

**Supplementary Table 1. Summary of Clinical and Pathological Data for the TCGA and MDACC Cohorts.**

Stage	Subtype	Gender F/M	Total	Age, yr, mean $\pm$ SD
<b>TCGA cohort</b>				
Invasive (T2 and higher)	Luminal	48/164	212	68.3 $\pm$ 11.0
	Basal	56/123	179	68.0 $\pm$ 10.1
	Double negative	3/14	17	65.6 $\pm$ 10.2
<b>MD Anderson fresh frozen tissue cohort</b>				
Noninvasive (Ta-Tis)	Luminal	8/24	32	64.8 $\pm$ 12.1
Invasive (T1 and higher)	Luminal	11/49	60	67.4 $\pm$ 10.6
	Basal	12/23	35	69.6 $\pm$ 10.6
	Double negative	1/4	5	63.8 $\pm$ 7.5
<b>MD Anderson formalin-fixed paraffin-embedded tissue cohort</b>				
Invasive (T2 and higher)	Luminal	8/38	46	70.2 $\pm$ 11.6
	Basal	11/18	29	69.2 $\pm$ 11.0
	Double Neg	3/11	14	68.0 $\pm$ 8.4

F, female; M, male; yr, year; SD, standard deviation