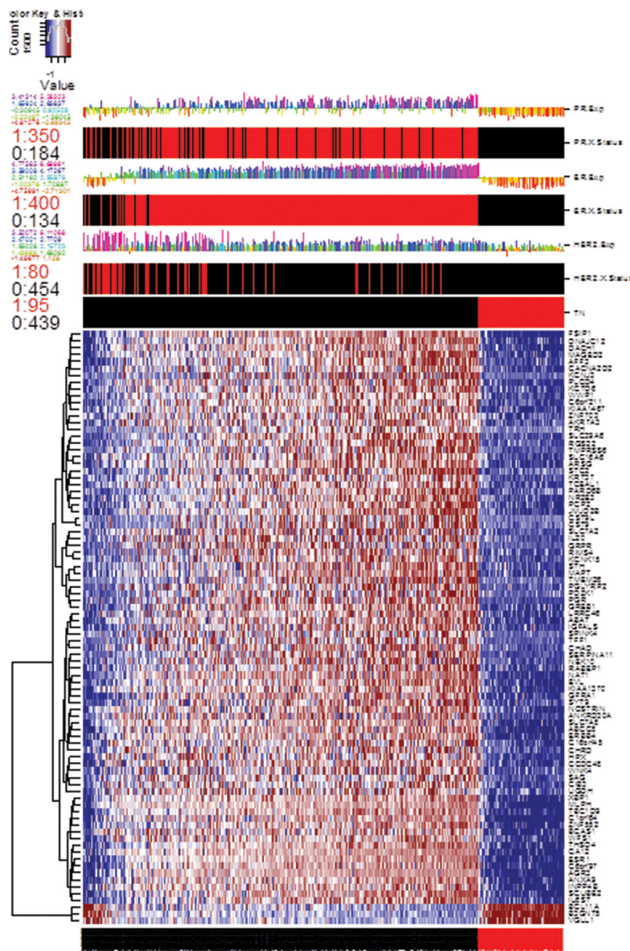


Supplemental Data

A New Gene Expression Signature for Triple-Negative Breast Cancer Using Frozen Fresh Tissue before Neoadjuvant Chemotherapy

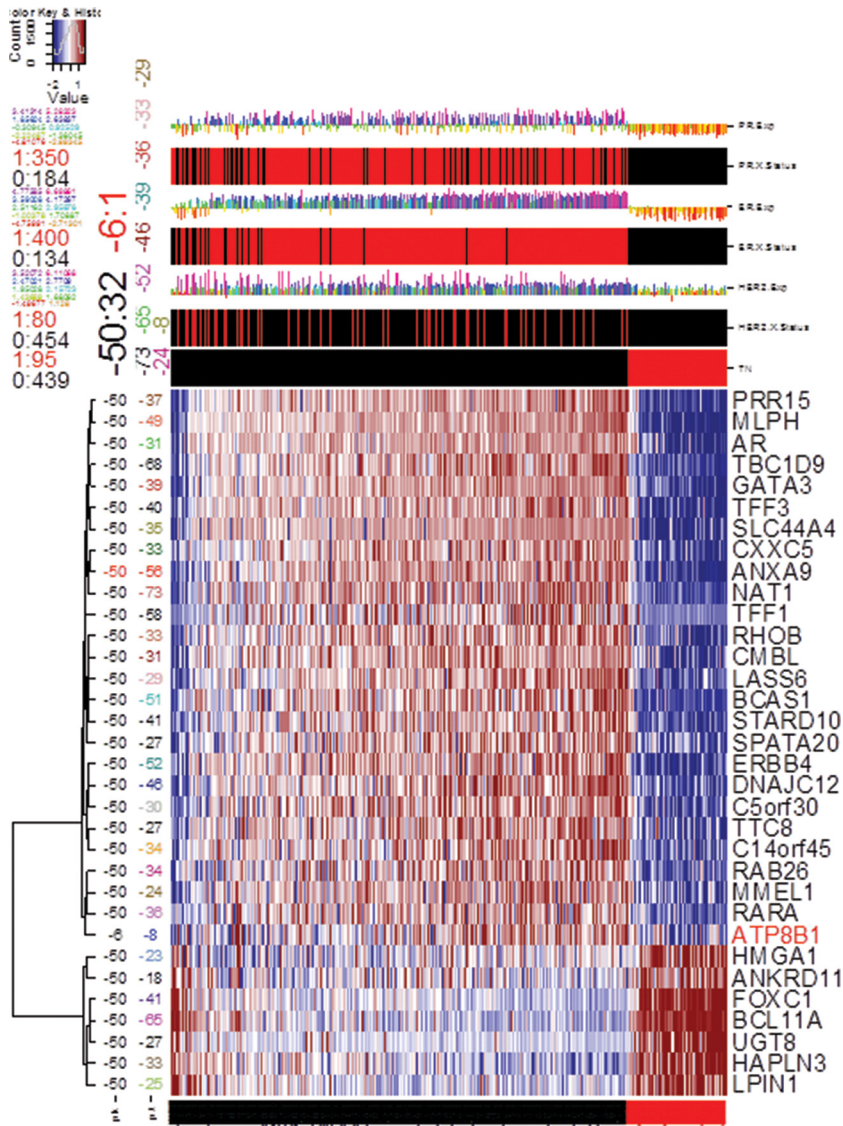
Sandra K Santuario-Facio,<sup>1,7\*</sup> Servando Cardona-Huerta,<sup>2\*</sup> Yadira X Perez-Paramo,<sup>3</sup> Victor Trevino,<sup>4</sup> Francisco Hernandez-Cabrera,<sup>5</sup> Augusto Rojas-Martinez,<sup>1,7</sup> Grecia Uscanga-Perales,<sup>6</sup> Jorge L Martinez-Rodriguez,<sup>2,8</sup> Lizeth Martinez-Jacobo,<sup>6</sup> Gerardo Padilla-Rivas,<sup>6</sup> Gerardo Muñoz-Maldonado,<sup>9</sup> Juan Francisco Gonzalez-Guerrero,<sup>10</sup> Javier Valero-Gomez,<sup>2</sup> Ana L Vazquez-Guerrero,<sup>2</sup> Herminia G Martinez-Rodriguez,<sup>6</sup> Alvaro Barboza-Quintana,<sup>2</sup> Oralia Barboza-Quintana,<sup>11</sup> Raquel Garza-Guajardo,<sup>11</sup> and Rocio Ortiz-Lopez<sup>1,7</sup>

Online address: <http://www.molmed.org>



Supplementary Figure S1. TCGA Validation, first approach. Gene signature of 87 differentially expressed genes with high significance ( $q < 1e-40$ ), only nine genes overlap with our signature.

GENE SIGNATURE OF TRIPLE NEGATIVE BREAST CANCER



**Supplementary Figure S2.** TCGA Validation, second approach. Differential expression of TNBC versus nTNBC using only 33 of 40 genes of our signature.

Supplemental Table S1. TNBC signature (40 genes and 49 probes).

Expression TNBC	Gene Symbol	ID Probe	Gene Title	t test p value	Kolmogorov test p value	FDR test p value	FDR k p value
Over	ANKRD11	238538_at	Ankyrin repeat domain 11	2.01e-07	9.19e-07	0.0001076	0.00048764
Over	BCL11A	1559078_at	B-cell cll/lymphoma 11a (zinc finger protein)	4.12e-08	3.49e-08	4.00e-05	0.00014361
Over	BCL11A	210347_s_at	B-cell cll/lymphoma 11a (zinc finger protein)	2.66e-10	1.15e-07	4.52e-06	0.00020489
Over	BCL11A	222891_s_at	B-cell cll/lymphoma 11a (zinc finger protein)	7.57e-09	4.04e-07	1.71e-05	0.00029192
Over	FAM171A1	212771_at	Family with sequence similarity 171, member a1	7.42e-08	6.21e-07	6.14e-05	0.00036378
Over	FOXC1	1553613_s_at	Forkhead box c1	2.71e-09	2.52e-07	1.32e-05	0.00023999
Over	HAPLN3	227262_at	Hyaluronan and proteoglycan link protein 3	2.29e-07	4.23e-08	0.0001132	0.00014361
Over	HMGAT1	210457_x_at	High mobility group at-hook 1	8.39e-07	6.79e-08	0.000259	0.00016469
Over	LPIN1	212276_at	Lipin 1	9.37e-08	1.15e-07	6.69e-05	0.00020489
Over	PRKX /// PRKY	204060_s_at	Protein kinase, x-linked /// protein kinase, y-linked	4.79e-09	1.04e-08	1.58e-05	8.84E-05
Over	UGT8	208358_s_at	Udp glycosyltransferase 8	4.54e-07	1.87e-08	0.0001837	0.00010584
Over	UGT8	228956_at	Udp glycosyltransferase 8	3.17e-07	1.72e-07	0.0001418	0.00023999
Under	ANXA9	211712_s_at	Annexin a9	2.45e-08	3.79e-09	2.98e-05	6.43E-05
Under	ANXA9	210085_s_at	Annexin a9	1.35e-06	2.52e-07	0.0003491	0.00023999
Under	AR	211110_s_at	Androgen receptor	1.72e-08	2.93e-07	2.65e-05	0.00024238
Under	ATP8B1	226302_at	ATPase, aminophospholipid transporter, class I, type 8b, member 1	2.49e-07	9.19e-07	0.000121	0.00048764
Under	BCAS1	204378_at	Breast carcinoma amplified sequence 1	2.15e-06	1.15e-07	0.0004628	0.00020489
Under	C5orf30	221823_at	Chromosome 5 open reading frame 30	1.46e-06	6.21e-07	0.0003575	0.00036378
Under	CCDC96	238682_at	Coiled-coil domain containing 96	5.08e-06	2.61e-07	0.0007376	0.00023999
Under	CMBL	227522_at	Carboxymethylenebutenolidase homolog (pseudomonas)	1.40e-08	6.21e-07	2.32e-05	0.00036378
Under	CXC5	224516_s_at	Cxcr5 finger protein 5	6.75e-07	2.22e-07	0.0002293	0.00023999
Under	DNAJC12	223721_s_at	Dnaj (hsp40) homolog, subfamily c, member 12	2.11e-07	4.23e-08	0.0001102	0.00014361
Under	DNAJC12	218976_at	Dnaj (hsp40) homolog, subfamily c, member 12	2.19e-08	6.79e-08	2.87e-05	0.00016469
Under	ERBB4	214053_at	V-erb-a erythroblastic leukemia viral oncogene homolog 4 (avian)	5.30e-07	6.21e-07	0.0001998	0.00036378
Under	FAM174B	51158_at	Family with sequence similarity 174, member b	7.50e-09	2.22e-07	1.71e-05	0.00023999
Under	FAM174B	221880_s_at	Family with sequence similarity 174, member b	7.34e-08	2.93e-07	6.14e-05	0.00024238
Under	FLJ38379	1556474_a_at	Hypothetical flj38379	8.76e-09	2.73e-08	1.75e-05	0.00013249
Under	GATA3	209603_at	Gata binding protein 3	5.34e-09	9.19e-07	1.58e-05	0.00048764
Under	GATA3	209602_s_at	Gata binding protein 3	8.54e-10	9.19e-07	7.25e-06	0.00048764
Under	LASS6	212446_s_at	Lag1 homolog, ceramide synthase 6	5.95e-08	1.72e-07	5.32e-05	0.00023999
Under	LASS6	242019_at	Lag1 homolog, ceramide synthase 6	1.88e-05	6.21e-07	0.001669	0.00036378
Under	LOC100505614	237339_at	Hypothetical loc100505614	3.79e-08	6.21e-07	3.84e-05	0.00036378
Under	LOC692247	1569453_a_at	Hypothetical locus loc692247	9.45e-08	6.06e-10	6.69e-05	2.06E-05
Under	MILPH	218211_s_at	Melanophilin	5.19e-10	1.15e-07	5.88e-06	0.00020489
Under	MIMEL1	1552930_at	Membrane metallo-endoropeptidase-like 1	1.21e-06	2.19e-07	0.0003269	0.00023999
Under	NAT1	214440_at	N-acetyltransferase 1 (arylamine n-acetyltransferase)	3.04e-07	1.04e-08	0.0001397	8.84E-05
Under	TTC6	1556666_a_at	Non-protein coding rna 291	1.57e-06	6.79e-08	0.0003728	0.00016469
Under	PRR15	226961_at	Proline rich 15	2.15e-09	4.04e-07	1.24e-05	0.00029192
Under	RAB26	50965_at	Rab26, member ras oncogene family	1.19e-07	2.22e-07	7.80e-05	0.00023999
Under	RARA	203749_s_at	Retinoic acid receptor, alpha	7.68e-08	4.04e-07	6.21e-05	0.00029192
Under	RHOB	212099_at	Ras homolog gene family, member b	3.99e-07	2.93e-07	0.0001693	0.00024238
Under	SLC44A4	205597_at	Solute carrier family 44, member 4	1.65e-07	1.87e-08	9.20e-05	0.00010584
Under	SPATA20	218164_at	Spermatogenesis associated 20	1.82e-08	2.49e-07	2.68e-05	0.00023999
Under	STARD10	223103_at	Star-related lipid transfer (start) domain containing 10	2.91e-08	1.72e-07	3.40e-05	0.00023999
Under	TBC1D9	212960_at	Tbc1 domain family, member 9 (with gram domain)	1.44e-07	2.22e-07	8.44e-05	0.00023999
Under	TBC1D9	212956_at	Tbc1 domain family, member 9 (with gram domain)	4.32e-09	6.21e-07	1.58e-05	0.00036378
Under	TFF1	205009_at	Trefoil factor 1	4.91e-07	2.22e-07	0.0001877	0.00023999
Under	TFF3	204623_at	Trefoil factor 3 (intestinal)	6.32e-09	2.93e-07	1.65e-05	0.00024238
Under	TTC8	226120_at	Tetraficopeptide repeat domain 8	1.29e-06	6.21e-07	0.0003421	0.00036378