Supplementary Data

JUNB governs a feed-forward network of TGFβ signaling that aggravates breast cancer invasion

Anders Sundqvist[1,2,6,*], Masato Morikawa[1,3,6], Jiang Ren[5], Eleftheria Vasilaki[1,2], Natsumi Kawasaki[3], Mai Kobayashi[3], Daizo Koinuma[1,3], Hiroyuki Aburatani[4], Kohei Miyazono[1,2,3], Carl-Henrik Heldin[1,2], Hans van Dam[1,5] & Peter ten Dijke[1,2,5,*]

[1]Ludwig Cancer Research, [2]Department of Medical Biochemistry and Microbiology, Science for Life Laboratory, Box 582, Biomedical Center, Uppsala University, SE-751 23 Uppsala, Sweden, [3]Department of Molecular Pathology, Graduate School of Medicine, The University of Tokyo, Tokyo 113-0033, Japan, [4]Genome Science Division, Research Center for Advanced Science and Technology, The University of Tokyo, Tokyo 153-8904, Japan, [5]Department of Molecular Cell Biology, Cancer Genomics Centre Netherlands, Leiden University Medical Center, P.O. Box 9600, 2300 RC, Leiden, The Netherlands.

*To whom correspondence should be addressed. Tel: +46 18 160408; Fax: +46 18 160420; Email: anders.sundqvist@imbim.uu.se. Correspondence may also be addressed to: Tel: +31 71 5269271; Fax: +31 71 5268270; Email: P.ten_Dijke@lumc.nl.

[6]The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors.

Supplementary Figure Legends
Supplementary Figures S1-S7
Supplementary Tables S1-S2

Supplementary Figure Legends

Supplementary Figure S1. SMAD2/3 are redirected to different sites in MCF10A MII after prolonged TGFβ treatment. (**A**) ChIP-qPCR showing time-dependent recruitment of SMAD2/3 to the SMAD binding site close to the *SERPINE1* transcription start site (TSS), the binding site approximately 5 kb upstream of the *LAMB3* TSS, and to the intronic binding site in the *MMP2* gene locus, in MCF10A MII cells after no treatment, or treatment with TGFβ (5 ng/ml) for 1.5 h or 16 h. Results of three independent experiments are shown by dot plot chart; *p < 0.05, **p < 0.01. (**B**) Quantification of overlap between the SMAD2/3 binding sites and histone marks in breast epithelial cells, related to Fig 1E. (**C**) Functional annotation of SMAD2/3 binding regions, performed using GREAT (35). The top five over-represented categories belonging to Gene Ontology (GO) biological process, which describes the biological processes associated with gene function, are presented. The x axis represents binomial raw (uncorrected) P-values in (-log10).

Supplementary Figure S2. JUNB is a critical AP1 component for SMAD2/3 binding after TGFβ stimulation. (**A**) Expression levels of indicated genes in MCF10A MII cells after 1.5 h and 16 h TGFβ (5 ng/ml) treatment are shown in FPKM (fragments per kilobase of exon per million fragments mapped) values (data represent FPKM ± 95% confidence interval). (**B** and **C**) Frequency of *JUNB* gene alterations (mutation, amplification and deletion) in breast cancer datasets using cBioPortal (40,42,43). Patients with *JUNB* amplification had a trend of poorer prognosis (**C**), although this was not statistically significant because of the small number of the cases. (**D**) Western blot control of JUNB knockdown efficiency of Figure 3C.

Supplementary Figure S3. A JUNB-mediated feed-forward mechanism regulates genes associated with cell adhesion and invasion. (**A**) A list of genes whose induction after 16 h TGFβ (5 ng/ml) treatment was attenuated more than 50 % with siJUNB treatment. See also Fig 4A. (**B**) qRT-PCR validation of identified JUNB target genes. MCF10A MII cells were transfected with non-targeting control (siNTC) or specific JUNB siRNA and stimulated for 1.5 h or 16 h with TGFβ (5 ng/ml). Results of three independent experiments are shown; *p < 0.05, **p < 0.01. (**C**) GSEA of expression changes of SMAD2/3 target genes after manipulation of JUNB expression. The SMAD2/3 target genes were pre-rank-ordered according to their fold change (log2) between siNTC and siJUNB, and analyzed based on KEGG signaling pathway enrichment. Gene sets with p-value < 5% and FDR q-value < 25% were considered significant. See also Fig 4B.

Supplementary Figure S4. JUNB regulates genes associated with EMT and and invasion. (**A**) qRT-PCR analysis of late TGFβ-induced gene expression. MCF10A MII cells were stimulated for 16, 48 or 72 h with TGFβ (5 ng/ml). A representative results of three independent experiments is shown. (**B**) Western blot (left) and qRT-PCR (right) analysis of A549 human pulmonary adenocarcinoma cells transfected with non-targeting control (siNTC) or specific JUNB siRNA and treated with TGFβ (5 ng/ml) for 48 h (left) or 16 h (right) as indicated. (**C**) A549 mCherry cells transfected with non-targeting control (siNTC) or specific JUNB siRNA were injected into the ducts of

Cuvier (DoC) of 48 hours post-fertilization (hpf) zebrafish embryos. Left: representative images of zebrafish at 6 days post-injection (dpi). Right: quantification of invasive cell cluster numbers in non-targeting and JUNB knock-down cells injected zebrafish larvae.

Supplementary Figure S5. Activation of the WNT signaling pathway strengthens the TGFβ-induced migratory phenotype. (**A**) Genomic locus of the *WNT7B* gene shown together with the results of SMAD2/3 ChIP-seq data obtained in MCF10A MII cells. The direction of transcription is shown by the arrow beginning at the TSS. Statistically significant regions are marked by a gray-colored box. (B) Western blot for WNT7A and WNT7B in MCF10A MII cells stable expressing control GFP or MYC-tagged WNT7A (WNT7A-MYC) or WNT7B (WNT7B-MYC). (**C**) Collagen invasion assay of MCF10A MII spheroids. Spheroids were embedded in collagen in the absence or presence of TGFβ (5 ng/ml), recombinant WNT7A (300 ng/ml) or the WNT inhibitor (WNTi) IWP-2 (5 μM), as indicated. Left: representative pictures of spheroids taken 36 h after being embedded in collagen. Right: relative invasion was quantified as the mean area that the spheroids occupied 36 h after being embedded in collagen. Data represent means ± SD ($n \geq 6$ spheroids per condition) and are representative of two independent experiments; *p < 0.05, **p < 0.01, ***p < 0.001. (**D**) Collagen invasion assay of MCF10A MII spheroids stably expressing control GFP, WNT7A or WNT7B. Spheroids were embedded in collagen in the absence or the presence of TGFβ (5 ng/ml). Left: representative pictures of spheroids taken 36 h after being embedded in collagen. Right: relative invasion was quantified as the mean area that the spheroids occupied 36 h after being embedded in collagen. Data represent means ± SD ($n \geq 6$ spheroids per condition) and are representative of two independent experiments; ***p < 0.001. (**E**) Canonical WNT signaling activity as measured by a TCF/LEF driven transcriptional luciferase reporter plasmid system in the MCF10A MII cells stably expressing control GFP or ectopic WNT7B-MYC.

Supplementary Figure S6. ChIP-seq of TNBC cell lines and meta-analysis of published microarray datasets of Breast Cancer patients.(**A**) Kaplan-Meyer analysis of relapse-free survival of breast cancer datasets, generated using KM plotter [35]; all subtypes, n=3,557; ER^+ subjects, n=2,036; ER^- subjects, n=807; luminal A subtype, n=2,069; luminal B subtype, n=1,166; HER2- subtype, n=239; basal-like subtype, n=668). Survival analysis was performed using a log-rank test. (**B**) Genomic loci of *SMAD7* and *WNT7B* are shown together with the results of SMAD2/3 ChIP-seq data obtained in the TNBC cells Hs-578-T and BT-549. The direction of transcription is shown by the arrow beginning at the TSS. Statistically significant regions are marked by a gray-colored box. (**C**) The number of SMAD2/3 binding sites and overlap between 1.5 h and 16 h. The number of ChIP-seq peaks in each time point is presented. The number of peaks overlapping with other conditions is also presented, together with the percent to the total. (**D**) Motifs enriched in the SMAD2/3 binding sites in TNBCs treated with TGFβ for 1.5 h.

Supplementary Figure S7. Working model. The TGFβ/SMAD-mediated induction of JUNB in premalignant cells causes redirection of SMAD binding to different sites on the genome which results in the activation of an invasion-mediating transcriptional program via a self-enabling mechanism. This

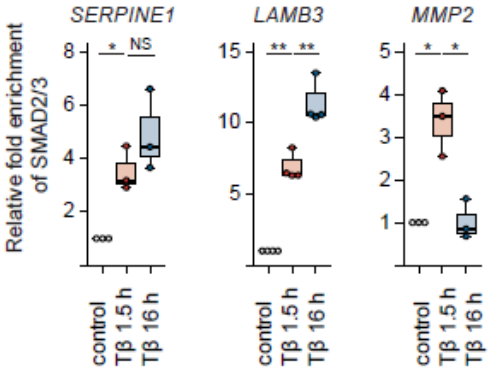
self enabling TGFβ/SMAD/JUNB-dependent transcriptional program will contribute to make the cell more migratory/invasive. One example of a TGFβ and JUNB-induced target gene activated by this mechanism is WNT7B. We suggest that in late phases of breast cancer the JUNB/WNT7B signaling pathway contributes to the tumor promoting function of TGFβ.

Supplementary Table S1. Primer sequences used for qRT-PCR. Primer sequences used for qRT-PCR are shown. Fw, forward primer; Rev, reversed primer.

Supplementary Table S2. Primer sequences used for ChIP-qPCR. Primer sequences used for ChIP-qPCR are shown. Fw, forward primer; Rev, reversed primer.

Supplementary Figure S1

A



B

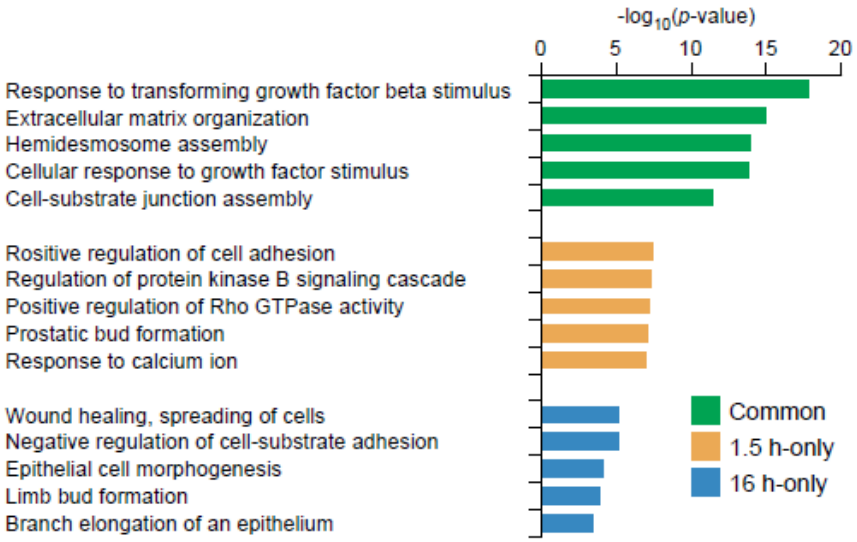
Enhancer (H3K4me1)

Condition	No. of peaks	HMEC (ENCODE)		MCF10A	
		No. of overlap	%	No. of overlap	%
1.5 h-only	1,475	1,140	77.3%	972	65.9%
16 h-only	2,631	946	36.0%	655	24.9%
Common	649	621	95.7%	552	85.1%

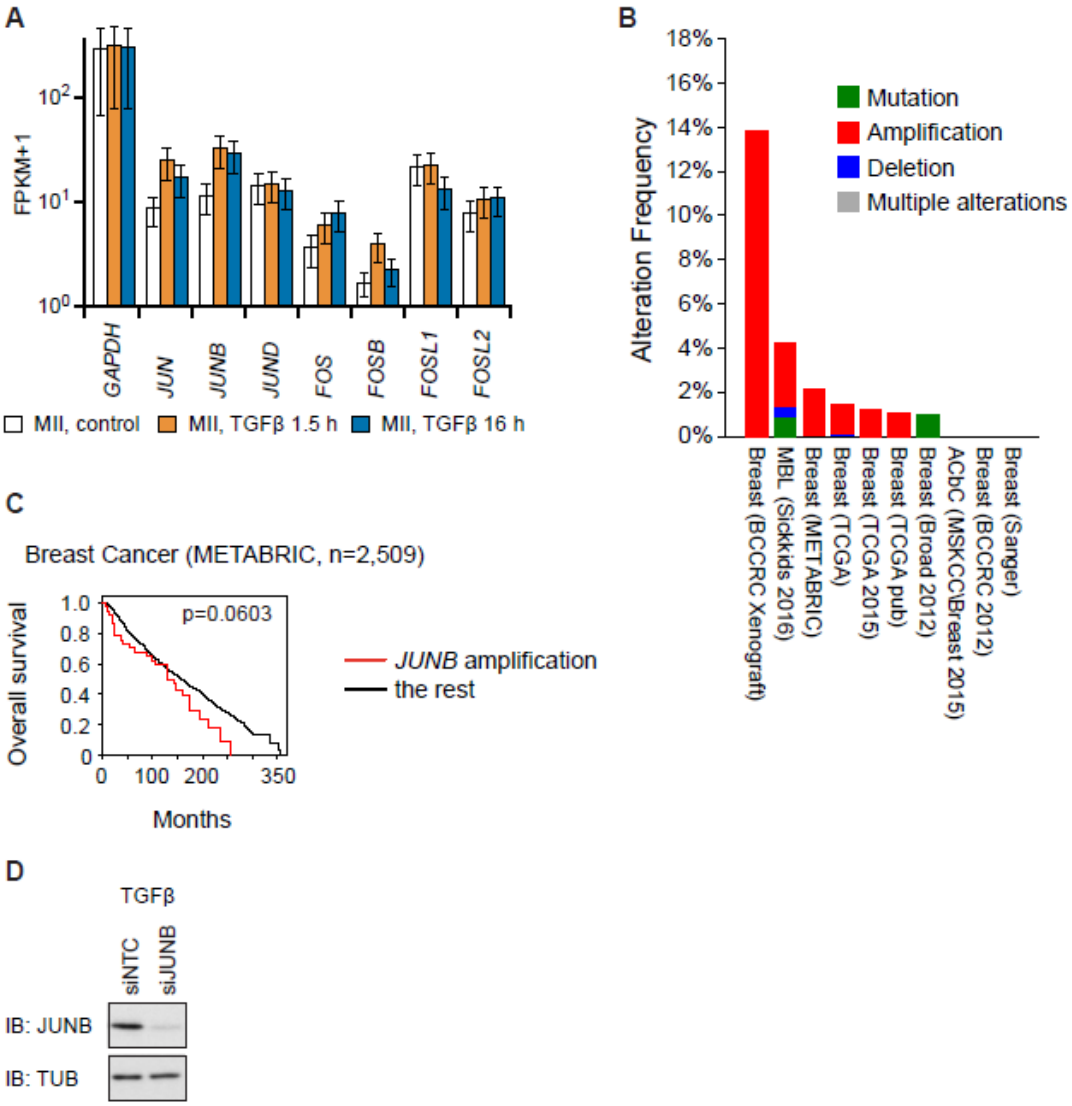
Promoter (H3K4me3)

Condition	No. of peaks	HMEC (ENCODE)		MCF10A	
		No. of overlap	%	No. of overlap	%
1.5 h-only	1,475	447	30.3%	74	5.0%
16 h-only	2,631	361	13.7%	96	3.6%
Common	649	297	45.8%	60	9.2%

C



Supplementary Figure S2



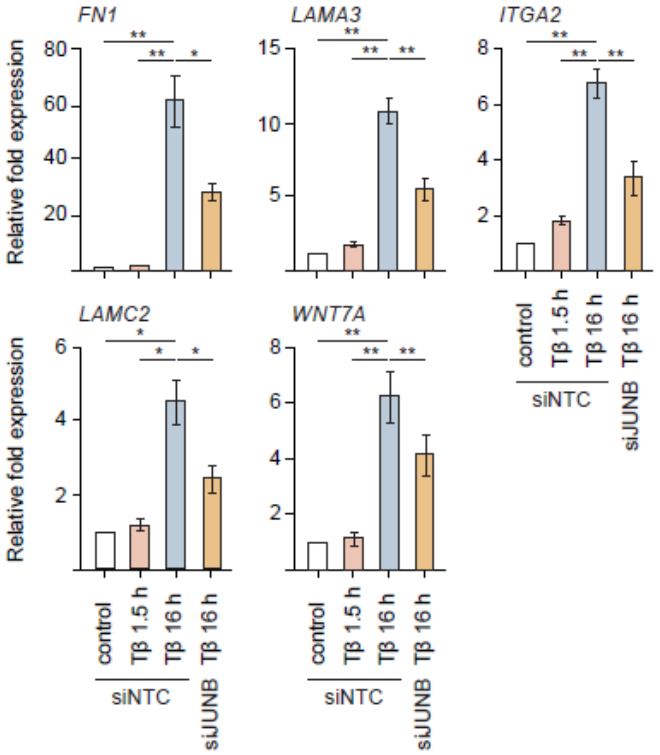
Supplementary Figure S3

A

Gene Symbol	TGFβ 16 h		Ratio
	siNTC	siJUNB	
PTGS1	3.50	0.64	0.18
FERMT1	8.10	1.69	0.21
PALM2-AKAP2	7.91	1.67	0.21
FN1	136.03	35.65	0.26
BPGM	7.31	2.09	0.29
ANTXR2	4.94	1.69	0.34
SERPINE1	49.18	17.09	0.35
LRRC8C	4.98	1.77	0.36
TGM2	12.27	4.47	0.36
SUSD1	2.36	0.88	0.37
WNT7B	5.43	2.06	0.38
COL1A1	9.87	3.83	0.39
WFDC3	2.05	0.83	0.40
SPOCK1	4.06	1.66	0.41
TMEM173	2.10	0.89	0.42
NT5E	4.41	1.94	0.44
ATXN1	2.89	1.28	0.44
TMEM117	2.63	1.19	0.45
FAM83A	2.85	1.30	0.46
ANPEP	3.24	1.53	0.47
FANK1	3.92	1.85	0.47
YAF2	2.48	1.18	0.48
FBXL17	2.23	1.06	0.48
ISG20	3.60	1.73	0.48
LAMA3	10.69	5.20	0.49
EPHB2	8.63	4.26	0.49
TGFBI	9.32	4.66	0.50

(fold change)

B

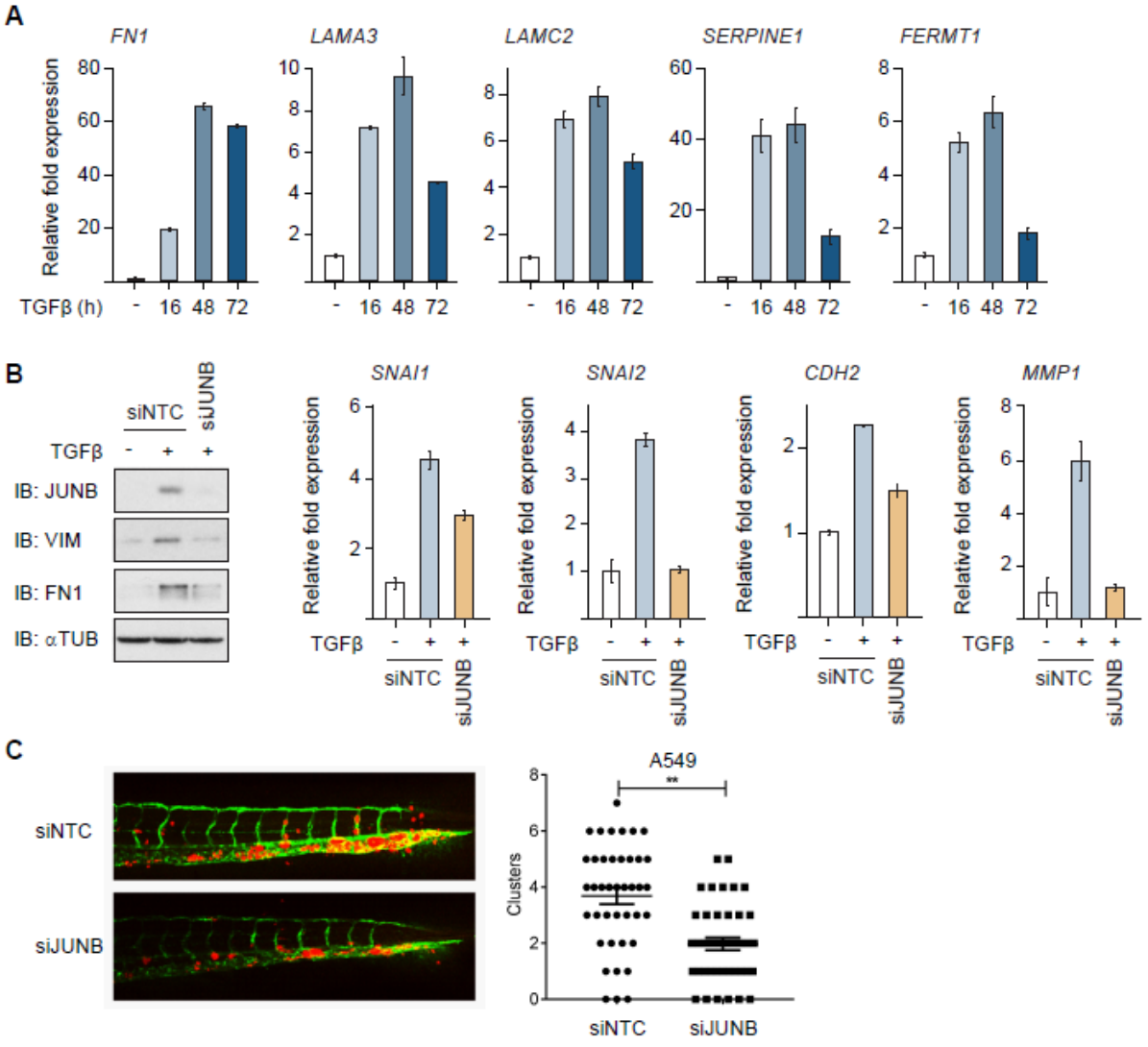


C

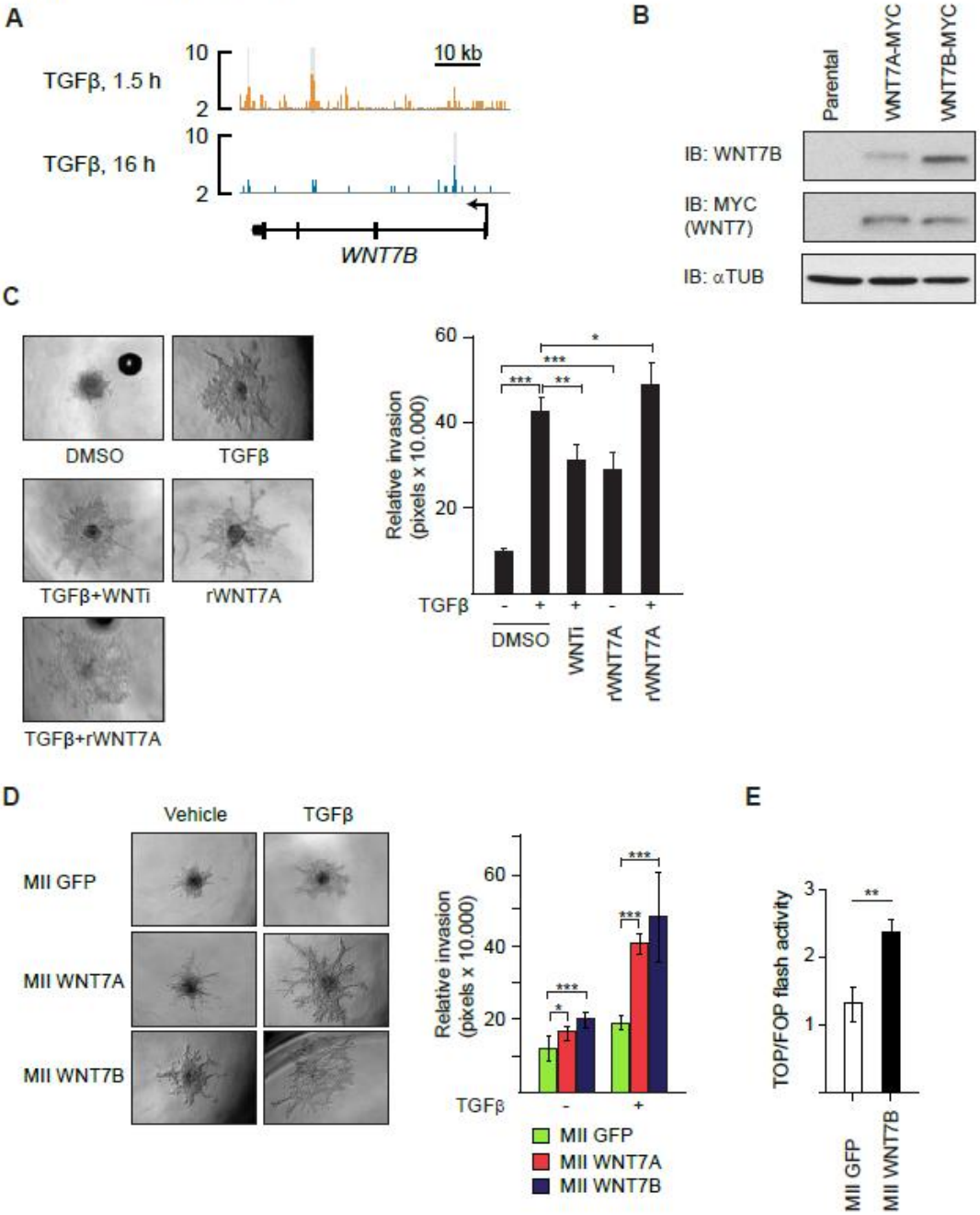
KEGG pathway	ES	p-value	FDR q-value
ECM receptor interaction	0.737	0.000	0.002
Focal adhesion	0.617	0.000	0.001
Dilated cardiomyopathy	0.781	0.000	0.006
Pathways in cancer	0.547	0.000	0.016
Small cell lung cancer	0.673	0.000	0.026
Hypertrophic cardiomyopathy	0.766	0.002	0.033
Arrhythmogenic right ventricular cardiomyopathy	0.724	0.003	0.066
Melanogenesis	0.670	0.009	0.049
Vasopressin regulated water reabsorption	0.816	0.009	0.122
Hedgehog signaling pathway	0.756	0.018	0.136
Cell adhesion molecules	0.785	0.021	0.144
WNT signaling pathway	0.544	0.025	0.147
Tight junction	0.561	0.040	0.189
Prostate cancer	0.609	0.041	0.204
Complement and coagulation cascades	0.780	0.044	0.211
Spliceosome	-0.684	0.007	0.057

(n=1,436 genes, with SMAD2/3-binding after TGFβ 16 h)

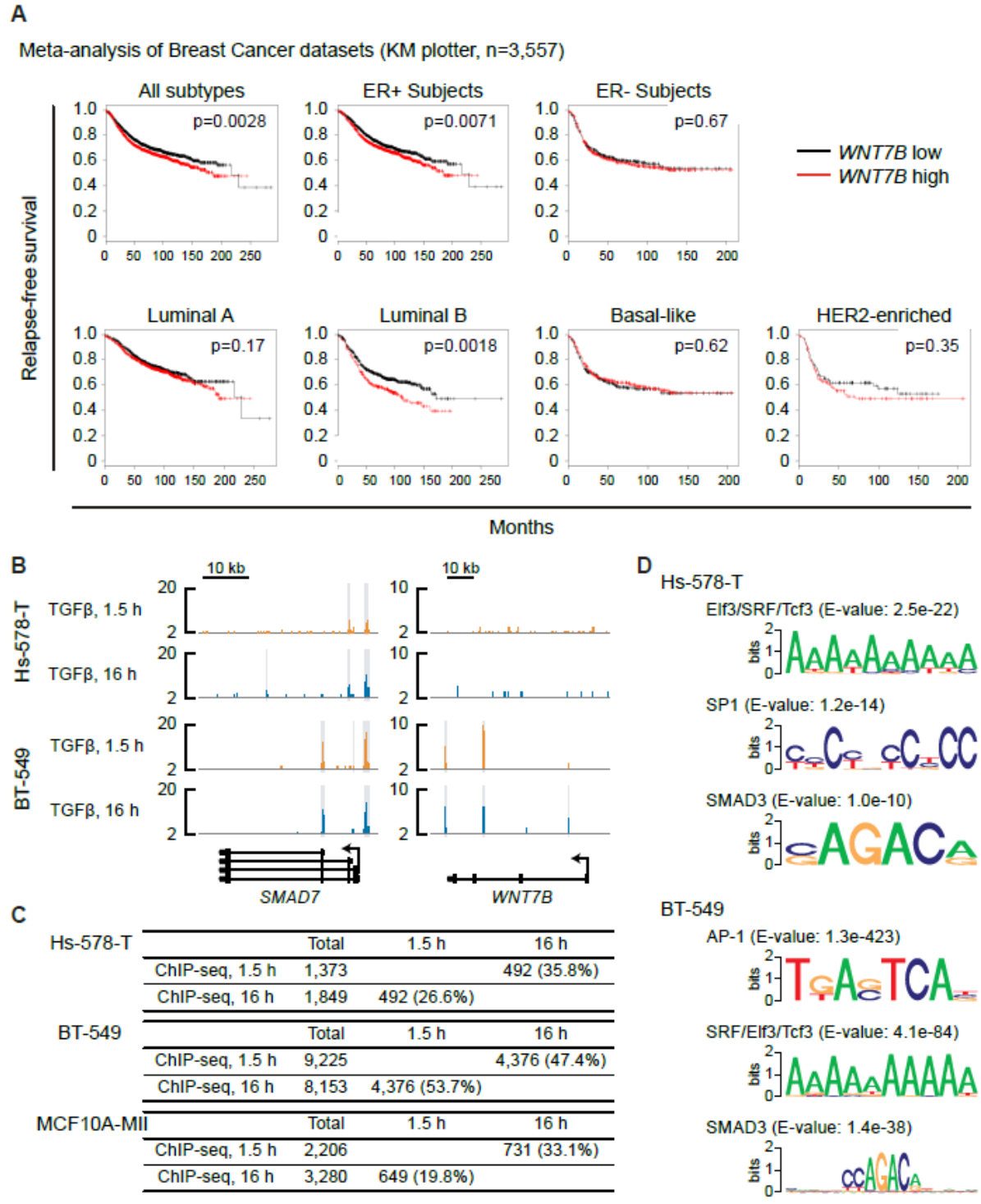
Supplementary Figure S4



Supplementary Figure S5



Supplementary Figure S6

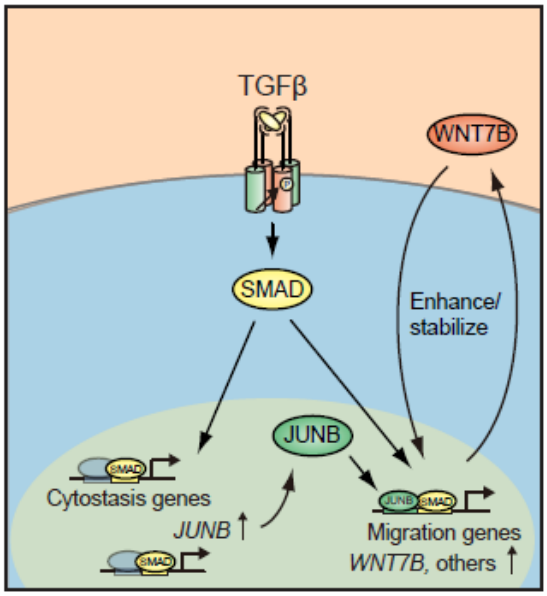


Hs-578-T	Total	1.5 h	16 h
ChIP-seq, 1.5 h	1,373		492 (35.8%)
ChIP-seq, 16 h	1,849	492 (26.6%)	

BT-549	Total	1.5 h	16 h
ChIP-seq, 1.5 h	9,225		4,376 (47.4%)
ChIP-seq, 16 h	8,153	4,376 (53.7%)	

MCF10A-MII	Total	1.5 h	16 h
ChIP-seq, 1.5 h	2,206		731 (33.1%)
ChIP-seq, 16 h	3,280	649 (19.8%)	

Supplementary Figure S7



Supplementary Table S1

Primer sequences for RT-qPCR	
Name	**Sequence**
***CDH2* Fw**	5'-CCTGCTTCAGGCGTCTGTAGA-3'
***CDH2* Rev**	5'-TCATGCACATCCTTCGATAAGACT-3'
***FERMT1* Fw**	5'-CTTGGTTCAGTGACAGCCCT-3'
***FERMT1* Rev**	5'-GGAGTCTAGCCAACCTGCAT-3'
***FN1* Fw**	5'-CATCGAGCGGATCTGGCCC-3'
***FN1*Rev**	5'-GCAGCTGACTCCGTTGCCCA-3'
***GAPDH* Fw**	5'-GGAGTCAACGGATTTGGTCGTA-3'
***GAPDH* Rev**	5'-GGCAACAATATCCACTTTACCA-3'
***ITGA2* Fw**	5'-GCTGGTGCTCCTCGGGCAAA-3'
***ITGA2* Rev**	5'-TGGTCACCTCGGTGAGCCTGA-3'
***LAMA3* transcript variants 2and 4 Fw**	5'-CCTGGGGCAGTGTCTGGGCT-3'
***LAMA3* transcript variants 2and 4 Rev**	5'-TCCCGCGGTGTTGTGCTGAC-3'
***LAMB3* Fw**	5'-ACGGCAGAACACACAGCAAGGA-3'
***LAMB3* Rev**	5'-ACCGGGTCCTCCCAACAAGCA-3'
***LAMC2* transcript variant 1 Fw**	5'-CATCTGATGGACCAGCCTCTC-3'
***LAMC2* transcript variant 1 Rev**	5'-GCAGTTGGCTGTTGATCTGG-3'
***MMP1* Fw**	5'-CCAAATGGGCTTGAAGCT-3'
***MMP1* Rev**	5'-GTAGCACATTCTGTCCCTAA-3'
***MMP2* Fw**	5'-AGATGCCTGGAATGCCAT-3'
***MMP2* Rev**	5'-GGTTCTCCAGCTTCAGGTAAT-3'
***SERPINE1* Fw**	5'-GAGACAGGCAGCTCGGATTC-3'
***SERPINE1* Rev**	5'-GGCCTCCCAAAGTGCATTAC-3'
***SNAI1* Fw**	5'-CACTATGCCGCGCTCTTTC-3'
***SNAI1* Rev**	5'-GCTGGAAGGTAAACTCTGGATTAGA-3'
***SNAI2* Fw**	5'-AGACCCTGGTTGCTTCAAGGA-3'
***SNAI2* Rev**	5'--3'
***WNT7B* Fw**	5'-AAGCTCGGAGCACTGTCATC-3'
***WNT7B* Rev**	5'-ACTGGTACTGGCACTCGTTG-3'

Supplementary Table S2

Primer sequences for ChIP-qPCR	
Name	**Sequence**
***HBB* Fw**	5´-AACGTGATCGCCTTTCTC-3´
***HBB* Rev**	5´-GAAGCAGAACTCTGCACTTC-3´
***HPRT1* Fw**	5´-TGTTTGGGCTATTTACTAGTTG- 3’
***HPRT1* Rev**	5-ATAAAATGACTTAAGCCCAGAG-3’
***SERPINE1* Fw**	5'-GCAGGACATCCGGGAGAGA-3'
***SERPINE1* Rev**	5'-CCAATAGCCTTGGCCTGAGA-3'
***LAMB3* Fw**	5'-TTGCCCTGCACTACAACACA-3'
***LAMB3* Rev**	5'-GTAACACACCAGGCCCACTT-3'
***MMP2* Fw**	5'-TCCCAGGCCTGCCCATGTCA-3'
***MMP2* Rev**	5'-GGAGCTGGTGGGTGGAAAGCC-3'
***WNT7B* Fw**	5'-TCACCCATGACTCACTTGGC-3'
***WNT7B* Rev**	5'-AGGTCTCTTCCGCTCTCAGT-3'