Supporting Information

Mizuno et al. 10.1073/pnas.1017001108

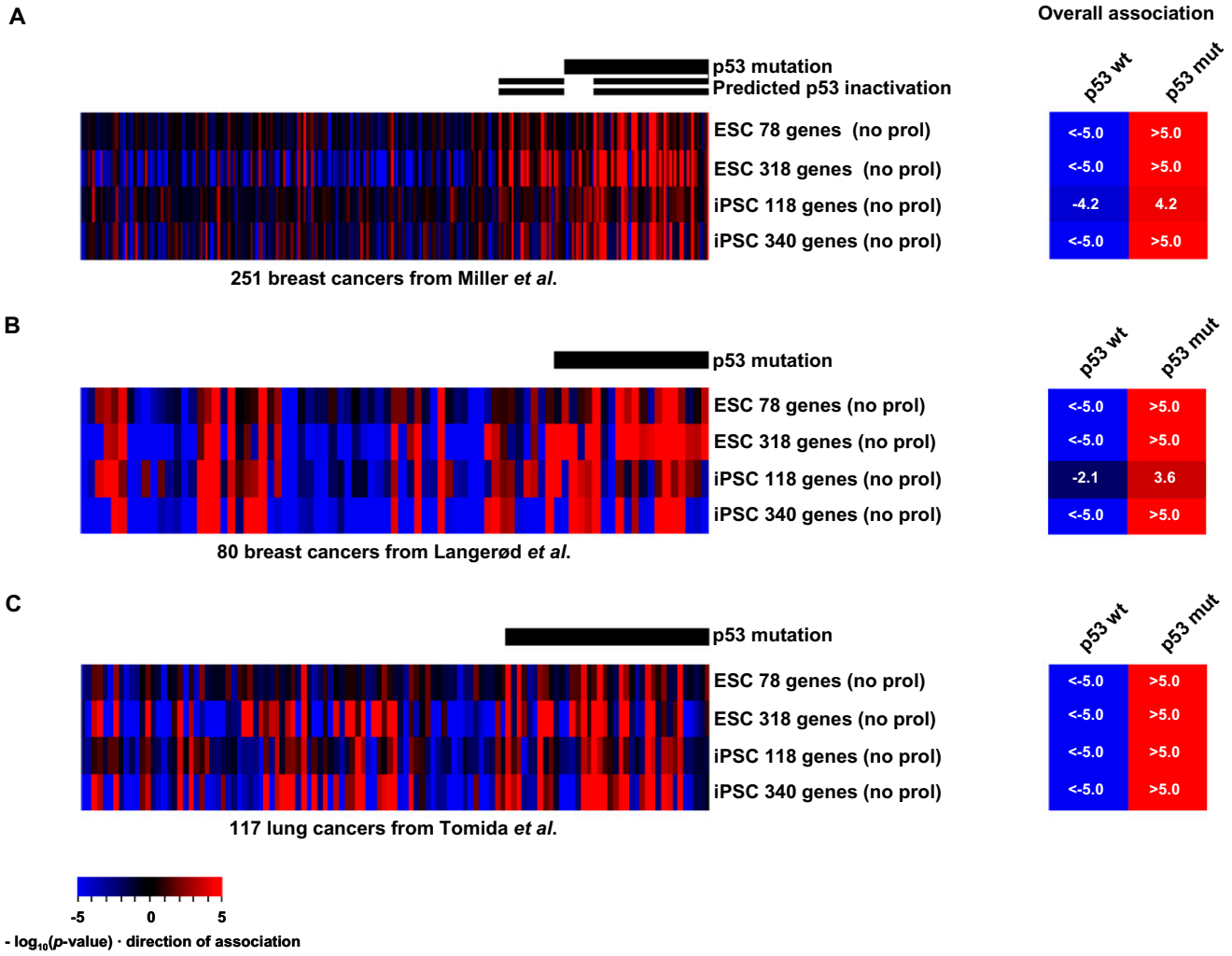


Fig. S1. Association scores from different sizes of ESC/iPSC signatures in (*A*) the Miller et al. (1) and (*B*) the Langerød et al. (2) datasets, and (*C*) Tomida et al. (3) datasets. Overall associations (*Right*) were assessed using representative profiles.

1. Miller LD, et al. (2005) An expression signature for p53 status in human breast cancer predicts mutation status, transcriptional effects, and patient survival. *Proc Natl Acad Sci USA* 102: 13550–13555.
2. Langerød A, et al. (2007) TP53 mutation status and gene expression profiles are powerful prognostic markers of breast cancer. *Breast Cancer Res* 9:R30.
3. Tomida S, et al. (2008) Relapse-related molecular signature in lung adeno carcinomas identify patients with dismal prognosis. *J Clin Oncol* 27:2793–2799.

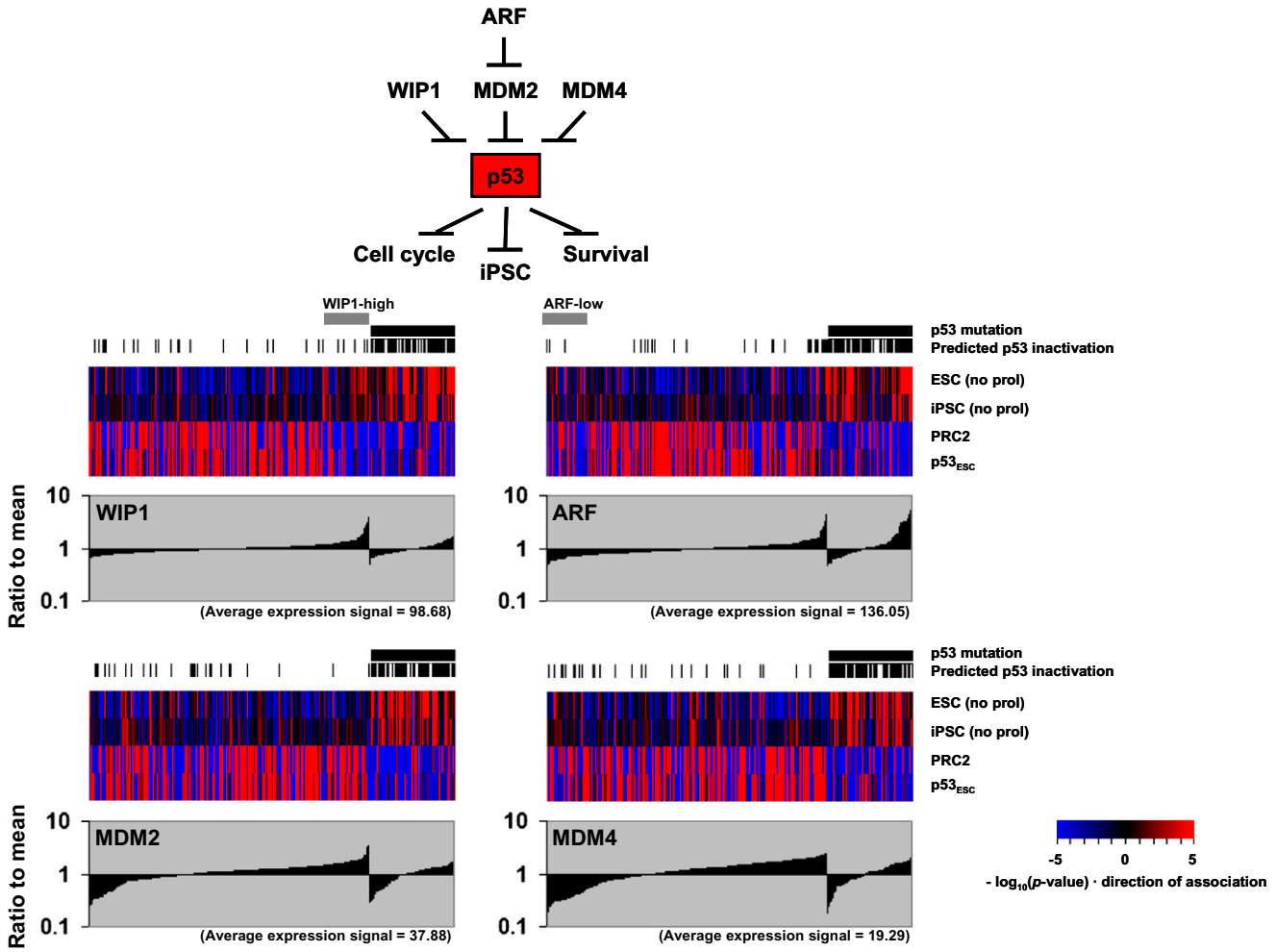
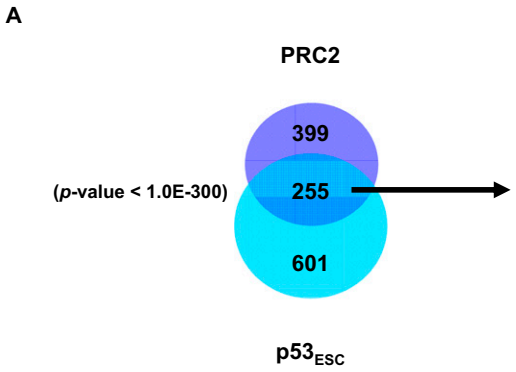
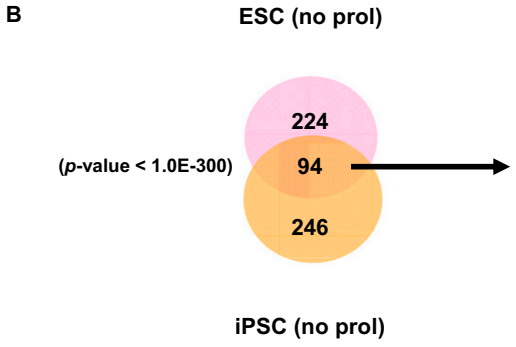


Fig. S2. *WIP1*, *ARF*, *MDM2*, and *MDM4* are known upstream regulators of p53 activity and are shown diagrammatically at the top. Tumors are sorted according to p53 mutation status and relative expression levels of these upstream p53 regulators. Association scores for ESC, iPSC, PRC2, and $p53_{ESC}$ signatures are shown. Gray areas over the p53 wild-type tumors in the *WIP1* and *ARF* plots show the demarcation of the highest and lowest 15% of expressers, respectively.



Top 20 categories	p-value
multicellular organismal development	6.37E-75
system development	3.15E-74
anatomical structure development	5.31E-73
developmental process	3.05E-71
organ development	6.31E-62
multicellular organismal process	1.58E-60
anatomical structure morphogenesis	2.42E-54
cell differentiation	2.78E-54
nervous system development	1.84E-52
cellular developmental process	3.03E-52
regulation of transcription, DNA-dependent	1.48E-48
regulation of RNA metabolic process	1.71E-47
regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolic process	3.34E-42
neurogenesis	4.35E-42
regulation of biosynthetic process	8.99E-42
regulation of nitrogen compound metabolic process	9.79E-42
regulation of cellular biosynthetic process	2.49E-41
generation of neurons	4.90E-40
central nervous system development	5.54E-40
pattern specification process	1.20E-39



Top 20 categories	p-value
DNA metabolic process	4.80E-05
developmental process	7.16E-05
organelle organization	9.64E-05
multicellular organismal development	1.63E-04
multicellular organismal process	3.84E-04
DNA recombination	4.23E-04
anatomical structure development	5.91E-04
cellular component organization	0.00114
cell development	0.00162
system development	0.00208
DNA-dependent DNA replication	0.00512
somatic recombination of immunoglobulin gene segments	0.00515
chromatin silencing	0.00515
cellular developmental process	0.00545
negative regulation of gene expression, epigenetic	0.00699
female gamete generation	0.00704
interspecies interaction between organisms	0.00718
somatic diversification of immunoglobulins	0.00765
cell differentiation	0.00788
somatic diversification of immune receptors via germline recombination within a single locus	0.00908

Fig. S3. Venn diagram for (*A*) $p53_{ESC}$ versus PRC2 signatures and for (*B*) ESC versus iPSC signatures. In each case, the overlap in genes is highly significant (χ^2 test). The top 20 gene-annotation categories identified among overlapping genes are listed at the right. These represent biological states and activities with likely functional relevance for samples that exhibit their coordinate up- or down-regulation.

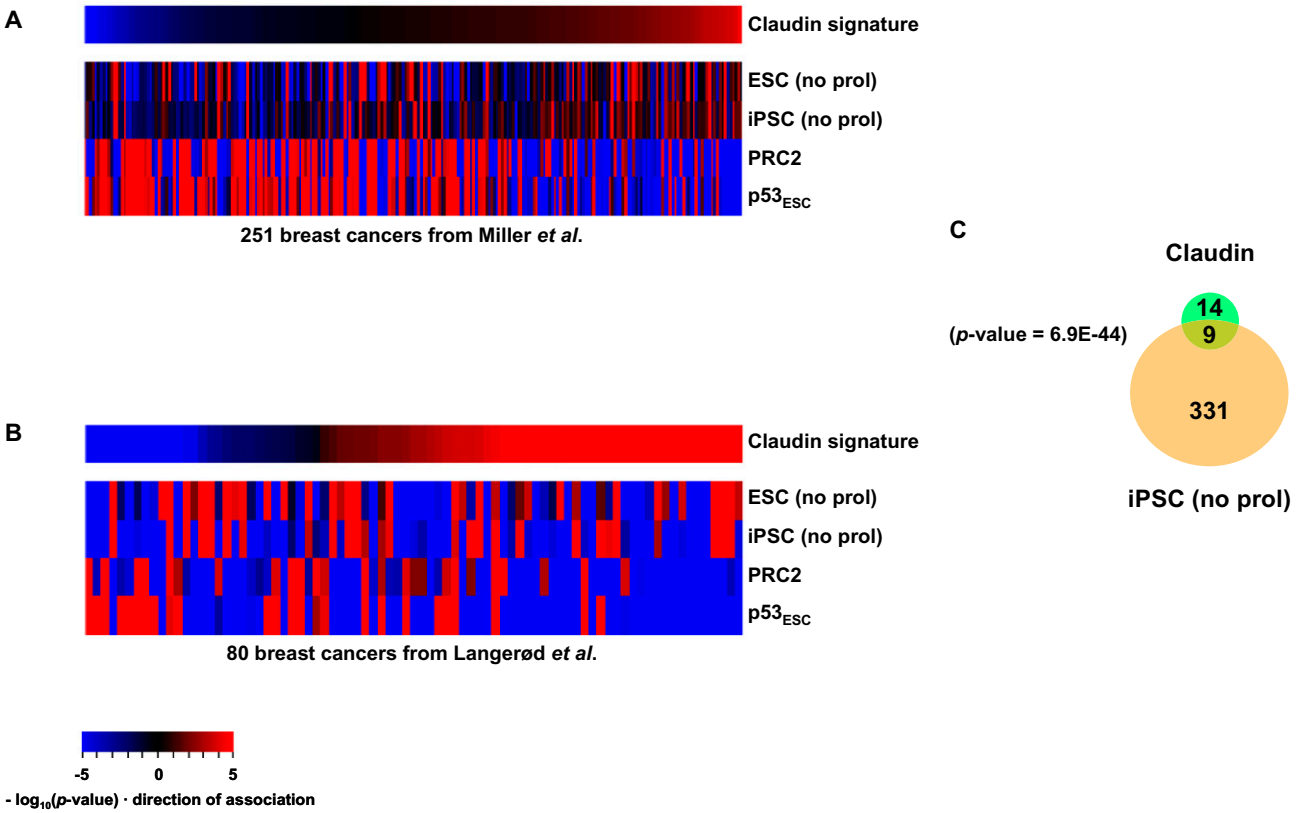


Fig. S4. Breast cancers were ordered according to association scores for the claudin signature (1) in (*A*) the Miller et al. (2) and (*B*) the Langerød et al. (3) datasets. Scores for ESC, iPSC, PRC2, and $p53_{ESC}$ signatures are also shown. (*C*) Venn diagram and overlap statistics (χ^2 test) for the claudin signature (1) versus the iPSC signature. The claudin signature used comprises the following genes: *APOE, BSRPY, CDH1, CGN, CLDN3, CLDN4, CLDN7, ELF3, EPCAM, EPN3, ESRP1, FXYD3, KRT19, MAL2, MB, MY06, NEBL, OCLN, PRR15L, SHROOM3, SPINT2, TOMIL1,* and *TRPS1* (1).

1. Hennessy BT, et al. (2009) Characterization of a naturally occurring breast cancer subset enriched in epithelial-to-mesenchymal transition and stem cell characteristics. *Cancer Res* 69: 4116–4124.
2. Miller LD, et al. (2005) An expression signature for p53 status in human breast cancer predicts mutation status, transcriptional effects, and patient survival. *Proc Natl Acad Sci USA* 102: 13550–13555.
3. Langerød A, et al. (2007) TP53 mutation status and gene expression profiles are powerful prognostic markers of breast cancer. *Breast Cancer Res* 9:R30.

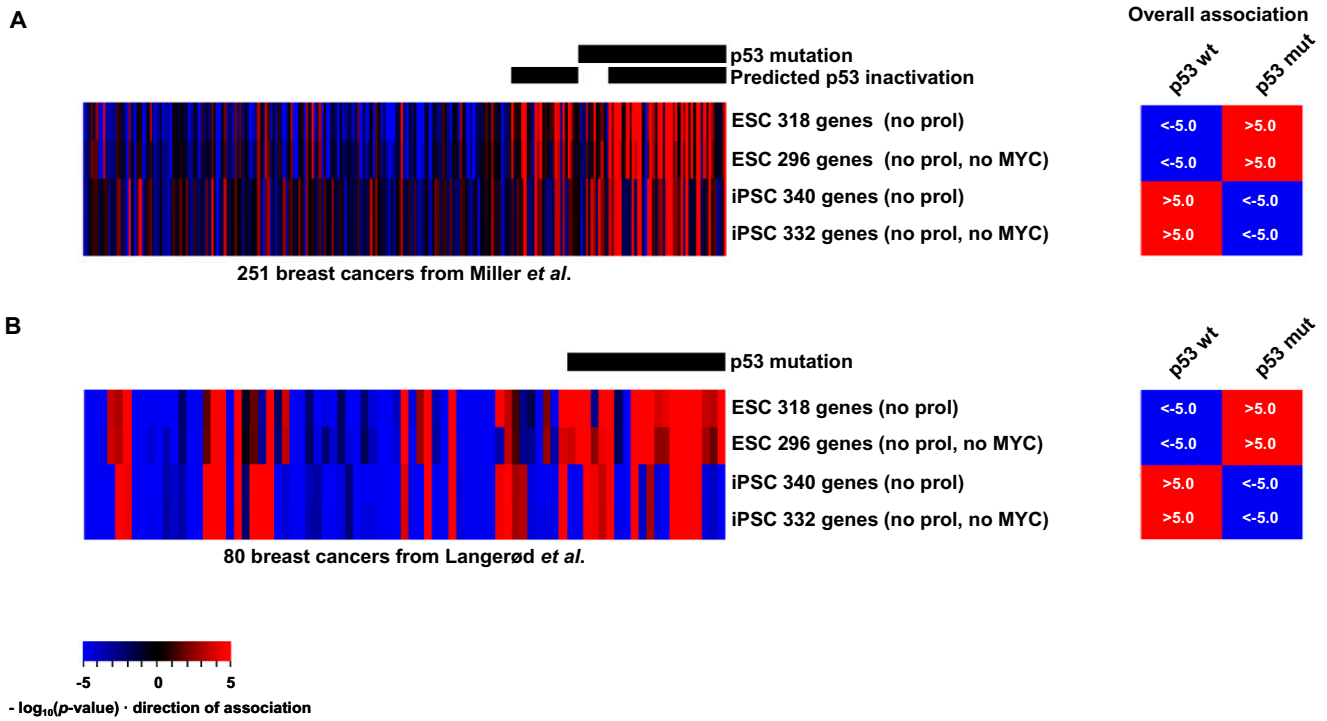


Fig. S5. Association scores from ESC/iPSC signatures after the removal of *MYC*-network genes (1) for (*A*) the Miller et al. (2) and (*B*) the Langerød et al. (3) datasets. Overall associations (*Right*) were assessed using representative profiles.

1. Kim J, et al. (2010) A Myc network accounts for similarities between embryonic stem and cancer cell transcription programs. *Cell* 143:313–324.
2. Miller LD, et al. (2005) An expression signature for p53 status in human breast cancer predicts mutation status, transcriptional effects, and patient survival. *Proc Natl Acad Sci USA* 102: 13550–13555.
3. Langerød A, et al. (2007) TP53 mutation status and gene expression profiles are powerful prognostic markers of breast cancer. *Breast Cancer Res* 9:R30.

Table S1. Datasets used for iPSC signature generation

Table S1

Table S2. Genes up-regulated in various iPSCs

Table S2