Supplementary Data

TITLE Activation of epithelial-mesenchymal transition process during breast cancer progression – the impact of molecular subtype and stromal composition

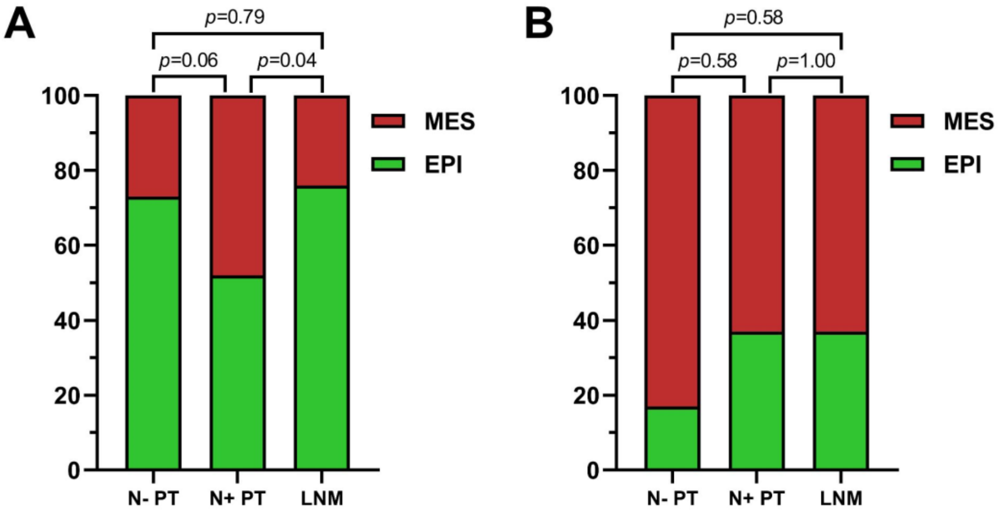


Figure S1. Percentages of epithelial (EPI) and mesenchymal (MES) phenotypes of PT which have not (N-) or have metastasized to the lymph nodes (N+) as well as lymph node metastases themselves (LNM) divided into luminal (panel A) or TNBC molecular subtype (B). Significance levels were values calculated with Pearson's chi-squared or Fisher's exact test.

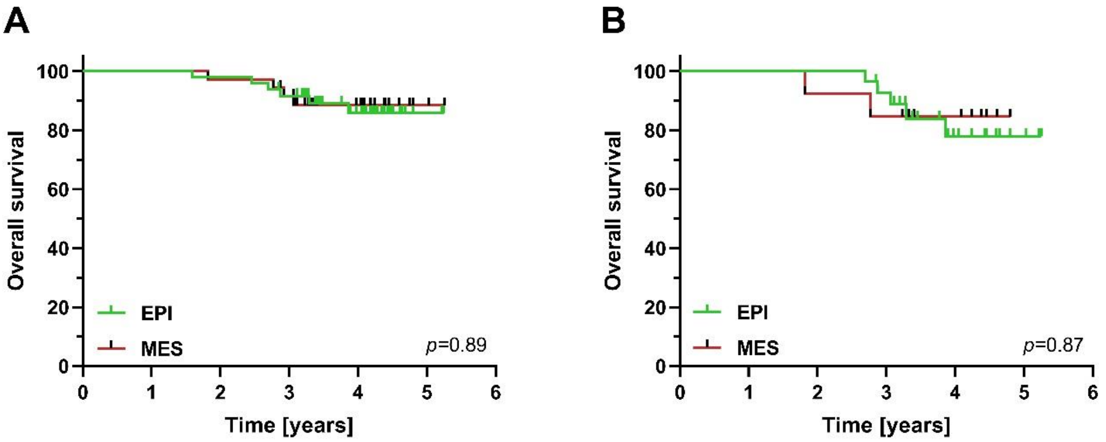


Figure S2. Prognostic significance of EMT phenotypes in primary tumours (A) and lymph node metastases (B) - effect on overall survival of the patients (log-rank test).

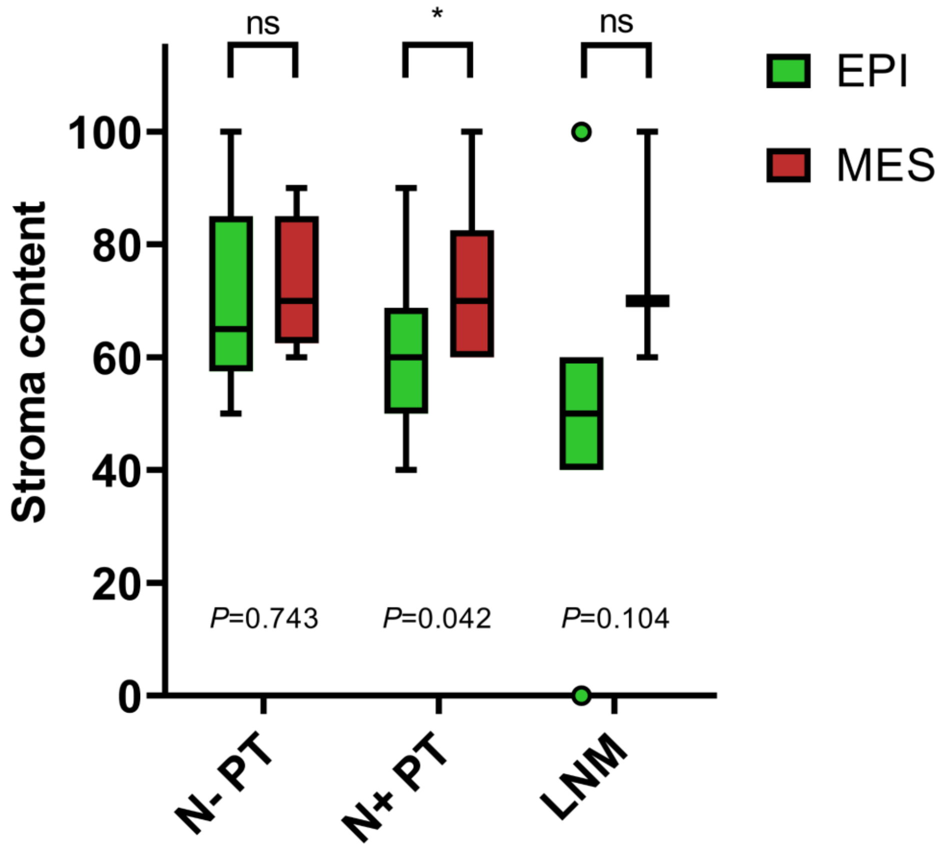
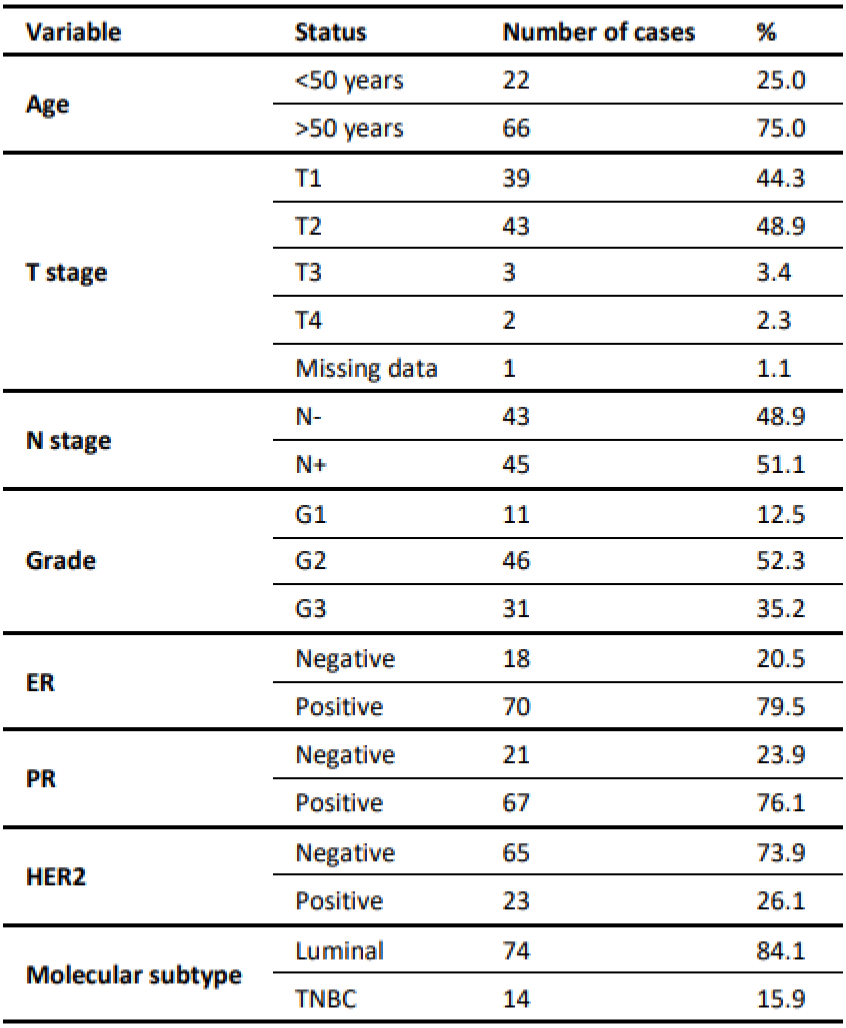


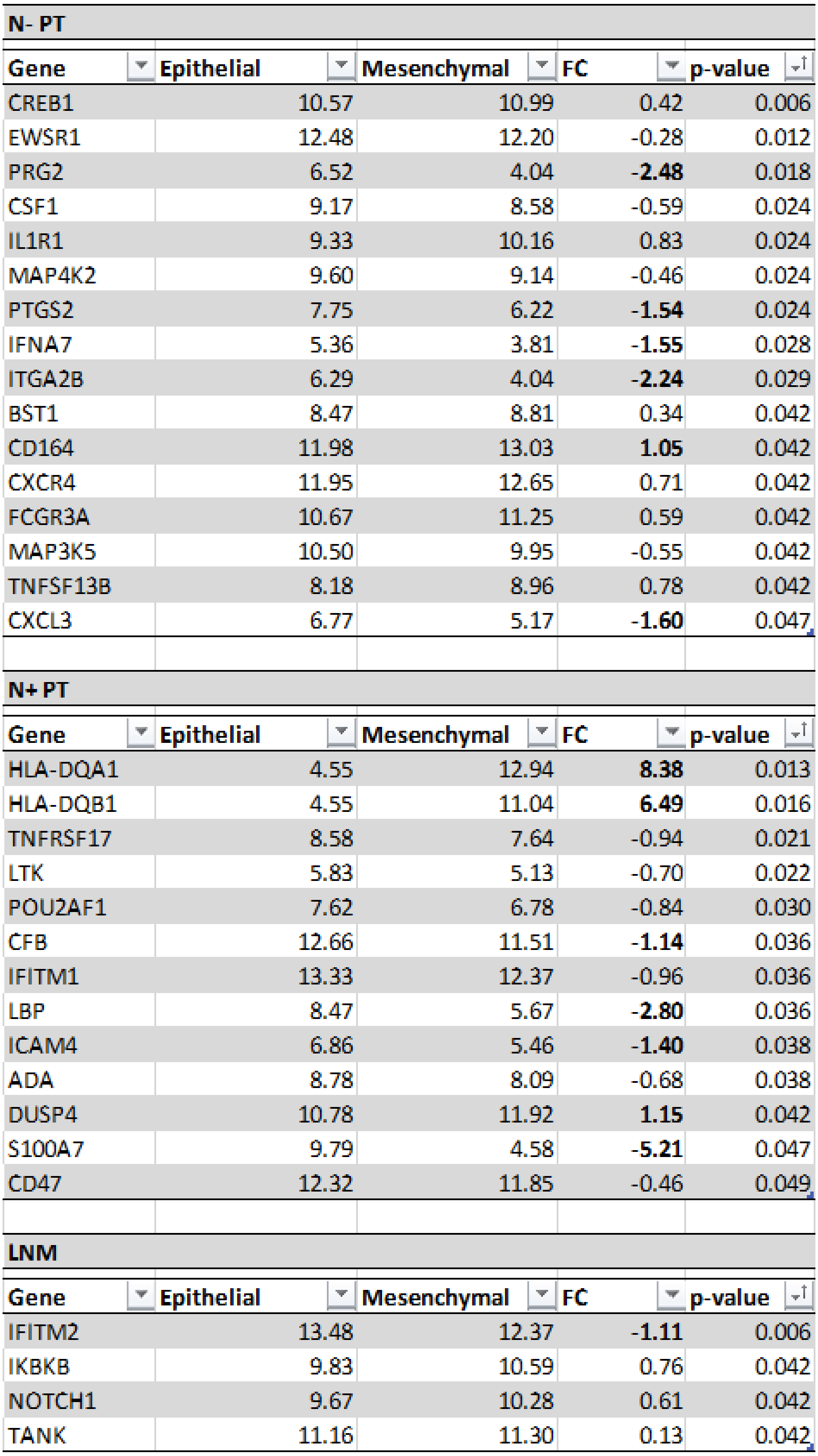
Figure S3. Distribution of stroma content in N- PT, N+ PT and LNM specimens.
The maximum record out of 5 evaluated tissue cores is plotted for each specimen. Differences in median stroma content between EMT groups were analysed with the Mann–Whitney U test; the bars correspond to the interquartile range (IQR), the whiskers cover 1.5 IQR from the median.

Table S1. Clinico-pathological characteristics of the breast cancer patients included in the study.



Variable	Status	Number of cases	%
Age	<50 years	22	25.0
	>50 years	66	75.0
T stage	T1	39	44.3
	T2	43	48.9
	T3	3	3.4
	T4	2	2.3
	Missing data	1	1.1
N stage	N-	43	48.9
	N+	45	51.1
Grade	G1	11	12.5
	G2	46	52.3
	G3	31	35.2
ER	Negative	18	20.5
	Positive	70	79.5
PR	Negative	21	23.9
	Positive	67	76.1
HER2	Negative	65	73.9
	Positive	23	26.1
Molecular subtype	Luminal	74	84.1
	TNBC	14	15.9

Table S2. List of all DEGs in each group of tissues – N- PT, N+ PT and LNM, analysed according to their EMT phenotype.



N- PT

Gene	Epithelial	Mesenchymal	FC	p-value
CREB1	10.57	10.99	0.42	0.006
EWSR1	12.48	12.20	-0.28	0.012
PRG2	6.52	4.04	**-2.48**	0.018
CSF1	9.17	8.58	-0.59	0.024
IL1R1	9.33	10.16	0.83	0.024
MAP4K2	9.60	9.14	-0.46	0.024
PTGS2	7.75	6.22	**-1.54**	0.024
IFNA7	5.36	3.81	**-1.55**	0.028
ITGA2B	6.29	4.04	**-2.24**	0.029
BST1	8.47	8.81	0.34	0.042
CD164	11.98	13.03	**1.05**	0.042
CXCR4	11.95	12.65	0.71	0.042
FCGR3A	10.67	11.25	0.59	0.042
MAP3K5	10.50	9.95	-0.55	0.042
TNFSF13B	8.18	8.96	0.78	0.042
CXCL3	6.77	5.17	**-1.60**	0.047

N+ PT

Gene	Epithelial	Mesenchymal	FC	p-value
HLA-DQA1	4.55	12.94	**8.38**	0.013
HLA-DQB1	4.55	11.04	**6.49**	0.016
TNFRSF17	8.58	7.64	-0.94	0.021
LTK	5.83	5.13	-0.70	0.022
POU2AF1	7.62	6.78	-0.84	0.030
CFB	12.66	11.51	**-1.14**	0.036
IFITM1	13.33	12.37	-0.96	0.036
LBP	8.47	5.67	**-2.80**	0.036
ICAM4	6.86	5.46	**-1.40**	0.038
ADA	8.78	8.09	-0.68	0.038
DUSP4	10.78	11.92	**1.15**	0.042
S100A7	9.79	4.58	**-5.21**	0.047
CD47	12.32	11.85	-0.46	0.049

LNM

Gene	Epithelial	Mesenchymal	FC	p-value
IFITM2	13.48	12.37	**-1.11**	0.006
IKBKB	9.83	10.59	0.76	0.042
NOTCH1	9.67	10.28	0.61	0.042
TANK	11.16	11.30	0.13	0.042

Table S3. List of GO BP and Reactome terms enriched in each list of DEGs - in N- PT, N+ PT and LNM tissues.

N- PT

Category	Term	Count	%	PValue	Genes	List Tot	Pop Hi	Pop To	Fold Er	Bonfer	Benjan	FDR
GOTERM_BP_DIRECT	GO:0006955~immune response	7	43.75	**9.90E-07**	CD164, MA	16	421	16792	17.45	**2.31E-04**	**2.31E-04**	**2.31E-04**
GOTERM_BP_DIRECT	GO:0006954~inflammatory response	4	25.00	**4.24E-03**	CSF1, CXCR	16	379	16792	11.08	6.29E-01	4.69E-01	4.69E-01
GOTERM_BP_DIRECT	GO:0045087~innate immune response	4	25.00	**6.03E-03**	MAP4K2, IFI	16	430	16792	9.76	7.56E-01	4.69E-01	4.69E-01
GOTERM_BP_DIRECT	GO:0045672~positive regulation of osteoclast differentiation	2	12.50	**1.68E-02**	CREB1, CSF1	16	19	16792	110.47	9.81E-01	8.97E-01	8.97E-01
GOTERM_BP_DIRECT	GO:0043065~positive regulation of apoptotic process	3	18.75	**2.87E-02**	CREB1, PTG	16	300	16792	10.50	9.99E-01	8.97E-01	8.97E-01
GOTERM_BP_DIRECT	GO:0040018~positive regulation of multicellular organism growth	2	12.50	**3.08E-02**	CREB1, CSF1	16	35	16792	59.97	9.99E-01	8.97E-01	8.97E-01
GOTERM_BP_DIRECT	GO:0007257~activation of JUN kinase activity	2	12.50	**3.34E-02**	MAP4K2, M	16	38	16792	55.24	1.00E+00	8.97E-01	8.97E-01
GOTERM_BP_DIRECT	GO:0030890~positive regulation of B cell proliferation	2	12.50	**3.43E-02**	BST1, TNFS	16	39	16792	53.82	1.00E+00	8.97E-01	8.97E-01
GOTERM_BP_DIRECT	GO:0007254~JNK cascade	2	12.50	**4.29E-02**	MAP4K2, M	16	49	16792	42.84	1.00E+00	8.97E-01	8.97E-01
GOTERM_BP_DIRECT	GO:0006959~humoral immune response	2	12.50	**4.97E-02**	BST1, IFNA7	16	57	16792	36.82	1.00E+00	8.97E-01	8.97E-01
GOTERM_BP_DIRECT	GO:0030097~hemopoiesis	2	12.50	5.14E-02	CD164, CSF1	16	59	16792	35.58	1.00E+00	8.97E-01	8.97E-01
GOTERM_BP_DIRECT	GO:0007613~memory	2	12.50	5.40E-02	CREB1, PTG	16	62	16792	33.85	1.00E+00	8.97E-01	8.97E-01
GOTERM_BP_DIRECT	GO:0050727~regulation of inflammatory response	2	12.50	5.48E-02	IL1R1, PTGS	16	63	16792	33.32	1.00E+00	8.97E-01	8.97E-01
GOTERM_BP_DIRECT	GO:0046330~positive regulation of JNK cascade	2	12.50	5.65E-02	MAP4K2, M	16	65	16792	32.29	1.00E+00	8.97E-01	8.97E-01
REACTOME_PATHWAY	R-HSA-380108~Chemokine receptors bind chemokines	2	12.50	5.70E-02	CXCR4, CXC	10	59	9075	30.76	9.15E-01	1.00E+00	1.00E+00
GOTERM_BP_DIRECT	GO:0006468~protein phosphorylation	3	18.75	6.12E-02	MAP4K2, CR	16	456	16792	6.90	1.00E+00	8.97E-01	8.97E-01
GOTERM_BP_DIRECT	GO:0070098~chemokine-mediated signaling pathway	2	12.50	6.16E-02	CXCR4, CXC	16	71	16792	29.56	1.00E+00	8.97E-01	8.97E-01
GOTERM_BP_DIRECT	GO:0009615~response to virus	2	12.50	9.39E-02	IFNA7, CXC	16	110	16792	19.08	1.00E+00	1.00E+00	1.00E+00

N+ PT

Category	Term	Count	%	PValue	Genes	List Tot	Pop Hi	Pop To	Fold Er	Bonfer	Benjan	FDR
GOTERM_BP_DIRECT	GO:0008228~opsonization	2	15.38	**7.13E-03**	CD47, LBP	13	10	16792	258.34	6.33E-01	8.46E-01	8.46E-01
GOTERM_BP_DIRECT	GO:0002504~antigen processing and presentation of peptide or poly	2	15.38	**1.21E-02**	HLA-DQA1,	13	17	16792	151.96	8.18E-01	8.46E-01	8.46E-01
REACTOME_PATHWAY	R-HSA-202430~Translocation of ZAP-70 to Immunological synapse	2	15.38	**2.16E-02**	HLA-DQA1,	10	22	9075	82.50	3.82E-01	1.87E-01	1.87E-01
REACTOME_PATHWAY	R-HSA-202427~Phosphorylation of CD3 and TCR zeta chains	2	15.38	**2.45E-02**	HLA-DQA1,	10	25	9075	72.60	4.21E-01	1.87E-01	1.87E-01
REACTOME_PATHWAY	R-HSA-389948~PD-1 signaling	2	15.38	**2.55E-02**	HLA-DQA1,	10	26	9075	69.81	4.34E-01	1.87E-01	1.87E-01
REACTOME_PATHWAY	R-HSA-202433~Generation of second messenger molecules	2	15.38	**3.52E-02**	HLA-DQA1,	10	36	9075	50.42	5.45E-01	1.93E-01	1.93E-01
GOTERM_BP_DIRECT	GO:0019882~antigen processing and presentation	2	15.38	**3.86E-02**	HLA-DQA1,	13	55	16792	46.97	9.96E-01	1.00E+00	1.00E+00
GOTERM_BP_DIRECT	GO:0050829~defense response to Gram-negative bacterium	2	15.38	**3.86E-02**	LBP, S100A7	13	55	16792	46.97	9.96E-01	1.00E+00	1.00E+00
GOTERM_BP_DIRECT	GO:0060333~interferon-gamma-mediated signaling pathway	2	15.38	**4.96E-02**	HLA-DQA1,	13	71	16792	36.39	9.99E-01	1.00E+00	1.00E+00
GOTERM_BP_DIRECT	GO:0031295~T cell costimulation	2	15.38	5.44E-02	HLA-DQA1,	13	78	16792	33.12	1.00E+00	1.00E+00	1.00E+00
GOTERM_BP_DIRECT	GO:0019886~antigen processing and presentation of exogenous pep	2	15.38	6.38E-02	HLA-DQA1,	13	92	16792	28.08	1.00E+00	1.00E+00	1.00E+00
REACTOME_PATHWAY	R-HSA-216083~Integrin cell surface interactions	2	15.38	8.22E-02	ICAM4, CD4	10	86	9075	21.10	8.48E-01	3.01E-01	3.01E-01
REACTOME_PATHWAY	R-HSA-877300~Interferon gamma signaling	2	15.38	8.40E-02	HLA-DQA1,	10	88	9075	20.63	8.55E-01	3.01E-01	3.01E-01
REACTOME_PATHWAY	R-HSA-202424~Downstream TCR signaling	2	15.38	9.59E-02	HLA-DQA1,	10	101	9075	17.97	8.91E-01	3.01E-01	3.01E-01

LNM

Category	Term	Count	%	PValue	Genes	List Tot	Pop Hi	Pop To	Fold Er	Bonfer	Benjan	FDR
GOTERM_BP_DIRECT	GO:0035666~TRIF-dependent toll-like receptor signaling pathway	2	50.00	**4.99E-03**	IKBKB, TAN	4	28	16792	299.86	5.84E-01	8.54E-01	8.54E-01
GOTERM_BP_DIRECT	GO:0007249~I-kappaB kinase/NF-kappaB signaling	2	50.00	**1.07E-02**	IKBKB, TAN	4	60	16792	139.93	8.47E-01	8.54E-01	8.54E-01
GOTERM_BP_DIRECT	GO:0009615~response to virus	2	50.00	**1.95E-02**	IKBKB, IFIT	4	110	16792	76.33	9.68E-01	8.54E-01	8.54E-01
GOTERM_BP_DIRECT	GO:0071356~cellular response to tumor necrosis factor	2	50.00	**1.95E-02**	IKBKB, TAN	4	110	16792	76.33	9.68E-01	8.54E-01	8.54E-01
GOTERM_BP_DIRECT	GO:0006955~immune response	2	50.00	7.33E-02	NOTCH1, IFI	4	421	16792	19.94	1.00E+00	1.00E+00	1.00E+00
GOTERM_BP_DIRECT	GO:0045893~positive regulation of transcription, DNA-templated	2	50.00	8.92E-02	IKBKB, NOT	4	515	16792	16.30	1.00E+00	1.00E+00	1.00E+00