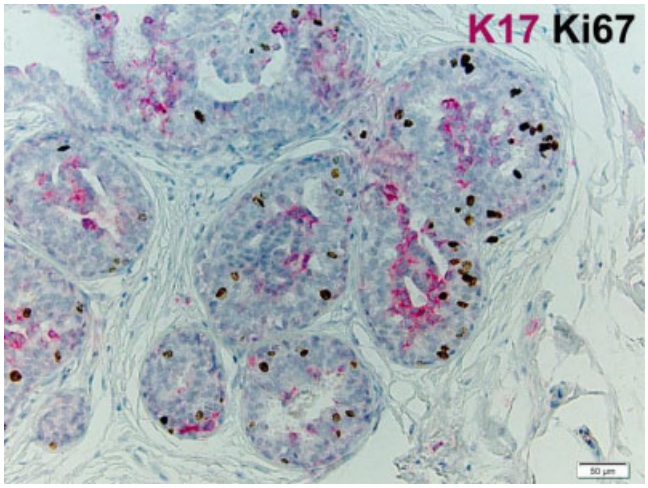
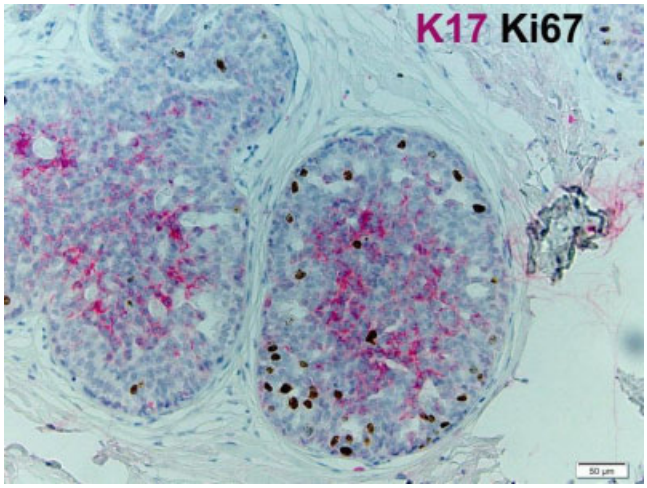
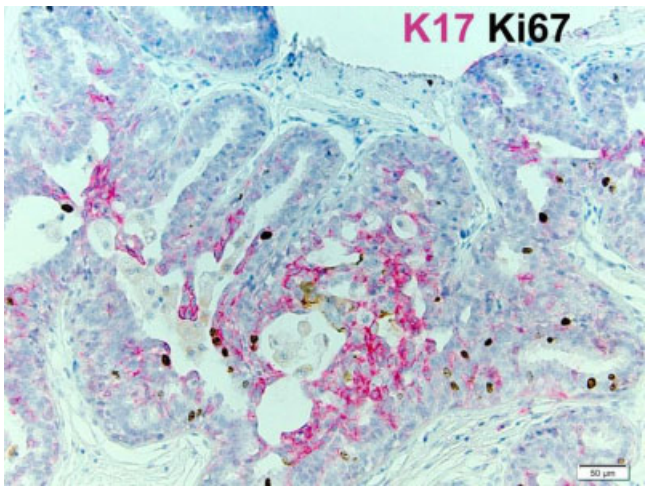
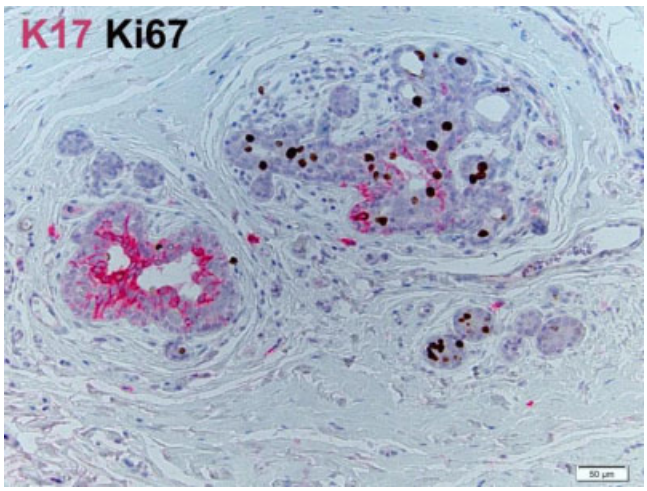


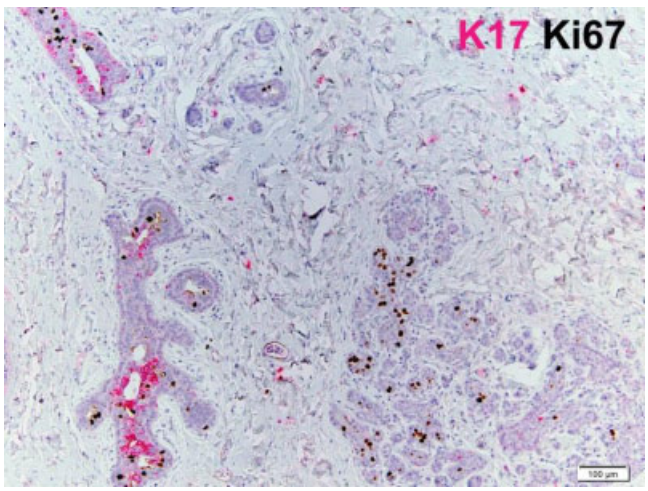
Table S2B K17+Ki67











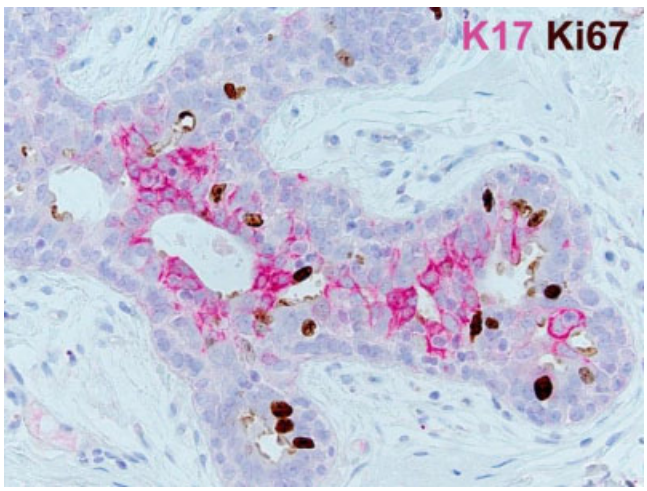
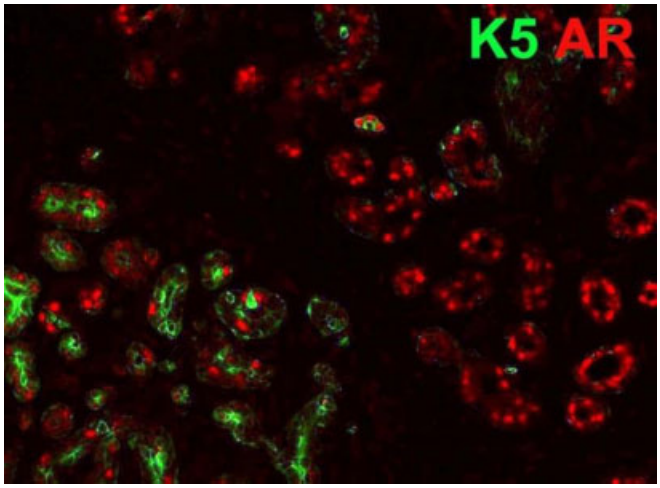
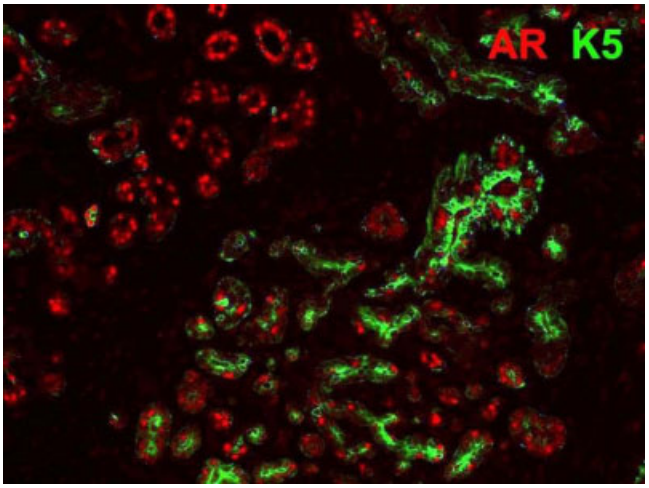
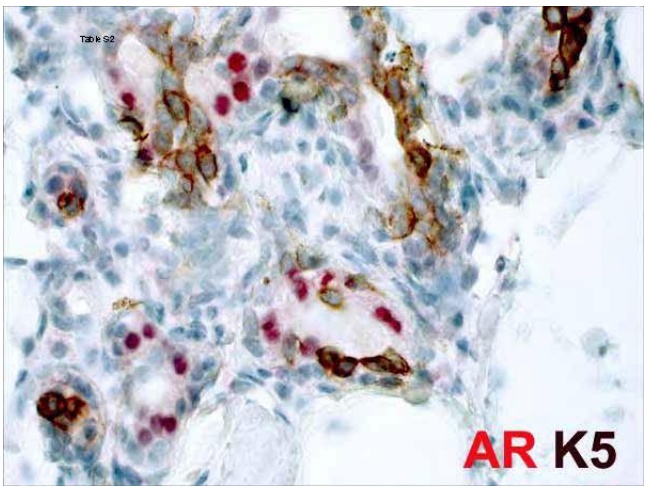
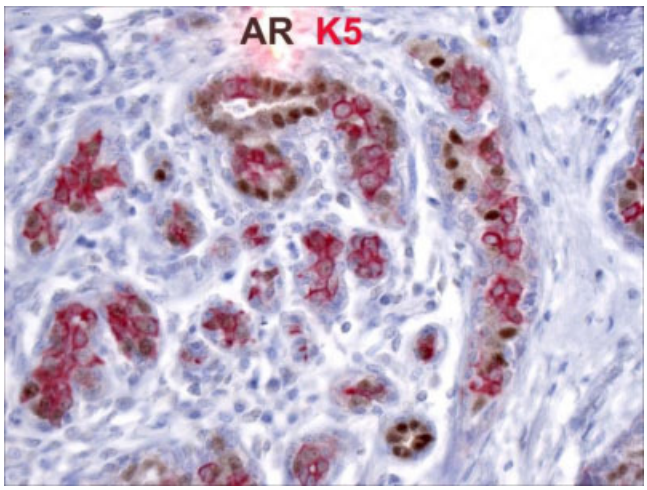


Table S2C K5+AR









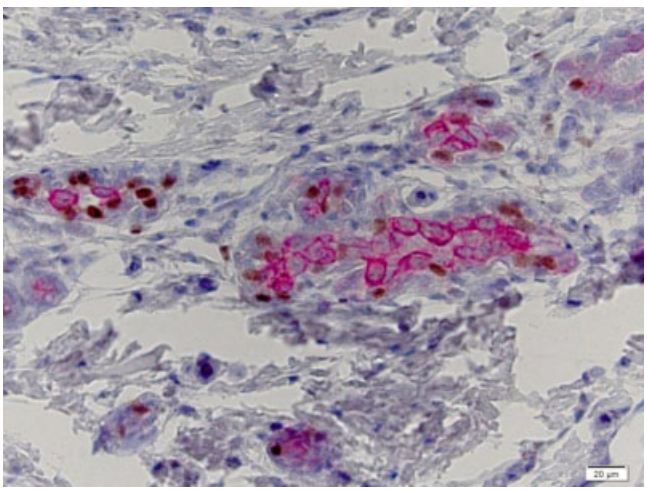
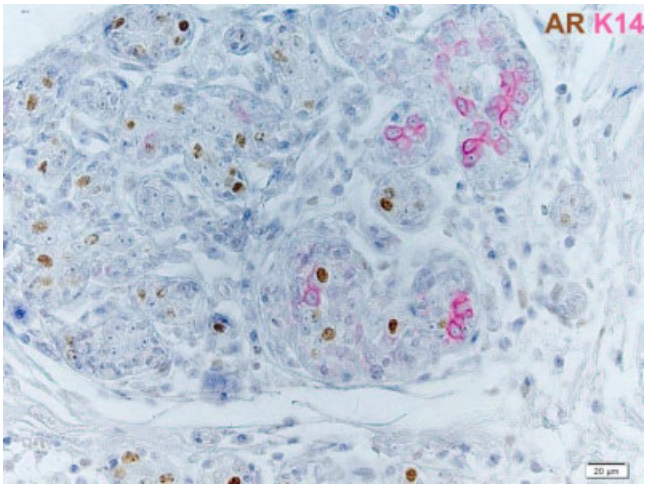
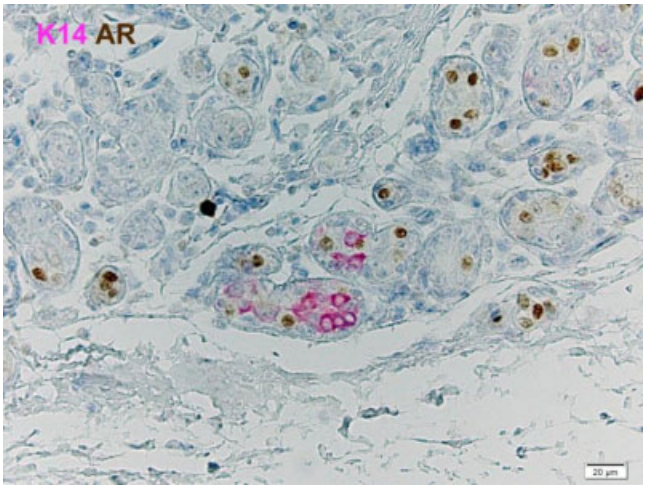
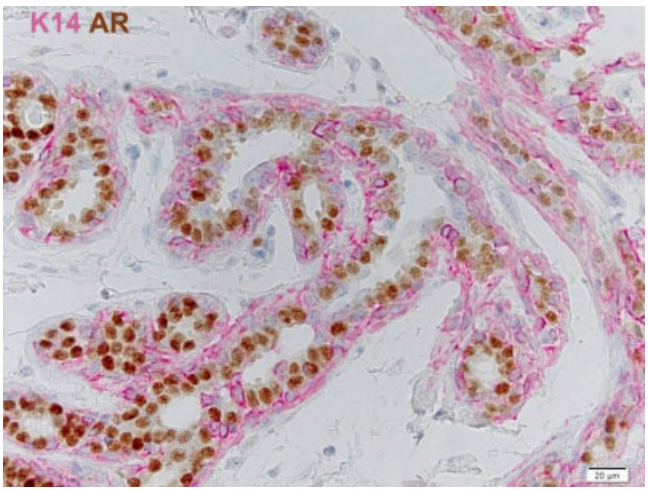
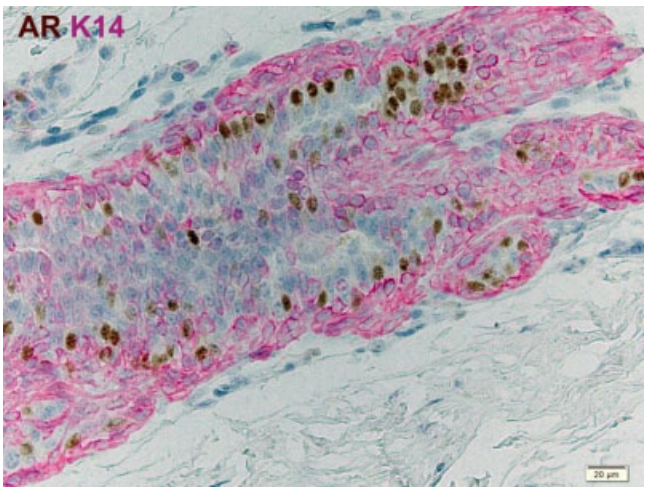


Table S2C K14+AR









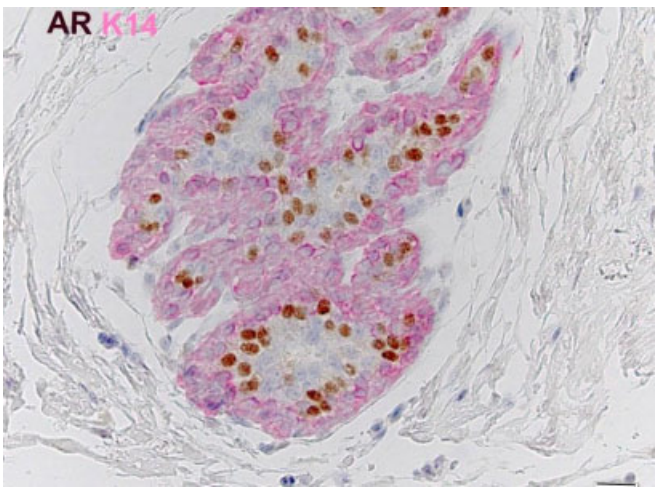
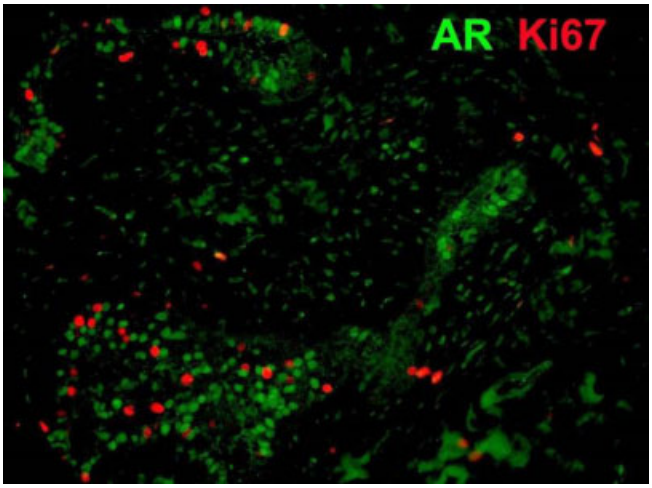
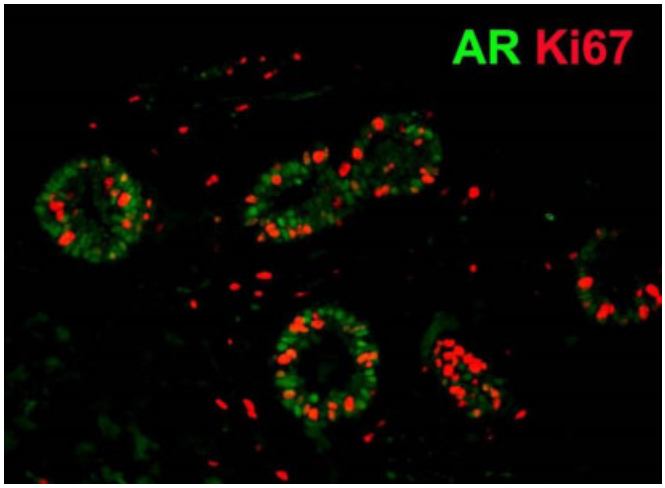
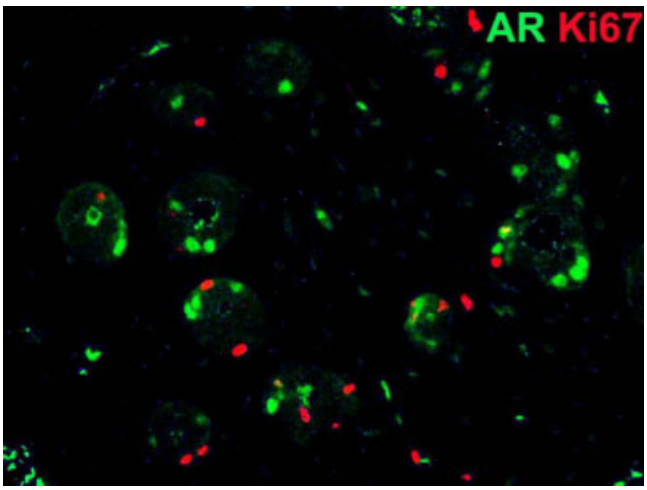
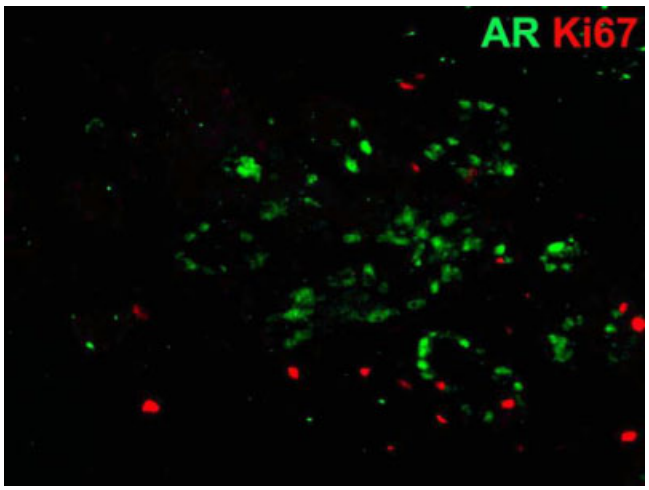


Table S2C AR+Ki67









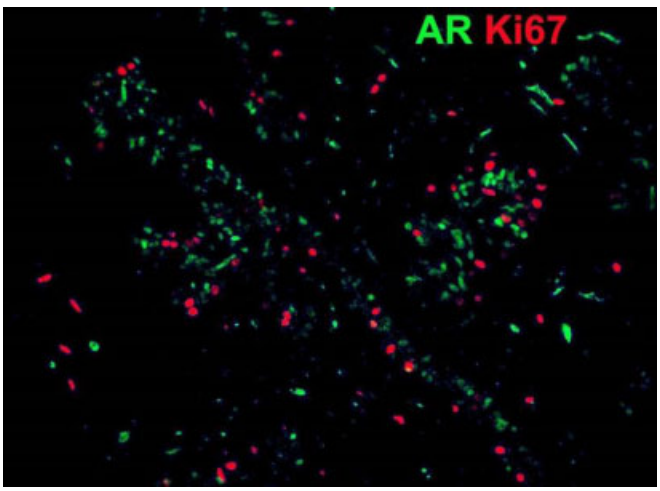
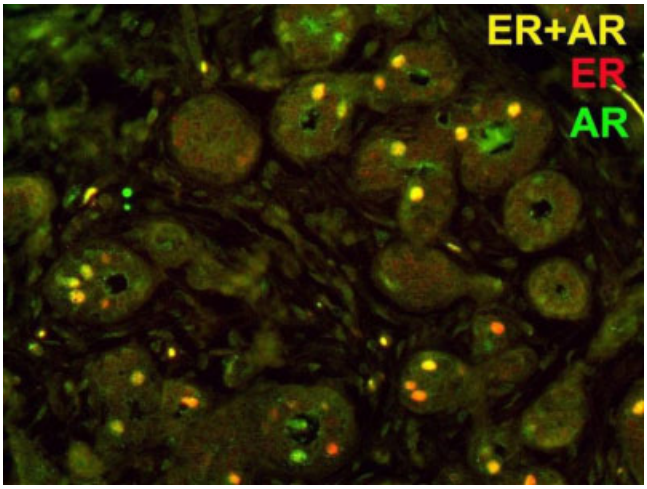
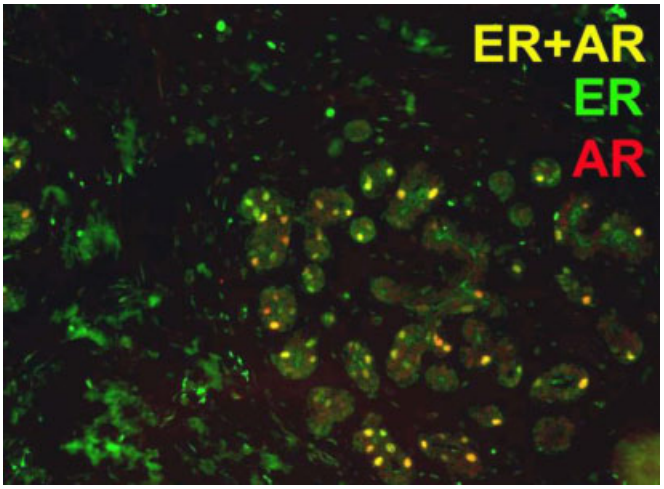
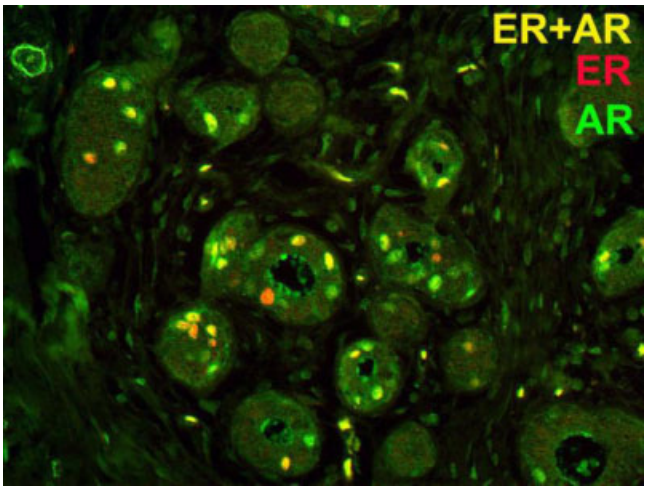
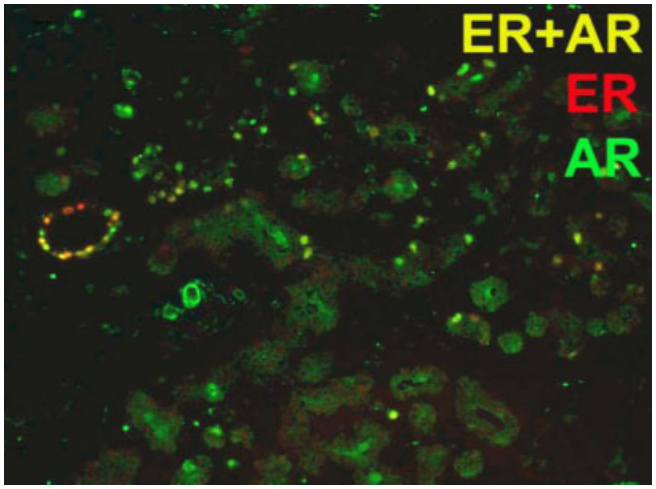


Table S2C ER+AR









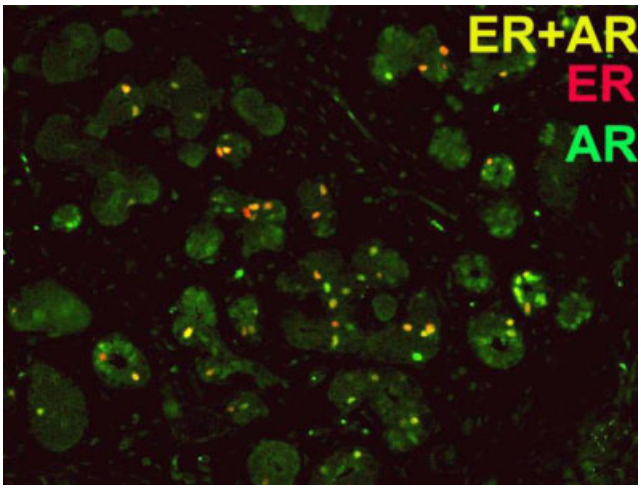
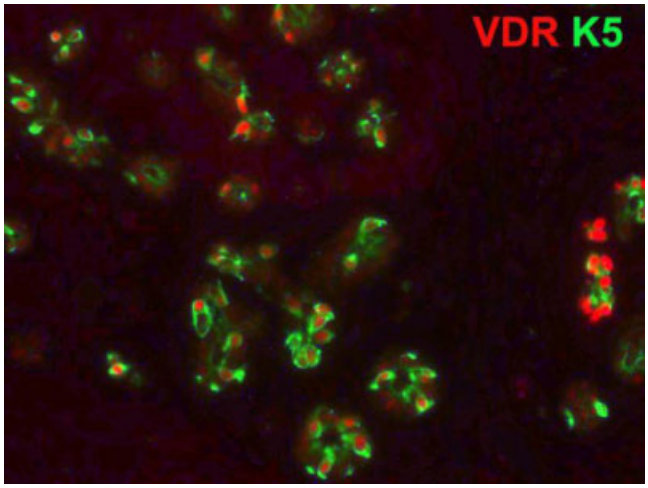
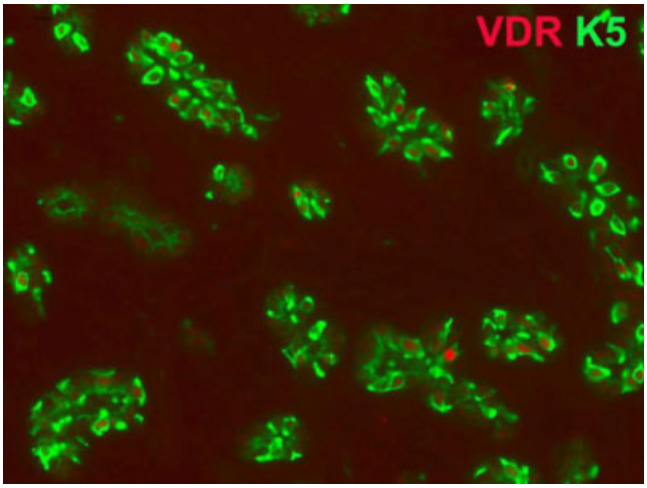
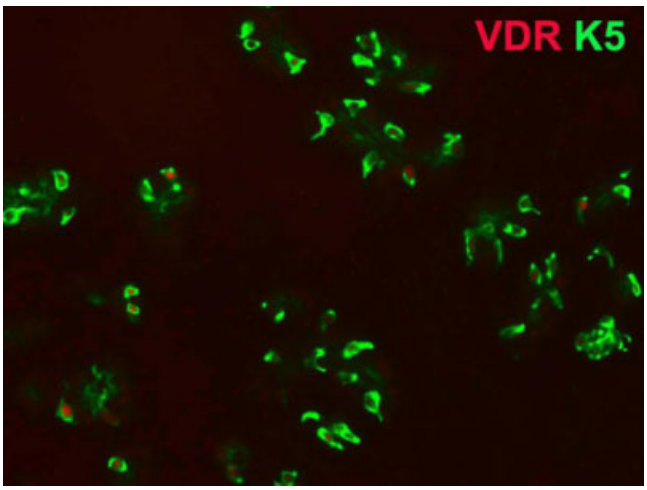
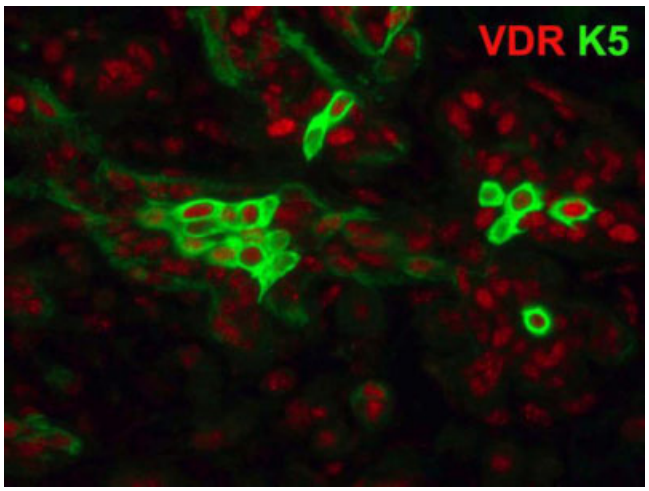


Table S2D K5+VDR









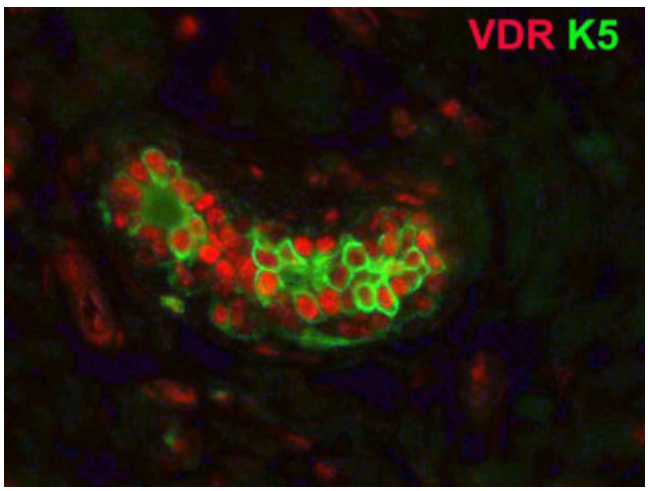
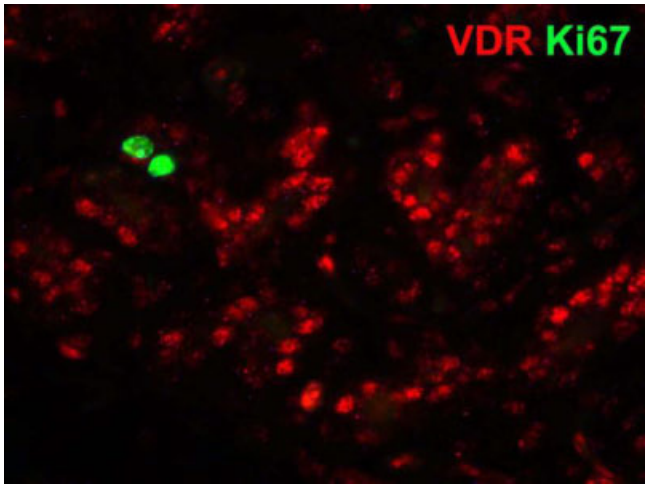
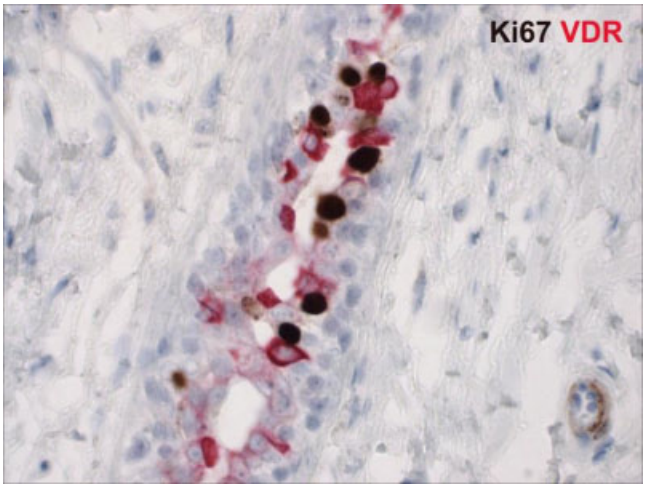
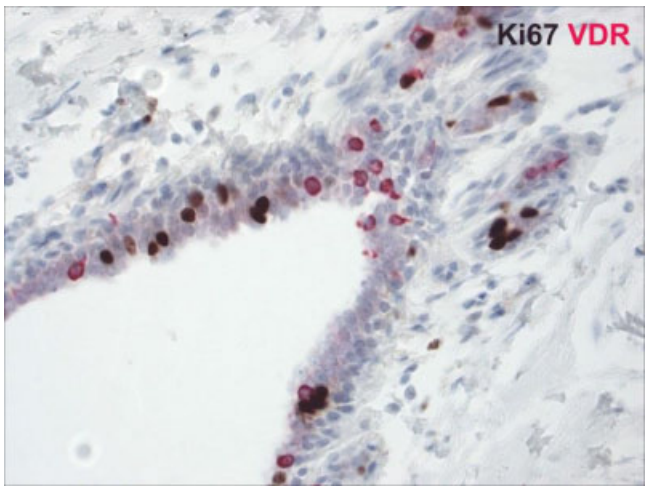


Table S2D VDR+Ki67







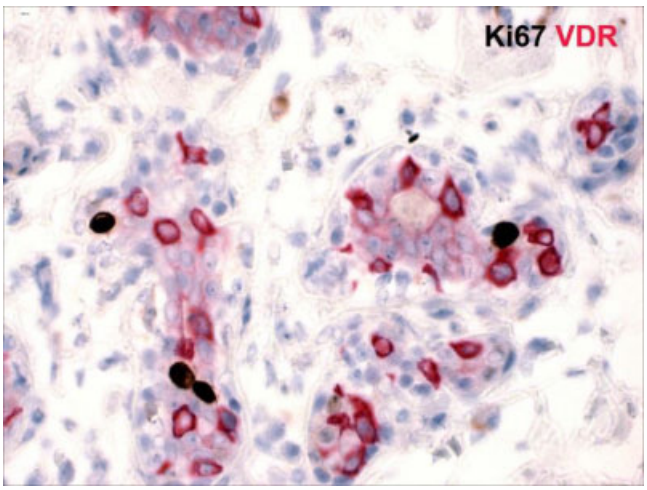
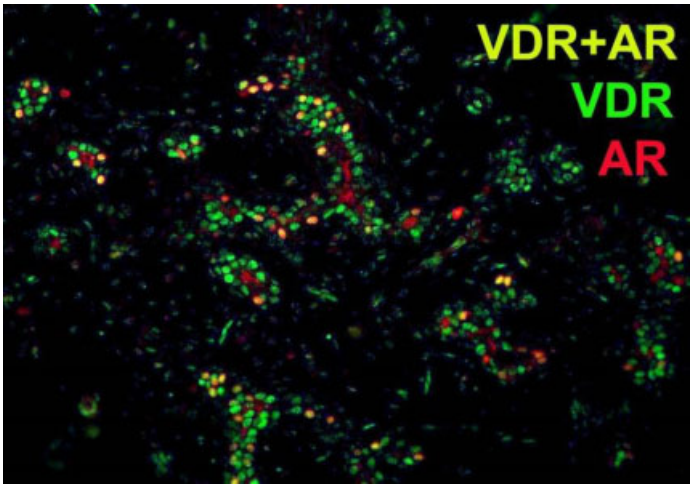
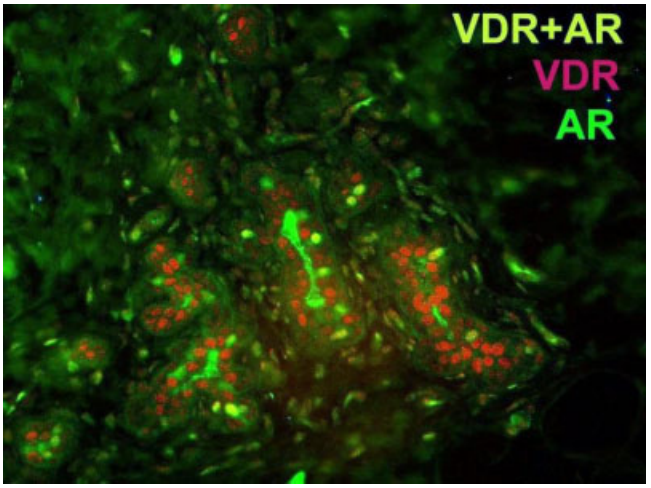
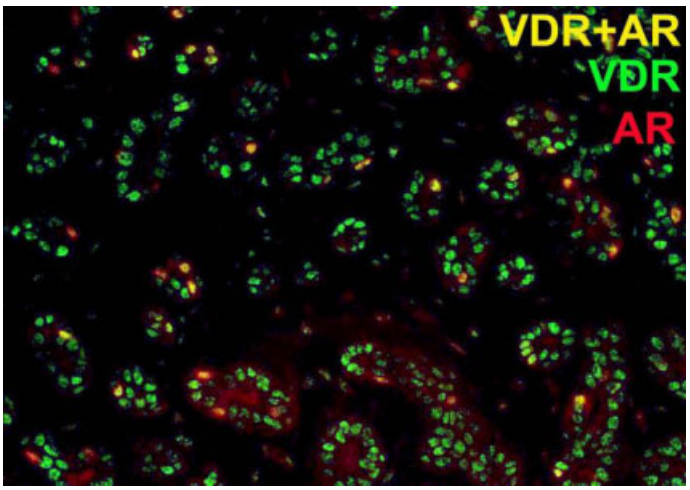
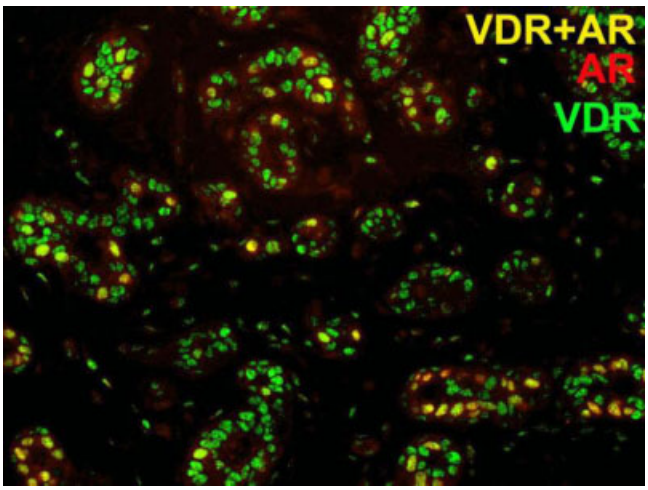


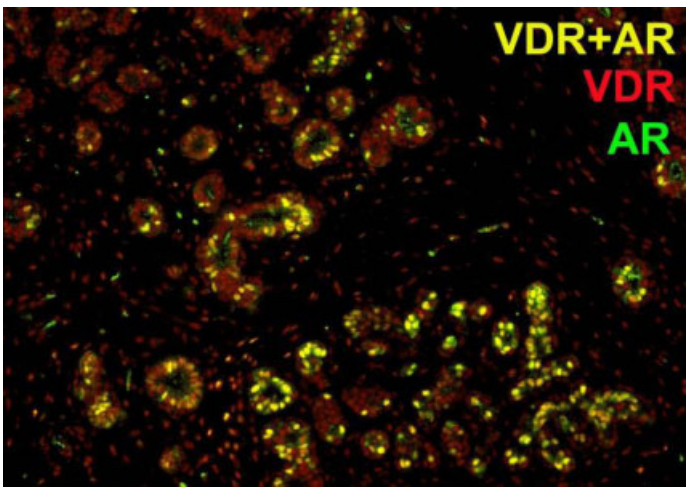
Table S2D AR+VDR











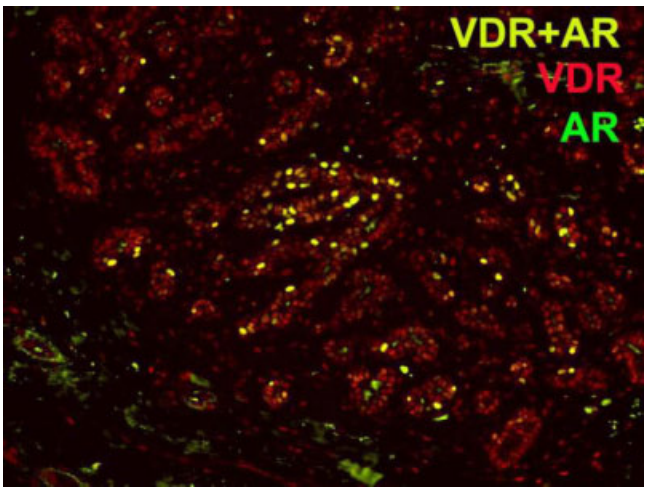
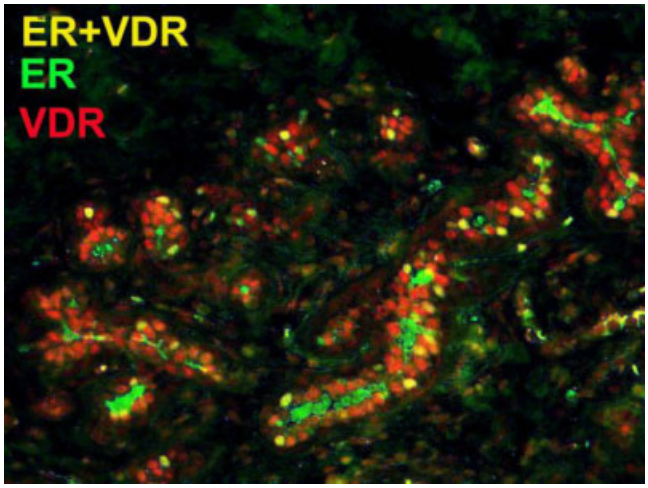
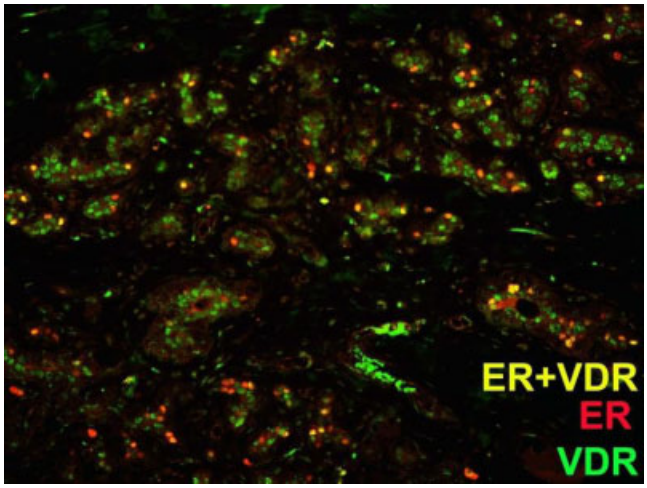
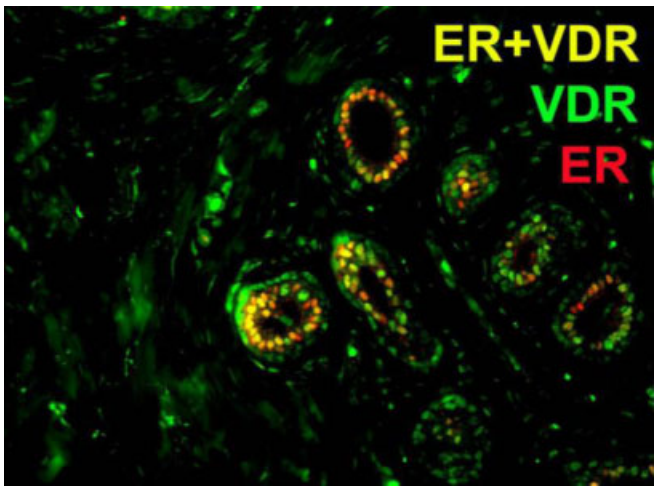
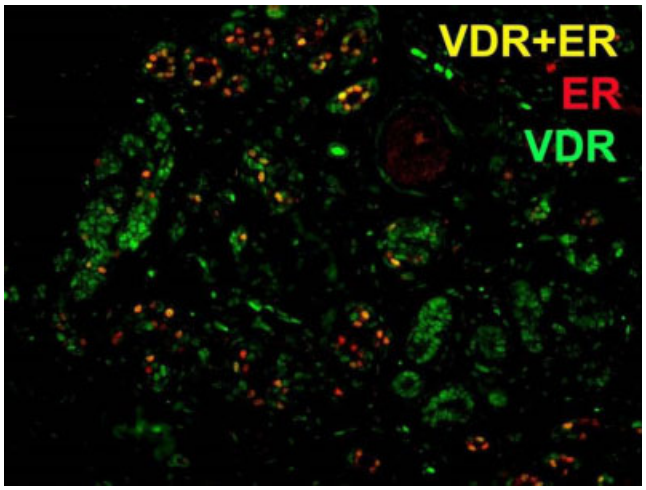


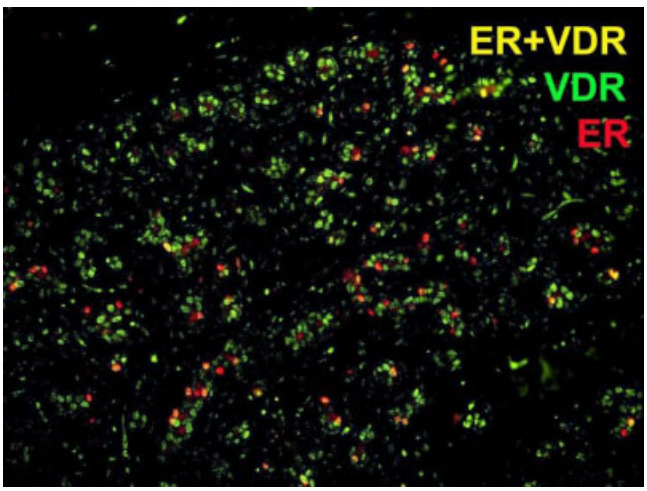
Table S2D ER+VDR











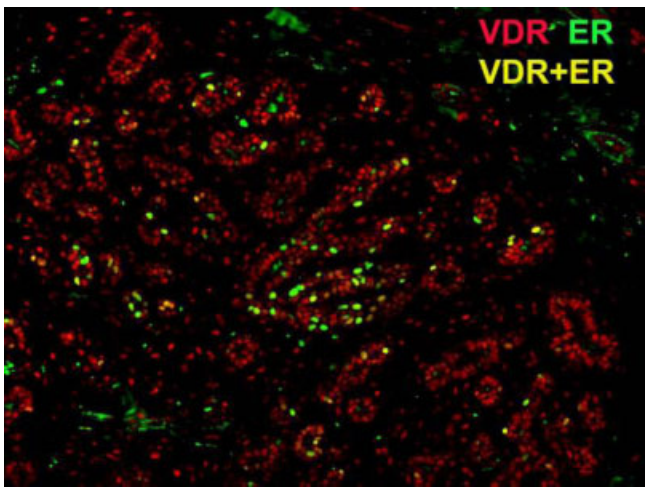
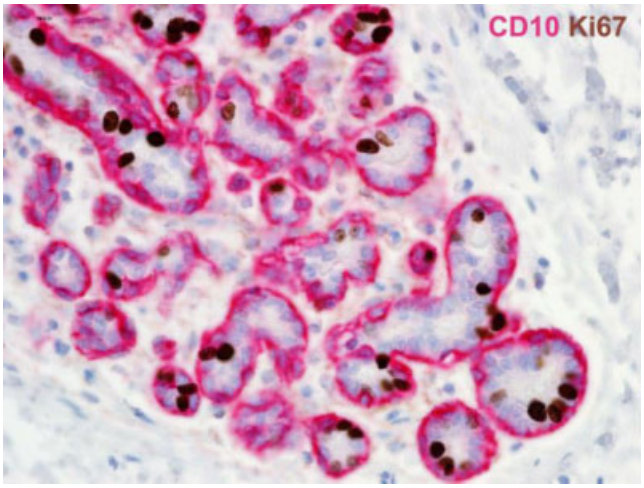
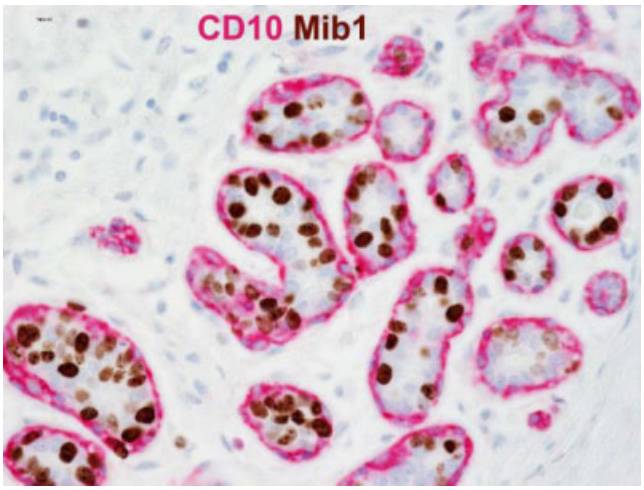
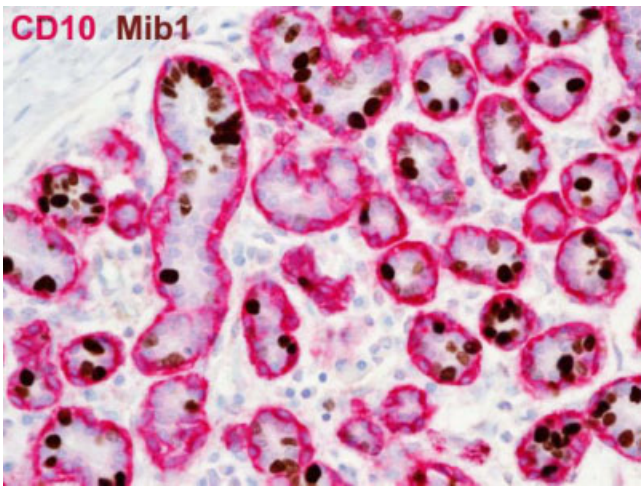
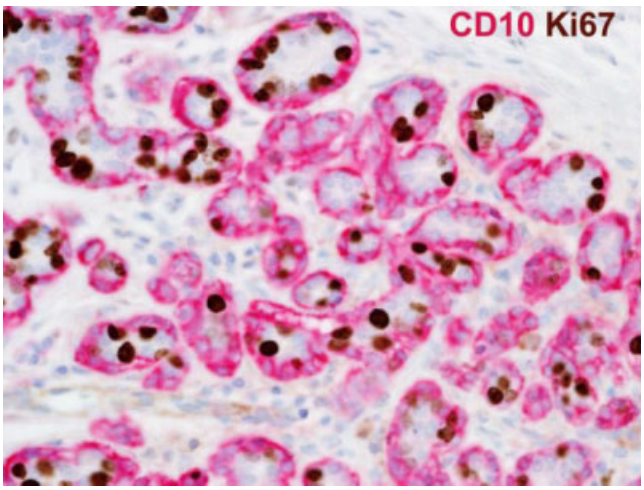


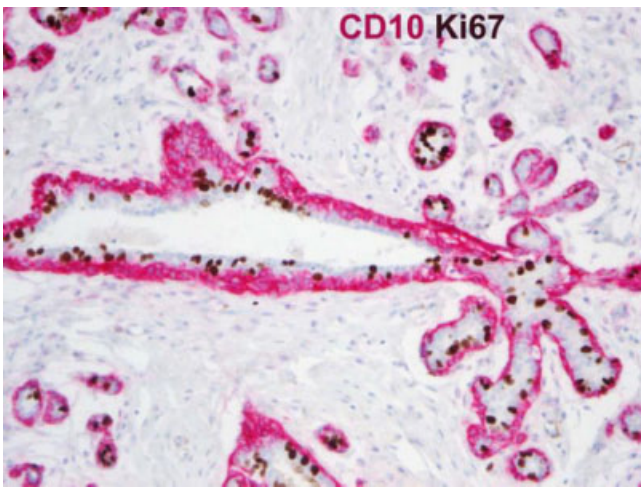
Table S2E CD10+Ki67











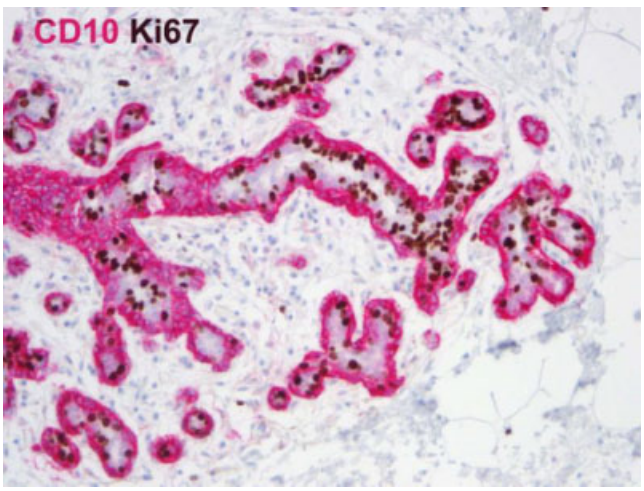


Table S2F K14+K18

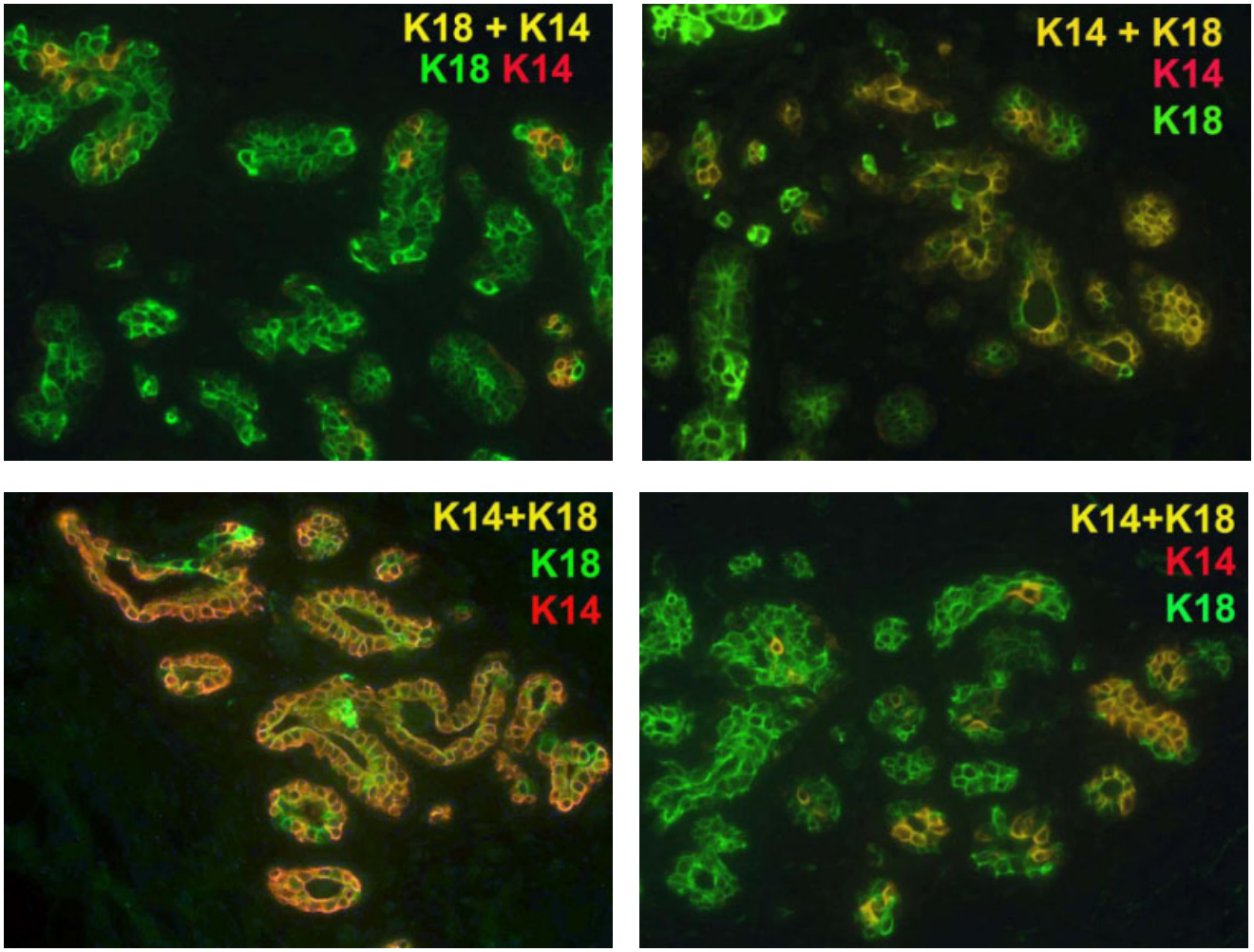


Table S2F K5+K18