SUPPLEMENTARY MATERIAL

Supplementary Methods

Study population

The retrospective consecutive cohort of 2,453 female BC patients was obtained from a systematic retrieval of information from a dedicated database containing all consecutive BC patients operated at the European Institute of Oncology (IEO) in Milan between 1997 and 2000, which led to the selection of 3,396 patients based on the following inclusion criteria: i) operable BC, ii) female, iii) BC as first primary tumour, with no history of a previous malignancy, iv) no distant metastasis at diagnosis and v) no neoadjuvant therapy (Supplementary Table S3).

Available clinicopathological information included age, date at surgery, tumour characteristics (histological type, tumour size, nodal involvement, grade, perivascular infiltration, Ki-67 and ER/PgR expression) and treatment modality (type of surgery, adjuvant radiotherapy, hormonal therapy, chemotherapy) (Supplementary Table S3). In this cohort, ER, PgR and Ki-67 were measured by immunohistochemistry (IHC) on whole tissue sections, and were retrieved from histopathological reports. Conversely, IHC analysis of HER2 expression was repeated *ad hoc* for the purpose of this study.

Paraffin blocks were available for 2,935 patients, due to the exclusion of four hundred sixty one blocks showing biological degradation of the paraffin-embedded tissue material. After review of haematoxylin and eosin-stained slides, four hundred eighty-two tissue blocks were excluded for insufficient tumour cellularity (<60% cellularity), minimal areas of infiltrating carcinoma with respect to *in situ* carcinoma areas, and/or presence of massive inflammatory infiltration or massive necrosis, resulting in the final selection of 2,453 patients who represent the "IEO BC 97-00" study cohort (described in Supplementary Table S3).

The comparison of the demographics and clinicopathological characteristics between the initial selection of 3,396 patients and the final "IEO BC 97-00" study cohort showed that the two populations were similar, with no substantial differences in any of the variables analysed (Supplementary Table S3).

From the study cohort of 2,453 patients, successful extraction of a sufficient amount of RNA suitable to multiplex RT-qPCR analysis was possible for 2,335 patients. After quality control of RT-qPCR data, a total of 19 samples (19/2,335=0·8%) were excluded from statistical analyses because of spurious RT-qPCR results, likely due to poor quality mRNA. Therefore, a total of 2,316 patients were finally included in the statistical analyses.

Patient follow-up included physical examination, annual mammography and breast ultrasound every 6 months, blood tests every 6-12 months, and further evaluations only in the case of symptoms. When possible, the status of women who had not attended their scheduled follow-up visits for more than one year was obtained by telephone contact.

The different molecular subtypes of BC (i.e., Luminal, HER2+ and Triple-Negative) were defined according to the St. Gallen 2015 classification, and identified through the immunohistochemical surrogate panel including ER/PgR, HER2 and Ki-67.[8,19]

Quantitative real-time PCR analysis

Total mRNA was extracted from FFPE samples using a single core of 1·5 mm in diameter (2,157 samples), or at least two 10-µm thick sections (178 samples) when tumour size was limited. Tissue cores and/or sections were selected by a pathologist. Total RNA was extracted using the AllPrep DNA/RNA FFPE Kit automated on the QIAcube (Qiagen), following manufacturer's instructions.

RT-qPCR was performed with an in-house custom designed TaqMan® Array using hydrolysis probes (ThermoFisher Scientific) in combination with the SsoAdvanced Universal Probes Supermix (Bio-Rad Laboratories) in a final volume of 10 µl in 384-well plates. PCR reactions were run in LightCycler (LC) 480 real-time PCR instruments (Roche) using the following thermal cycling conditions: 1 cycle at 95°C for 30 sec, 45 cycles at 95°C for 5 sec and 60°C for 30 sec. TaqMan gene expression assays were selected based on amplicon size (<100 bp), and on their ability to detect the Ref Seq identified in the Affymetrix meta-analysis and as many isoforms as possible. Custom TaqMan assays were designed, when possible, in the 3' region of the gene using the Primer Express Software V3·0 (ThermoFisher Scientific) (see Supplementary Table S4 for a detailed list). Standard methods for RT-qPCR data mining and manufacturer's recommendations for quality control and sample rejection were used in the analysis of RT-qPCR data. Each target was assayed in triplicate and average Cq (AVG Cq) values were calculated from triplicate values, when the standard deviation was <0·4, or from the best duplicate values when the standard deviation was ≥0·4. AVG Cq values were calculated and normalized using four reference genes (*HPRT1*, *GAPDH*, *GUSB* and *TBP*). The normalized Cq ($Cq_{normalized}$) of each target gene was calculated using the following formula: $Cq_{normalized}$ = AVG Cq - SF, where: SF is the difference between the AVG Cq value of reference genes for each patient and a constant reference value K, which represents the mean of the AVG Cq of the 4 reference genes calculated across all samples (K = 25·012586069). This

normalization strategy allows the retention of information about the abundance of the original transcript, as measured by RT-qPCR (i.e., in Cq scale), which is conversely lost when using the more classical ΔCq method. A Cq=35 was considered as our limit of detection and Cq values beyond this limit were set to 35 and normalization was omitted.

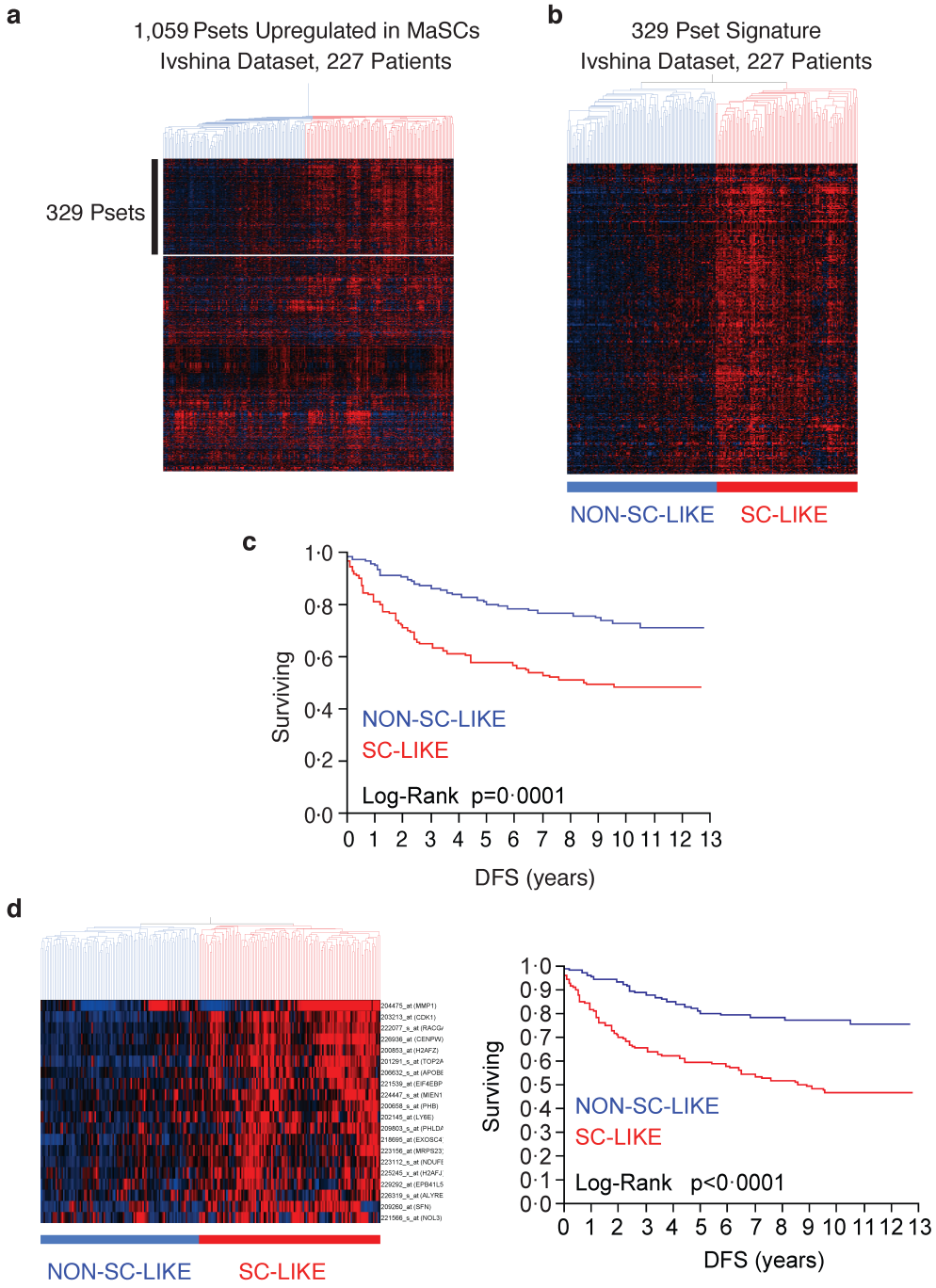
Rationale for the use of the tumoursphere serial propagation assay as a tool to investigate the biological basis of the 20-gene SC signature

The tumoursphere propagation assay was performed as previously described,[6,7] starting from patient-derived primary tumour cells freshly isolated from biopsy tumour specimens (the prospective cohort of 90 patients, described in Supplementary Table 9).

This assay is based on the ability of CSCs to survive in suspension culture conditions, giving origin to three-dimensional spheroid structures that can be serially dissociated and propagated through several passages. In previous work, we established that the efficiency of patient-derived primary tumour cells to sustain the formation of several generations of tumourspheres *in vitro* depends on their intrinsic propensity to self-renew and proliferate, and provides a quantitative estimate of the number of CSCs in individual BCs.[6,7] Indeed, upon tumoursphere dissociation and serial propagation, the cumulative number of tumourspheres over several passages can either exponentially increase (what we call an "unlimited" phenotype) or progressively extinguish ("self-limiting" phenotype") (examples of actual tumours showing these two opposite biological behaviours are in Fig. 3c of the main text). Of note, these *in vitro*-determined phenotypes stringently correlate with the biological aggressiveness of the tumour *in vivo*, as formally demonstrated in limiting-dilution transplantation experiments in our previous studies.[6,7] Mechanistically, we demonstrated that different modalities of the CSC self-renewing division are responsible for the two phenotypes: CSCs that preferentially divide symmetrically (1 CSC -> 2CSCs) give rise to an unlimited self-renewal phenotype, while CSCs that preferentially divide asymmetrically (1 CSC -> 1 CSC + 1 non-tumourigenic progenitor) produce a self-limiting phenotype.[6]

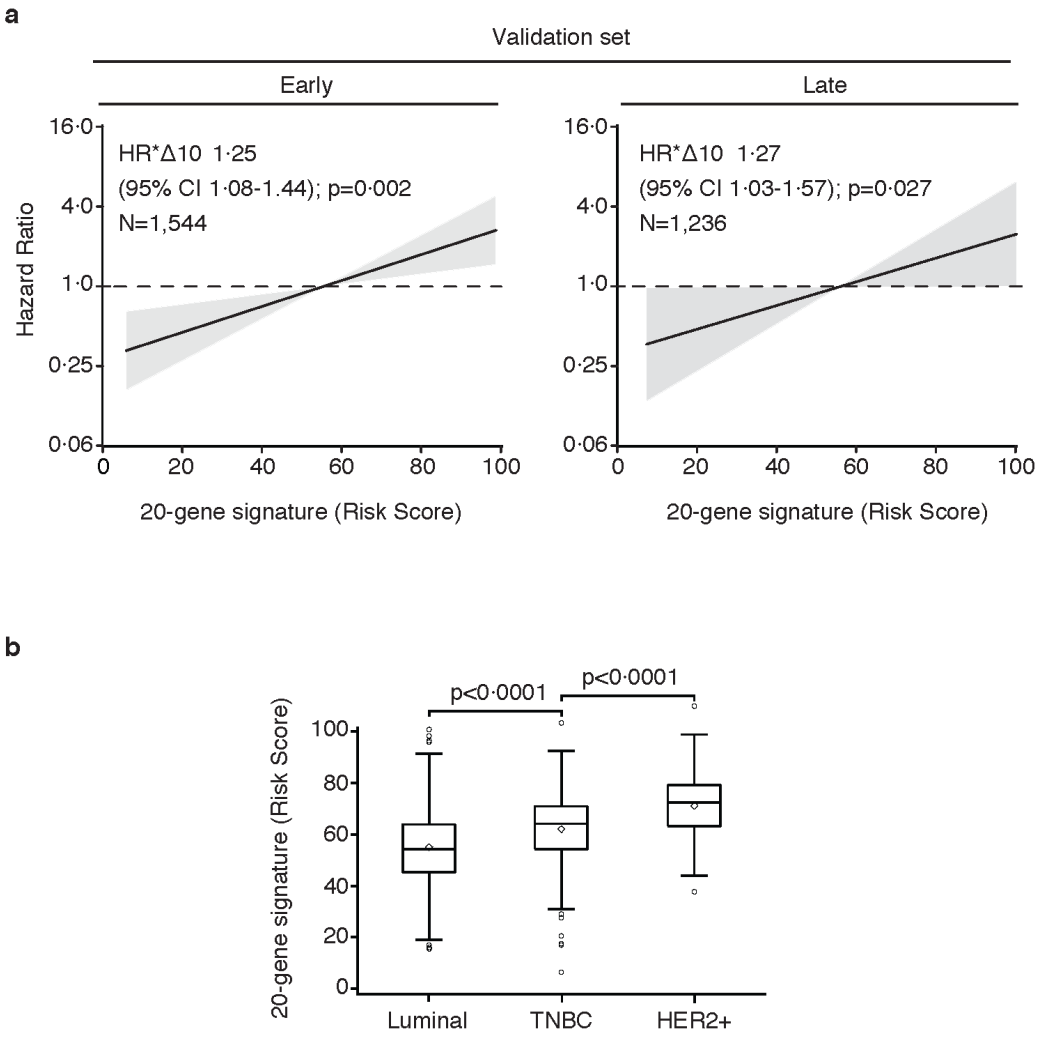
SUPPLEMENTARY FIGURES

Supplementary Figure S1



Legend to Supplementary Figure S1. Derivation and initial analytical validation of the 20-gene SC signature. **a.** Distribution of 227 BC patients from the Ivshina *et al*. dataset[9] by unsupervised hierarchical clustering analysis using the 1,059 probesets upregulated in MaSCs *vs*. progenitors, identified in previous studies.[7] Two groups of BC patients could be identified based on their similarity to the expression pattern of 329 probesets (indicated by the bar on the left of the cluster). **b.** Unsupervised hierarchical clustering analysis of the 227 BC patients from the Ivshina *et al.* dataset[9] using the 329-signature, which allowed the identification of two subgroups of patients with distinct "SC-like" and "Non-SC-like" characteristics. **c.** Kaplan-Meier analysis of disease-free survival (DFS) in the "SC-like" and "Non-SC-like" subgroups of BC patients from the Ivshina *et al.* dataset,[9] according to the 329-signature. **d.** The 20 most highly expressed probesets of the 329-signature (see Supplementary Table S1) were used for the hierarchical clustering analysis (left panel) and the Kaplan-Meier analysis of DFS (right panel) of the "SC-like" and "Non-SC"-like BC patients from the Ivshina *et al*. dataset.[9]

Supplementary Figure S2

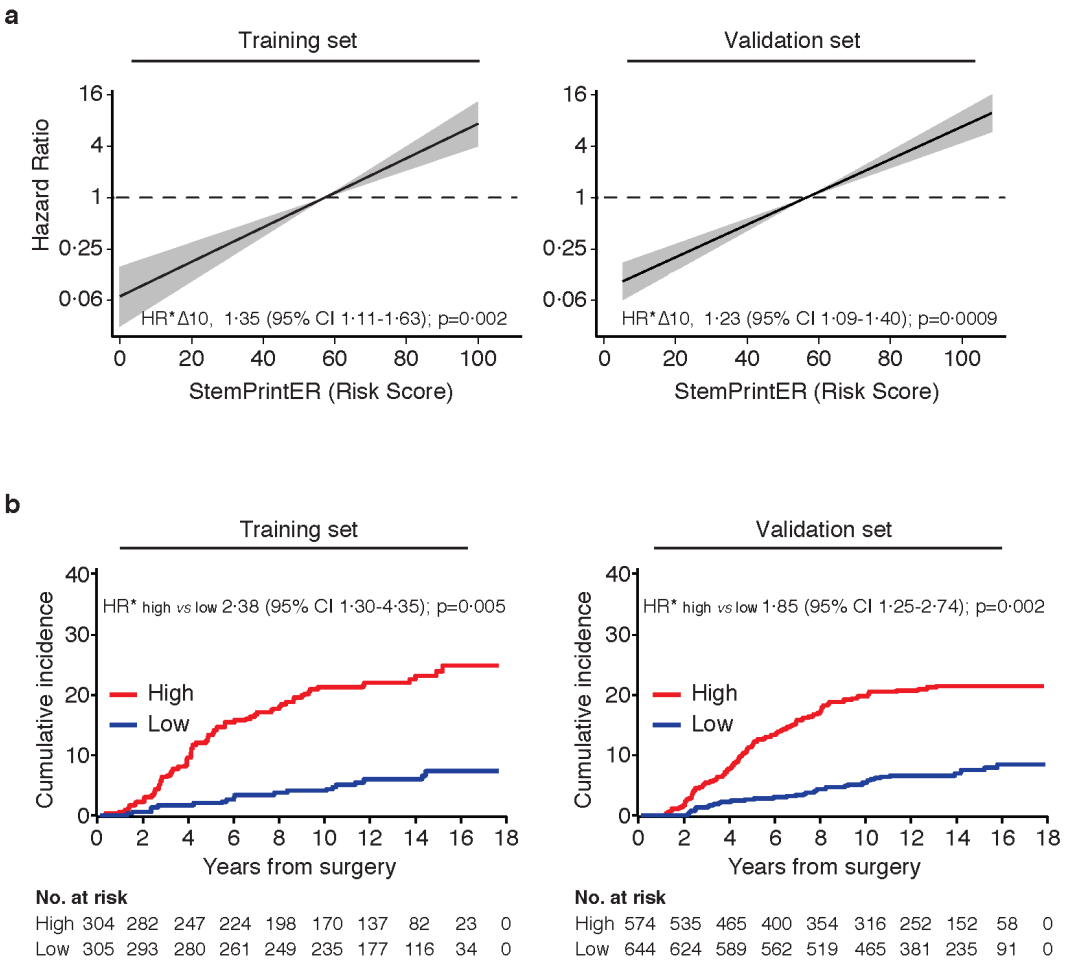


Legend to Supplementary Figure S2. Analysis of the clinical validity of the 20-gene SC signature in the IEO BC 97-00 study cohort.

a. The 20-signature risk score is an independent predictor of early (0-5 years) and late (5-10 years) distant metastasis in the "IEO BC 00-97" validation set (N=1,544). Shown are multivariable hazard ratios estimated using the risk score as a continuous variable with a 10-unit increase (HR*Δ10) in a multivariable model adjusted for Grade (G1, G2 and G3), Ki-67 (Ki-67<14% and Ki-67≥14%), HER2 status (positive and negative), ER/PgR status [not expressed (Both 0) and expressed (ER>0 or PgR>0)], tumour size (pT1 and pT2-3-4), number of positive lymph nodes (pN0, pN1-2-3 and pNX) and age at surgery (<50 and ≥50). Grey shaded areas represent the 95% CI. Vertical axis (HR) is plotted in logarithmic scale. HR, hazard ratio; CI, confidence interval; p, p -value.

b. Box plots showing the distribution of the 20-gene SC continuous risk score across the indicated BC molecular subtypes in the "IEO BC 97-00" validation set. TNBC, Triple-Negative BC; HER2+, HER2-positive BC. Differences in the distribution of continuous risk score between groups were evaluated using a linear regression model.

Supplementary Figure S3



Legend to Supplementary Figure S3. Prognostic value of StemPrintER in the ER+/HER2- training and validation sets.

a. The StemPrintER risk score, used as a continuous function (10-unit increase, Δ10) over the entire follow-up interval, provides prognostic information beyond that obtained from traditional clinicopathological parameters. Shown are multivariable hazard ratios (HR*) estimated with a multivariable model adjusted for Grade (G1, G2 and G3), Ki-67 (Ki-67<14% and Ki-67≥14%), HER2 status (positive and negative), ER/PgR status [not expressed (Both 0) and expressed (ER>0 or PgR>0)], tumour size (pT1 and pT2-3-4), number of positive lymph nodes (pN0, pN1-2-3 and pNX) and age at surgery (<50 and ≥50). Grey shaded areas represent the 95% CI. Vertical axis is plotted in logarithmic scale. HR, hazard ratio; CI, confidence interval; p, p -value.

b. Cumulative incidence of distant metastasis determined using StemPrintER as a 2-class risk model (High, Low) in the training and validation sets over the entire follow-up interval. Multivariable hazard ratios (High *vs.* Low) were estimated as in (a). HR*, multivariable hazard ratio; CI, confidence interval; p, p-value; "No. at risk", number of patients at risk.

SUPPLEMENTARY TABLES

Supplementary Table S1 (available as excels file). Derivation of a BC prognostic 20-signature based on the human MaSC profile using the Ivshina BC Dataset.

Sheet "Ivshina Dataset". Clinicopathological and molecular characteristics of the BC patient cohort from the Ivshina *et al*. dataset.[9] For each patient, the "SC-like" status for both the 20-signature (column D) and the 329-signature (column E) is also indicated.

Sheet "329 pset Ivshina". List of the probesets upregulated in human normal MaSCs *vs*. progenitors,[7] which identify a subgroup of a SC-like BC patients in the unsupervised hierarchical clustering of the Ivshina *et al*. dataset[9] (see Supplementary Fig. 1a). The genes highlighted in yellow constitute the 20-gene signature.

For each entry, the following information is reported (from left to right) for this and subsequent sheets:

- Affy probeset (corresponding to the detecting probeset on Affymetrix HG-U133_Plus_2 GeneChips).
- Gene name (HUGO nomenclature when available).
- Average FC (fold-change) in "SC-like" *vs*. "Non-SC-like" BC comparison, across all patients.
- Statistical analysis of the difference in expression of each gene comparing "SC-like" *vs*. "Non-SC-like" BC groups, using the Benjamini correction (False Discovery Rate, FDR) to estimate the corrected p value.
- Description, Gene Name.

Sheet "20 pset Ivshina". List of the 20 probesets, with the highest differential expression, from the 329-signature used to perform a new hierarchical clustering analysis of the BC patient cohort of the Ivshina *et al*. dataset[9] (see Supplementary Fig. 1b).

Sheet "Ivshina HR". Univariate and multivariable analyses of the prognostic power of the 329-gene signature and 20-gene signature in the Ivshina *et al*. dataset.[9] Clinical endpoint: disease-free survival (DFS) (Supplementary Fig. 1, c and d).

Supplementary Table S2 (available as excels file). Analytical validation of the prognostic power of the 20-signature in independent genome-wide BC expression datasets.

Sheet "Pawitan Dataset". Clinicopathological and molecular characteristics of the BC patient cohort from the Pawitan *et al*. dataset. [10] For each patient, the "SC-like" status according to the 20-signature (column B) is also indicated.

Sheet "20 pset Pawitan". The 20 probesets, as above, were used to perform the unsupervised hierarchical clustering analysis of the BC patient cohort of the Pawitan *et al*. dataset,[10] shown in Fig. 2 (top, left panel).

Sheet "Pawitan HR". Univariate and multivariable analyses of the prognostic power of the 20-signature in the Pawitan *et al*. dataset.[10] Clinical endpoint: disease-free survival (DFS) (Fig. 2; top, right panel).

Sheet "Loi KI Dataset". Clinicopathological and molecular characteristics of the BC cohort from the Loi *et al*. dataset.[11] For each patient, the "SC-like" status according to the 20-signature (column D) is also indicated.

Sheet "20 pset Loi KI". The 20 probesets, as above, were used to perform the unsupervised clustering of the BC patient cohort of the Loi *et al*. dataset,[11] shown in Fig. 2 (middle, left panel).

Sheet "Loi KI HR". Univariate and multivariable analyses of the prognostic power of the 20-signature in the Loi *et al*. dataset.[11] Clinical endpoints: disease-free survival, (DFS); distant metastasis (DM) (Fig. 2; middle, right panel).

Sheet "Metabric Dataset". Clinicopathological and molecular characteristics of the BC cohort from the METABRIC dataset.[12] For each patient, the "SC-like" status according to the 20-signature (column B) is also indicated.

Sheet "Metabric HR". Univariate and multivariable analyses of the prognostic power of the 20-signature in the METABRIC cohort.[12] Clinical endpoints: death from breast cancer (Fig. 2; bottom, right panel).

Supplementary Table S3. Comparison of the clinicopathological variables between the initial consecutive cohort of BC patients from years 1997-2000 and the final "IEO BC 97-00" study cohort.

Variable	Stratum	Original patient selection from years 1997-2000 (N=3,396) N (%)	Study cohort ("IEO BC 97-00") (N=2,453) N (%)
Year of Surgery	1997	609 (17·9)	484 (19·7)
	1998	825 (24·3)	603 (24·6)
	1999	961 (28·3)	684 (27·9)
	2000	1001 (29·5)	682 (27·8)
Age at surgery	<35	150 (4·4)	90 (3·7)
	35-50	1331 (39·2)	957 (39·0)
	51-65	1389 (40·9)	1001 (40·8)
	>65	526 (15·5)	405 (16·5)
Histology	Ductal	2618 (77·1)	1960 (79·9)
	Lobular	353 (10·4)	195 (7·9)
	Mixed	88 (2·6)	69 (2·8)
	Other	337 (9·9)	229 (9·3)
pT	pT1	2234 (65·8)	1616 (65·9)
	pT2	1046 (30·8)	756 (30·8)
	pT3/ pT4	116 (3·4)	81 (3·3)
pN	pNX	98 (2·9)	59 (2·4)
	pN0	1748 (51·5)	1207 (49·2)
	pN1	997 (29·4)	764 (31·1)
	pN2	344 (10·1)	266 (10·8)
	pN3	209 (6·2)	157 (6·4)
Grade	n/a	114 (3·4)	57 (2·3)
	1	690 (20·3)	468 (19·1)
	2	1485 (43·7)	1069 (43·6)
	3	1107 (32·6)	859 (35·0)
LVI	Absent	2424 (71·4)	1681 (68·5)
	Focal	571 (16·8)	452 (18·4)
	Diffuse	401 (11·8)	320 (13·0)
ER/PgR	n/a	11 (0·3)	0 (0·0)
	Not expressed (Both 0)	525 (15·5)	341 (13·9)
	Expressed (ER>0 or PgR>0)	2860 (84·2)	2112 (86·1)
HER2 Status	n/a	1197 (35·2)	254 (10·4)
	NEG	1935 (57·0)	1935 (78·9)
	POS	264 (7·8)	264 (10·8)
Ki-67	Unknown	37 (1·1)	2 (0·1)
	<14%	1096 (32·3)	720 (29·4)
	≥14%	2263 (66·6)	1731 (70·6)
CT/HT	Nil	185 (5·4)	116 (4·7)
	HT	1251 (36·8)	866 (35·3)
	CT	514 (15·1)	345 (14·1)
	CT- HT	1446 (42·6)	1126 (45·9)
Surgery	n/a	2 (0·1)	0 (0·0)
	Quadrantectomy	2727 (80·3)	1996 (81·4)
	Mastectomy	667 (19·6)	457 (18·6)
Radiotherapy	No	654 (19·3)	450 (18·3)
	Yes	2742 (80·7)	2003 (81·7)
First Event	No event	2157 (63·5)	1513 (61·7)
	Loco-Regional	276 (8·1)	193 (7·9)
	Distant Metastasis	474 (14·0)	356 (14·5)
	Other	489 (14·4)	391 (15·9)

Legend to Supplementary Table S3. Comparison of the clinicopathological variables between the initial consecutive cohort of 3,396 BC patients from years 1997-2000 and the final "IEO BC 97-00" cohort of 2,453 patients (study cohort). The number (N) of patients and percentage (%) in each group is indicated. The event Distant Metastasis was defined as the occurrence of distant metastasis or death from BC as a first event (see Methods). Other events include second primary cancer or death from unknown causes or other causes. pT, primary tumour size; pN, nodal status; LVI, lymphovascular invasion; ER, oestrogen receptor; PgR, progesterone receptor; Ki-67, proliferation index. CT, adjuvant chemotherapy; HT, adjuvant hormone therapy; Nil, no adjuvant therapy; n/a, not available.

Supplementary Table S4. Design details for each TaqMan gene expression assay used in the PCR analysis.

Gene Symbol	Assay ID	Ref Seq	Exon Boundary	Assay Location	Amplicon Length	Primer and Probe sequences
APOBEC3B	custom	NM_001270411.1	7	1095-1151	57	Forward Primer: GGCTGCGGGCCATTC Reverse Primer: CTTAGAGACTGAGGCCCATCCTT Probe-FAM: CCAGAATCAGGGAAAC
RACGAP1	custom	NM_001320007.1	17 - 18	1511-1578	68	Forward Primer: TGTTACAGGACATCAAGCGTCAA Reverse Primer: CCAATACTCCAGAGGCAAGGAA Probe-FAM: CCAAGGTGGTTGAGCG
CENPW	custom	NM_001012507.3	2	664-724	61	Forward Primer: CAAACGCTTGTGCGAGTAAATG Reverse Primer: TTTGCTGCGGCCAGTACA Probe-FAM: AGAGTCATTAACAAGGAGC
H2AFZ	custom	NM_002106.3	1	501-559	59	Forward Primer: GCTGGTGGTGGTGTCATTCC Reverse Primer: TGTTGTCCTTTCTTCCCAATCA Probe-FAM: CACATCCACAAATCT
EXOSC4	custom	NM_019037.2	3 - 4	432-499	69	Forward Primer: GAAGCAGCCATCCTCACACA Reverse Primer: GCCTGTAGCACCTGCACATAGA Probe-FAM: ACCCACGCTCCCAGAT
NOL3	custom	NM_001276312.1	5	1428-1482	55	Forward Primer: GCCCACCACGAGCATCA Reverse Primer: CCTGGACTCCTAAGGGCAGAT Probe-FAM: CCAGTCCTCAGCCC
PHB	custom	NM_001281496.1	8	1176-1237	62	Forward Primer: TCCACCTCCCTACCAAAAATTG Reverse Primer: CCCGAATTGGGACCTAAAGC Probe-FAM: CAAGTGCCTATGCAAAC
H2AFJ	custom	NM_177925.3	1	2131-2190	60	Forward Primer: CAAAGGTCAGGCCGTACACA Reverse Primer: ACATCTCGAACCTGCCCAAT Probe-FAM: CTCTGTTAGGAGGCAAAT
SFN	custom	NM_006142.3	1	1115-1177	63	Forward Primer: TGCCTCTGATCGTAGGAATTGA Reverse Primer: CCTGCCACTGTCCAGTTCTCA Probe-FAM: TGTCCCGCCTTGTGG
CDK1	custom	NM_001786.4	2 - 3	164-239	76	Forward Primer: GAGAAAATTGGAGAAGGTACCTATGG Reverse Primer: TCATGGCTACCACTTGACCTGTA Probe-FAM: TGTATAAGGGTAGACACAAAA
EIF4EBP1	Hs00607050_m1	NM_004095.3	2 - 3	395	69	Probe-FAM: ATAAGCGGGCGGGCGGTGAAGAGTC
EPB41L5	Hs01554426_m1	NM_001184937.1	14 - 15	1375	67	Probe-FAM: AACTTAGTGTTCACAATAATGTTTC
LY6E	Hs03045111_g1	NM_001127213.1	3 - 4	329	66	Probe-FAM: GCCGGCATTGGGAATCTCGTGACAT

MIEN1	Hs00260553_m1	NM_032339.3	2 - 3	229	83	Probe-FAM: CGGGGGCACAGGTGCCTTTGAGATA
MMP1	Hs00899658_m1	NM_001145938.1	7 - 8	1019	64	Probe-FAM: AAGTCCGGTTTTTCAAAGGGAATAA
MRPS23	Hs00950118_g1	NM_016070.3	4 - 5	484	79	Probe-FAM: AAGCAAGGACTCAACACGGAGGTAG
NDUFB10	Hs01018233_g1	NM_004548.2	2 - 3	375	83	Probe-FAM: AGTGGAAGAGGGACTACAAAGTCGA
PHLDA2	Hs04194980_s1	NM_003311.3	1 - 1	254	75	Probe-FAM: GCGCACGGGCAAGTACGTGTACTTC
TOP2A	Hs01032142_g1	NM_001067.3	26 - 27	3611	96	Probe-FAM: TAAGAAATGAAAAAGAACAAGAGCT
ALYREF	Hs01099193_g1_	NM_005782.3	3 – 4	543	70	Probe-FAM: CGTCCCTCTGGATGGCCGCCCCATG
GAPDH	Hs_03929097_g1	NM_001256799.1	8 - 8	1250	58	Probe-FAM: CAAGAGGAAGAGAGAGACCCTCACT
HPRT1	Hs02800695_m1	NM_000194.2	2 – 3	297	82	Probe-FAM: GGACTAATTATGGACAGGACTGAAC
GUSB	Hs99999908_m1	NM_000181.3	11 - 12	1925	81	Probe-FAM: TGAACAGTCACCGACGAGAGTGCTG
TBP	Hs00427621_m1	NM_001172085.1	3 - 4	666	65	Probe-FAM: AATCCCAAGCGGTTTGCTGCGGTAA

Legend to Supplementary Table S4. Gene name (Gene Symbol) and identification number (Assay ID) of each TaqMan assay, accession number of the transcripts (Ref Seq) recognized by the assay, exon boundary, assay location and amplicon length are indicated. For TaqMan custom assays, locations of 5' nucleotide start and 3' nucleotide end of the entire amplicon and oligonucleotide sequences of forward and reverse primers, as well as FAM-probes are indicated. For proprietary designed TaqMan assays, locations corresponding to the nucleotide base located in the center of the probe and oligonucleotide context sequences of FAM-probes released by the vendor are reported.

Supplementary Table S5. Risk prediction algorithm based on mRNA expression levels of the 20 SC genes.

Gene Symbol	Value
H2AFZ	-0·01859310295660840
CDK1	-0·06976848796306750
EXOSC4	-0·00637353864393068
PHLDA2	-0·05504809927914070
APOBEC3B	0·02139449407835210
EIF4EBP1	-0·06561519832001840
SFN	0·00241415261227383
PHB	-0·01856010067320840
EPB41L5	-0·03434054196123320
RACGAP1	-0·04784034389453400
MRPS23	-0·09612536213842390
TOP2A	-0·06253072902806960
H2AFJ	-0·00972700284313969
NOL3	-0·02212862243820850
MIEN1	-0·01645816728337240
CENPW	0·01038668250241190
LY6E	-0·05456883437314120
ALYREF	-0·02403611555643560
MMP1	-0·03688016578089580
NDUFB10	0·00295444133576209
Scale factors	
Maximum	-14·9663172
Minimum	-17·5085913

Legend to Supplementary Table S5. RT-qPCR expression data for the 20 SC genes in the training set (N = 772) of the "IEO BC 97-00" cohort were used to derive a prognostic risk model after adjustment of regression coefficients of the respective genes by ridge penalized Cox regression model using 10-fold cross-validation. The regression coefficients obtained from the training set for each gene are indicated. Scale factors used to scale the risk score in a 0-100 range are also reported.

Supplementary Table S6. Distribution of clinicopathological characteristics and behaviour of the 20-signature continuous risk score in the patients of the IEO BC 97-00 cohort that were randomly assigned to the training and validation sets.

Variable	All N (% col)	Training N (% row)	Validation N (% row)	p value
All	2316 (100)	772 (33·3)	1544 (66·7)	
Age at surgery				0·22
<50	911 (39·3)	290 (31·8)	621 (68·2)	
≥50	1405 (60·7)	482 (34·3)	923 (65·7)	
Histology				0·56
Ductal	1859 (80·3)	625 (33·6)	1234 (66·4)	
No Ductal	457 (19·7)	147 (32·2)	310 (67·8)	
pT				0·16
pT1a/b	298 (12·9)	109 (36·6)	189 (63·4)	
pT1c	1195 (51·6)	374 (31·3)	821 (68·7)	
pT2	744 (32·1)	259 (34·8)	485 (65·2)	
pT3/pT4	79 (3·4)	30 (38·0)	49 (62·0)	
pN				0·57
pN0	1124 (48·5)	373 (33·2)	751 (66·8)	
pN+	1137 (49·1)	377 (33·2)	760 (66·8)	
pNX	55 (2·4)	22 (40·0)	33 (60·0)	
Grade				0·71
1	427 (18·4)	146 (34·2)	281 (65·8)	
2	1009 (43·6)	324 (32·1)	685 (67·9)	
3	826 (35·7)	282 (34·1)	544 (65·9)	
n/a	54 (2·3)	20 (37·0)	34 (63·0)	
LVI				0·95
Absent	1567 (67·7)	523 (33·4)	1044 (66·6)	
Present	749 (32·3)	249 (33·2)	500 (66·8)	
ER/PgR				0·67
Not expressed (Both 0)	313 (13·5)	101 (32·3)	212 (67·7)	
Expressed (ER>0 or PgR>0)	2003 (86·5)	671 (33·5)	1332 (66·5)	
HER2 Status				0·84
NEG	1826 (78·8)	605 (33·1)	1221 (66·9)	
POS	253 (10·9)	84 (33·2)	169 (66·8)	
n/a	237 (10·2)	83 (35·0)	154 (65·0)	
Ki-67				0·56
<14%	653 (28·2)	222 (34·0)	431 (66·0)	
≥14%	1661 (71·7)	550 (33·1)	1111 (66·9)	
n/a	2 (0·1)	0 (0·0)	2 (100)	
Subtype				0·95
Luminal	1827 (78·9)	614 (33·6)	1213 (66·4)	
HER2+	253 (10·9)	84 (33·2)	169 (66·8)	
Triple-Negative	212 (9·2)	69 (32·5)	143 (67·5)	
n/a	24 (1·0)	5 (20·8)	19 (79·2)	
CT/HT				0·45
Nil	111 (4·8)	37 (33·3)	74 (66·7)	
HT	806 (34·8)	285 (35·4)	521 (64·6)	
CT	320 (13·8)	99 (30·9)	221 (69·1)	
HT-CT	1079 (46·6)	351 (32·5)	728 (67·5)	
Surgery				0·087
Quadrantectomy	1870 (80·7)	608 (32·5)	1262 (67·5)	
Mastectomy	446 (19·3)	164 (36·8)	282 (63·2)	
Radiotherapy				0·084
No	437 (18·9)	161 (36·8)	276 (63·2)	
Yes	1879 (81·1)	611 (32·5)	1268 (67·5)	
		Mean (sd)	**Mean (sd)**	**p value** ***(t-test)***
20-signature continuous risk score		57·7 (14·5)	57·0 (14·5)	0·28

Legend to Supplementary Table S6. Patients from the "IEO BC 97-00" study cohort for whom risk score data were available (N = 2,316) were randomly assigned to the training (N = 772) or validation (N = 1,544) set. The association between group (training/validation set) and the demographic, clinical, and pathological variables was evaluated with the chi-square test. The association between group (training/validation set) and 20-signature continuous risk score was evaluated with the *t*-test. The number (N) of patients and percentage (%) in each group is indicated. pT, primary tumour size; pN, nodal status; LVI, lymphovascular invasion; ER, oestrogen receptor; PgR, progesterone receptor; Ki-67, proliferation index; CT, adjuvant chemotherapy; HT, adjuvant hormone therapy; Nil, no adjuvant therapy; n/a, not available; SD, standard deviation; POS, positive; NEG, negative.

Supplementary Table S7. Correlation between 20-gene SC signature and clinicopathological features in the "IEO BC 97-00" cohort.

Variables	N	Mean (SD)	Univariate analysis	
			Coef. (95% CI)	p value
All	1544	57·0 (14·5)		
Age at surgery				
<50	621	58·3 (14·7)	Ref.	
≥50	923	56·2 (14·4)	-2·1 (-3·6 to -0·6)	0·005
Histology				
Ductal	1234	59·2 (14·0)	10·7 (9·0–12·4)	<0·0001
No Ductal	310	48·5 (13·3)	Ref.	
pT				
pT1a/b	189	49·7 (13·4)	-14·1 (-16·4 to -11·8)	<0·0001
pT1c	821	54·3 (13·6)	-9·5 (-11·0 to -8·0)	<0·0001
pT2	485	63·8 (13·5)	Ref.	
pT3/pT4	49	63·4 (15·5)	-0·4 (-4·4 to 3·6)	0·85
pN				
pN0	751	55·0 (15·0)	Ref.	
pN+	760	59·1 (13·9)	4·1 (2·6–5·5)	<0·0001
pNX	33	53·9 (11·7)	-1·1 (-6·1 to 3·9)	0·67
Grade				
1	281	45·9 (11·2)	Ref.	
2	685	53·8 (12·5)	7·9 (6·2–9·6)	<0·0001
3	544	66·7 (12·7)	20·8 (19·1–22·6)	<0·0001
n/a	34	58·0 (12·2)	12·1 (7·7–16·5)	<0·0001
LVI				
Absent	1044	55·1 (14·4)	Ref.	
Present	500	61·1 (14·0)	6·0 (4·5–7·5)	<0·0001
ER/PgR				
Non expressed (Both 0)	212	64·7 (14·3)	8·9 (6·9–11·0)	<0·0001
Expressed (ER>0 or PgR>0)	1332	55·8 (14·2)	Ref.	
HER2 Status				
NEG	1221	56·1 (13·9)	Ref.	
POS	169	70·6 (12·0)	14·5 (12·4–16·7)	<0·0001
UNKNOWN	154	49·7 (13·4)	-6·3 (-8·6 to -4·1)	<0·0001
Ki-67				
<14%	431	46·0 (11·6)	Ref.	
≥14%	1111	61·3 (13·3)	15·3 (13·9–16·7)	<0·0001
n/a	2	56·4 (0·7)	10·4 (-7·4 to 28·2)	0·25
Subtype				
n/a	19	62·2 (14·0)		
Luminal	1213	54·5 (13·6)	Ref.	
HER2+	169	70·6 (12·0)	16·1 (13·9–18·3)	<0·0001
Triple-Negative	143	61·6 (15·0)	7·1 (4·8–9·5)	<0·0001

Legend to Supplementary Table S7. Correlation between the 20-gene SC signature and clinicopathological features was assessed using the linear regression model in the validation set of the "IEO BC 97-00" cohort (N=1,544). Coefficient (Coef.) of the linear regression model is the average difference in risk score between the comparison group and the reference group (Ref.). For instance, in the correlation of the 20-signature with grade of differentiation (Grade), the difference in the average risk score observed in Grade 3 tumour patients (comparison group) *vs.* Grade 1 tumour patients (reference group) is 20.8. N, Number of patients; SD, standard deviation; Coef., coefficient of linear regression model; CI, confidence interval; n/a, not available; pT, primary tumour size; pN, nodal status; ER, oestrogen receptor; PgR, progesterone receptor; Ki-67, proliferation index; LVI, lymphovascular invasion; POS, positive; NEG, negative; Ref., reference group.

Supplementary Table S8. Summary of the performance of the 20-signature risk model used as a continuous function (10-unit increase) in the prediction of the likelihood of distant metastasis in a stratified analysis of the different molecular subtypes of BC in the "IEO BC 97-00" validation set.

	Subgroups	N[$]	DM	Univariate		Multivariable	
				HR (95% CI)	p value	HR* (95% CI)	p value
Any time	All	1544	231	1·47 (1·35;1·61)	<0·0001	1·24 (1·11–1·38)	0·0002
	Luminal	1213	163	1·64 (1·47;1·83)	<0·0001	1·30 (1·13–1·48)	0·0002
	Triple-Negative	143	29	1·29 (1·00;1·68)	0·050	1·42 (1·05–1·92)	0·023
	HER2+	169	36	1·01 (0·78;1·32)	0·92	1·00 (0·74–1·36)	0·99
Early metastasis	All	1544	140	1·59 (1·43;1·78)	<0·0001	1·25 (1·08–1·44)	0·002
	Luminal	1213	87	1·68 (1·45;1·94)	<0·0001	1·30 (1·08–1·56)	0·004
	Triple- Negative	143	24	1·50 (1·13;2·01)	0·006	1·38 (0·99–1·90)	0·054
	HER2+	169	29	1·10 (0·82;1·47)	0·54	1·08 (0·77–1·50)	0·67
Late metastasis	All	1236	67	1·38 (1·18;1·61)	<0·0001	1·27 (1·03–1·57)	0·027
	Luminal	998	60	1·68 (1·41;2·00)	<0·0001	1·37 (1·09–1·71)	0·006
	Triple- Negative	102	1	0·64 (0·24;1·70)	0·37	NE	NE
	HER2+	118	5	0·61 (0·30;1·23)	0·17	0·42 (0·11–1·63)	0·20

Legend to Supplementary Table S8. Univariate and multivariable analysis in all patients of the "IEO BC 97-00" validation set and in patients divided by molecular subtype, for the early (<5 years from surgery) and late (5-10 years post-surgery) time intervals, were performed. The hazard ratios were estimated with a Cox proportional hazards univariate model (HR) and multivariable model (HR*), adjusted for Grade (G1, G2 and G3), Ki-67 (Ki-67<14% and Ki-67≥14%), HER2 status (positive and negative), oestrogen/progesterone receptor status [not expressed (Both 0) and expressed (ER > 0 or PgR > 0)], tumour size (pT1 and pT2-3-4), number of positive lymph nodes (pN0, pN1-2-3 and pNX) and age at surgery (<50 and ≥50). N, number of patients; DM, number of distant metastasis events; HR, hazard ratio; CI, confidence interval; NE, Not estimable. [$]Nineteen patients could not be assigned to the subgroups.

Supplementary Table S9. Correlation between clinicopathological parameters and CSC proliferation kinetics (self-limiting *vs.* unlimited expansion) in a 90-BC patient cohort.

Variable	All N (% col)	Self-limiting expansion N (% row)	Unlimited expansion N (% row)	p value
All	90 (100)	40 (44·4)	50 (55·6)	
Age at surgery				0·10
<50	62 (68·9)	24 (38·7)	38 (61·3)	
≥50	28 (31·1)	16 (57·1)	12 (42·9)	
Histology				0·002
Ductal	75 (83·3)	28 (37·3)	47 (62·7)	
No Ductal	15 (16·7)	12 (80·0)	3 (20·0)	
pT				0·29
pT1a/b	4 (4·4)	3 (75·0)	1 (25·0)	
pT1c	26 (28·9)	13 (50·0)	13 (50·0)	
pT2	48 (53·3)	21 (43·7)	27 (56·2)	
pT3/pT4	12 (13·3)	3 (25·0)	9 (75·0)	
Positive lymph nodes				0·87
pN0	24 (26·7)	11 (45·8)	13 (54·2)	
pN+	66 (73·3)	29 (43·9)	37 (56·1)	
Grade				<0·001
1	13 (14·4)	12 (92·3)	1 (7·7)	
2	40 (44·4)	18 (45·0)	22 (55·0)	
3	37 (41·1)	10 (27·0)	27 (73·0)	
LVI				0·015
Absent	39 (43·3)	23 (59·0)	16 (41·0)	
Present	51 (56·7)	17 (33·3)	34 (66·7)	
ER/PgR				0·56
Not expressed (Both 0)	11 (12·2)	4 (36·4)	7 (63·6)	
Expressed (ER>0 or PgR>0)	79 (87·8)	36 (45·6)	43 (54·4)	
HER2 status				0·25
Negative	82 (91·1)	38 (46·3)	44 (53·7)	
Positive	8 (8·9)	2 (25·0)	6 (75·0)	
Ki-67				<0·001
<14%	17 (18·9)	14 (82·4)	3 (17·6)	
≥14%	73 (81·1)	26 (35·6)	47 (64·4)	
Subtype				0·36
Luminal	73 (81·1)	35 (47·9)	38 (52·1)	
HER2+	8 (8·9)	2 (25·0)	6 (75·0)	
Triple-Negative	9 (10·0)	3 (33·3)	6 (66·7)	

Legend to Supplementary Table S9. The association between the self-limiting or unlimited proliferative phenotype of CSCs and the demographic, clinical and pathological variables, was evaluated using the chi-square test. The number (N) of patients and percentage (%) in each group is indicated. pT, primary tumour size; pN, nodal status; LVI, lymphovascular invasion; Ki-67, proliferation index; p, p -value.

Supplementary Table S10. Clinicopathological characteristics of the ER+/HER2- patients of the "IEO BC 97-00" cohort randomly assigned to the training and validation set.

Variable	All N (% col)	Training N (% row)	Validation N (% row)	p value
All	1827 (100)	609 (33·3)	1218 (66·7)	
Age at surgery				0·51
<50	670 (36·7)	217 (32·4)	453 (67·6)	
≥50	1157 (63·3)	392 (33·9)	765 (66·1)	
Histology				0·91
Ductal	1407 (77·0)	470 (33·4)	937 (66·6)	
No Ductal	420 (23·0)	139 (33·1)	281 (66·9)	
pT				0·33
pT1a/b	266 (14·6)	97 (36·5)	169 (63·5)	
pT1c	989 (54·1)	312 (31·5)	677 (68·5)	
pT2	517 (28·3)	182 (35·2)	335 (64·8)	
pT3/pT4	55 (3·0)	18 (32·7)	37 (67·3)	
pN				0·83
pN0	910 (49·8)	303 (33·3)	607 (66·7)	
pN+	866 (47·4)	287 (33·1)	579 (66·9)	
pNX	51 (2·8)	19 (37·3)	32 (62·7)	
Grade				0·57
1	418 (22·9)	140 (33·5)	278 (66·5)	
2	910 (49·8)	291 (32·0)	619 (68)	
3	453 (24·8)	161 (35·5)	292 (64·5)	
n/a	46 (2·5)	17 (37·0)	29 (63·0)	
LVI				0·88
Absent	1276 (69·8)	424 (33·2)	852 (66·8)	
Present	551 (30·2)	185 (33·6)	366 (66·4)	
Ki-67				0·72
<14%	627 (34·3)	213 (34·0)	414 (66·0)	
≥14%	1199 (65·6)	396 (33·0)	803 (67·0)	
n/a	1 (0·1)	0 (0·0)	1 (100·0)	
CT/HT				0·64
Nil	81 (4·4)	26 (32·1)	55 (67·9)	
HT	786 (43·0)	272 (34·6)	514 (65·4)	
CT	55 (3·0)	15 (27·3)	40 (72·7)	
HT-CT	905 (49·5)	296 (32·7)	609 (67·3)	
Surgery				0·29
Quadrantectomy	1524 (83·4)	500 (32·8)	1024 (67·2)	
Mastectomy	303 (16·6)	109 (36·0)	194 (64·0)	
Radiotherapy				0·27
No	314 (17·2)	113 (36·0)	201 (64·0)	
Yes	1513 (82·8)	496 (32·8)	1017 (67·2)	

Legend to Supplementary Table S10. Comparison of the clinicopathological characteristics of ER+/HER2- patients of the "IEO BC 97-00" cohort that were randomly assigned to the training and validation sets. Risk score data were available for 1827 ER+/HER2- patients. The association between group (training or validation set) and the demographic, clinical, and pathological variables was evaluated using the chi-square test. The number (N) of patients and percentage (%) in each group is indicated. pT, primary tumour size; pN, nodal status; LVI, lymphovascular invasion; ER, oestrogen receptor; PgR, progesterone receptor; Ki-67, proliferation index; CT, adjuvant chemotherapy; HT, adjuvant hormone therapy; Nil, no adjuvant therapy; n/a, not available.

Supplementary Table S11. Development of the StemPrintER risk model.

Gene Symbol	Value
H2AFZ	-0·03833591325196550
CDK1	-0·06132455806571770
EXOSC4	-0·02105976326055420
PHLDA2	-0·06295739658169650
APOBEC3B	0·02341881674020150
EIF4EBP1	-0·13911217901125500
SFN	0·05788269046891110
PHB	-0·03538557745953510
EPB41L5	-0·04675539403890050
RACGAP1	-0·05097505893853430
MRPS23	-0·14201022110072700
TOP2A	-0·11290078348786600
H2AFJ	-0·04975471358452700
NOL3	-0·04193802459521500
MIEN1	0·01133668644106850
CENPW	-0·03717918353187610
LY6E	-0·02829256296234230
ALYREF	-0·09541915699494330
MMP1	-0·00911370427072023
NDUFB10	0·00626166874136819
2-class cut-off	
Median	56·31840823
Scale factors	
Maximum	-21·7767727
Minimum	-25·2349961

Legend to Supplementary Table S11. RT-qPCR expression data for the 20 SC genes in the ER+/HER2- BC subgroup were used to derive a prognostic risk model after adjustment of regression coefficients of the respective genes by the ridge penalized Cox regression model using a 10-fold cross-validation. The regression coefficients obtained from the training set for each gene are indicated. Scale factors used to scale the risk score in a 0 – 100 range and cut-offs used to categorize patients into 2 classes (low, high) of risk are also reported.

Supplementary Table S12. Stratification of the training and validation set of ER+/HER2- patients according to StemPrintER used as a continuous risk score or as a 2-class risk model.

Patient group	All N (% col)	Training N (% row)	Validation N (% row)	p value
All	1827 (100)	609 (33·3)	1218 (66·7)	
StemPrintER 2-class model				0·26
Low	949 (51·9)	305 (32·1)	644 (67·9)	
High	878 (48·1)	304 (34·6)	574 (65·4)	
Variable		**Mean (SD)**	**Mean (SD)**	**p value *(t-test)***
StemPrintER continuous ris score		57·2 (14·5)	56·5 (14·5)	0·36

Legend to Supplementary Table S12. Risk score data were available for 1,827 ER+/HER2- patients from the "IEO BC 97-00" cohort. StemPrintER was used as a 2-class risk model or as a continuous function. The association between group (training/validation set) and StemPrintER 2-class categorization was evaluated with the chi-square test. For the analysis of StemPrintER used as a continuous function, the difference in the mean risk score between the training and validation set was evaluated with the *t*-test. The number (N) of patients and percentage (%) in each group is indicated. SD, standard deviation; Int., intermediate.

Supplementary Table S13. Correlation between the StemPrintER continuous risk score and clinicopathological parameters in a univariate analysis of the ER+/HER2- validation set.

			Univariate analysis	
Variable	**N**	**Mean (SD)**	**Coef· (95% CI)**	**p value**
All	1218	56·5 (14·5)		
Age at surgery				
<50	453	57·0 (14·9)	Ref.	
≥50	765	56·2 (14·2)	-0·8 (-2·5 to 0·9)	0·36
Histology				
Ductal	937	58·4 (14·4)	8·0 (6·1–9·9)	<0·0001
No Ductal	281	50·4 (13·1)	Ref.	
pT				
pT1a/b	169	50·4 (12·9)	-14·2 (-16·7 to -11·7)	<0·0001
pT1c	677	53·8 (13·3)	-10·7 (-12·5 to -8·9)	<0·0001
pT2	335	64·5 (14·3)	Ref.	
pT3/pT4	37	61·7 (13·4)	-2·9 (-7·4 to 1·7)	0·22
pN				
pN0	607	54·3 (14·5)	Ref.	
pN+	579	59·0 (14·2)	4·7 (3·1–6·3)	<0·0001
pNX	32	54·3 (12·9)	0·0 (-5·1 to 5·1)	0·99
Grade				
1	278	47·4 (11·0)	Ref.	
2	619	55·0 (12·4)	7·7 (5·9–9·4)	<0·0001
3	292	68·1 (14·2)	20·7 (18·6–22·8)	<0·0001
n/a	29	59·2 (13·1)	11·8 (7·0–16·6)	<0·0001
LVI				
Absent	852	54·6 (14·2)	Ref.	
Present	366	60·9 (14·2)	6·3 (4·5–8·0)	<0·0001
Ki-67				
<14%	414	47·7 (11·3)	Ref.	
≥14%	803	61·1 (13·9)	13·4 (11·9–15·0)	<0·0001
n/a	1	54·5 (n/a)	6·8 (-18·7 to 32·4)	0·60

Legend to Supplementary Table S13. Correlation between the StemPrintER continuous risk score and clinicopathological features was assessed using the univariate linear regression model in the ER+/HER2- validation set (N = 1,218). Coefficient (Coef.) of the linear regression model is the average difference in risk score between comparison group and the reference group (Ref.). For instance, in the correlation of StemPrintER risk score and grade of differentiation (Grade), the difference in the average risk score observed in Grade 3 tumour patients (comparison group) *vs*. Grade 1 tumour patients (reference group) is 20.7. N, Number of patients; SD, Standard deviation; Coef., coefficient of linear regression model; CI, confidence interval; n/a, not available; pT, primary tumour size; pN, nodal status; Ki-67, proliferation index; LVI, lymphovascular invasion; Ref., reference group.

Supplementary Table S14. Summary of the performance of the StemPrintER risk model, used as a continuous (10-unit increase) function or according to a 2-class categorization, in predicting risk of recurrence in the ER+/HER2- training and validation sets.

		Univariate N=609		Multivariable N=609	
Training Set	**Risk Model**	**HR (95% CI)**	**p value**	**HR* (95% CI)**	**p value**
	Continuous risk score	1·59 (1·38–1·83)	<0·0001	1·35 (1·11–1·63)	0·002
	2-Class: High *vs* Low	4·25 (2·56–7·06)	<0·0001	2·38 (1·30–4·35)	0·005
		Univariate N=1218		**Multivariable N=1218**	
Validation Set	**Risk Model**	**HR (95% CI)**	**p value**	**HR* (95% CI)**	**p value**
	Continuous risk score	1·55 (1·41–1·71)	<0·0001	1·23 (1·09–1·40)	0·0009
	2-Class: High *vs* Low	3·37 (2·40–4·74)	<0·0001	1·85 (1·25–2·74)	0·002

Legend to Supplementary Table S14. StemPrintER was used as a continuous function or as a 2-class risk model to predict risk of distant metastasis in the ER+/HER2- BC training and validation sets. Table reports univariate and multivariable analyses for the entire follow-up interval. The hazard ratios were estimated with a Cox proportional hazards univariate (HR) and multivariable (HR*) model, adjusted for Grade (G1, G2 and G3), Ki-67 (Ki-67 < 14% and Ki-67 ≥ 14%), tumour size (pT1 and pT2-3-4), number of positive lymph nodes (pN0, pN1-2-3 and pNX) and age at surgery (<50 and ≥50). N, number of patients; HR, hazard ratio; CI, confidence interval; Int., Intermediate risk category.

Supplementary Table S15. Multivariable Cox proportional analysis of risk of early (0-5 years) or late (5-10 years) distant metastasis in the ER+/HER2- validation set according to the StemPrintER continuous risk model.

Variable	Early distant metastasis recurrence (0-5 years)			Late distant metastasis recurrence (5-10 years)		
	HR	95% CI	p value	HR	95% CI	p value
StemPrintER 10-unit increase	1·23	1·04–1·46	0·016	1·32	1·07–1·62	0·008
Age ≥50 *vs* <50	1·42	0·91–2·24	0·13	0·96	0·57–1·60	0·86
pT2-3-4 *vs* pT1	2·88	1·77–4·68	<·0001	1·84	1·06–3·21	0·032
pN+ *vs* pN0	1·91	1·17–3·13	0·010	4·13	2·11–8·08	<·0001
pNX *vs* pN0	1·81	0·42–7·74	0·42	NE	NE	NE
Gn/a *vs* G1	1·05	0·12–9·15	0·97	1·41	0·28–7·14	0·68
G2 *vs* G1	2·43	0·92–6·46	0·074	0·92	0·38–2·24	0·85
G3 *vs* G1	4·47	1·55–12·90	0·005	1·00	0·37–2·77	0·99
Ki-67≥14% *vs* <14%	0·80	0·43–1·49	0·47	1·95	0·86–4·42	0·108

Legend to Supplementary Table S15. StemPrintER was used as a continuous risk model to estimate the likelihood of developing distant metastasis in the early (0-5 years) and late (5-10 years) time interval after surgery in the validation cohort of 1,218 ER+/HER- patients. Hazard ratios for distant metastasis were estimated with a Cox proportional hazards multivariable model adjusted for Grade (G1, G2 and G3), Ki-67 (Ki-67<14% and Ki-67≥14%), tumour size (pT1 and pT2-3-4), number of positive lymph nodes (pN0, pN1-2-3 and pNX) and age at surgery (<50 and ≥50). HR, hazard ratio; CI, confidence interval; NE, not estimable; n/a, not available.

Supplementary Table S16. Summary of the predictive power of the StemPrintER risk model used as a continuous (10-unit increase) function to estimate risk of recurrence in all patients and in different subgroups of ER+/HER2- patients of the validation set.

Time	Subgroup	Univariate analysis		Multivariable analysis	
		HR (95% CI)	p value	HR* (95% CI)	p value
Early metastasis (0-5 years)	**All**	1·60 (1·40–1·82)	<0·0001	1·23 (1·04–1·46)	0·0158
	Premenopausal	1·58 (1·29–1·93)	<0·0001	1·32 (1·01–1·71)	0·0387
	Postmenopausal	1·64 (1·37–1·97)	<0·0001	1·21 (0·96–1·52)	0·11
	Ductal	1·56 (1·35–1·80)	<0·0001	1·20 (1·00–1·45)	0·0464
	No Ductal	1·64 (1·10–2·44)	0·0141	1·32 (0·80–2·19)	0·28
	pT1	1·52 (1·18–1·96)	0·0014	1·40 (1·02–1·92)	0·0363
	pT2/3/4	1·33 (1·12–1·58)	0·001	1·17 (0·96–1·44)	0·13
	pN0	2·02 (1·55–2·63)	<0·0001	1·71 (1·22–2·39)	0·002
	pN+	1·40 (1·19–1·64)	<0·0001	1·12 (0·91–1·37)	0·29
	G1	1·30 (0·62–2·72)	0·49	0·86 (0·34–2·15)	0·75
	G2	1·48 (1·14–1·91)	0·0027	1·43 (1·06–1·91)	0·0181
	G3	1·34 (1·09–1·64)	0·0052	1·19 (0·96–1·49)	0·11
	LVI Absent	1·53 (1·28–1·83)	<0·0001	1·21 (0·97–1·51)	0·0991
	LVI Present	1·60 (1·30–1·97)	<0·0001	1·26 (0·96–1·65)	0·0977
	Ki-67<14%	1·80 (1·28–2·53)	0·0008	1·48 (0·95–2·31)	0·0839
	Ki-67≥14%	1·53 (1·30–1·79)	<0·0001	1·20 (1·00–1·44)	0·0562
Late metastasis (5-10 years)	**All**	1·60 (1·37–1·88)	<0·0001	1·32 (1·07–1·62)	0·0084
	Premenopausal	1·72 (1·40–2·10)	<0·0001	1·48 (1·13–1·95)	0·0045
	Postmenopausal	1·44 (1·12–1·85)	0·0046	1·15 (0·83–1·59)	0·41
	Ductal	1·59 (1·33–1·89)	<0·0001	1·38 (1·10–1·73)	0·0052
	No Ductal	1·67 (1·10–2·54)	0·016	1·02 (0·58–1·78)	0·95
	pT1	1·77 (1·36–2·32)	<0·0001	1·81 (1·30–2·53)	0·0004
	pT2/3/4	1·26 (1·01–1·57)	0·041	1·04 (0·78–1·38)	0·79
	pN0	1·64 (1·12–2·41)	0·0115	1·49 (0·92–2·40)	0·1
	pN+	1·48 (1·24–1·77)	<0·0001	1·30 (1·03–1·65)	0·025
	G1	1·10 (0·57–2·11)	0·78	0·79 (0·37–1·70)	0·55
	G2	1·65 (1·23–2·20)	0·0008	1·60 (1·13–2·26)	0·0086
	G3	1·43 (1·09–1·87)	0·0101	1·26 (0·95–1·68)	0·11
	LVI Absent	1·56 (1·27–1·93)	<0·0001	1·29 (1·00–1·66)	0·0532
	LVI Present	1·60 (1·23–2·08)	0·0004	1·29 (0·90–1·87)	0·17
	Ki-67<14%	1·09 (0·61–1·97)	0·77	0·98 (0·52–1·85)	0·96
	Ki-67≥14%	1·58 (1·31–1·90)	<0·0001	1·37 (1·10–1·70)	0·0046

Legend to Supplementary Table S16. The StemPrintER risk model was used as a continuous variable (10-unit increase) in a univariate and multivariable analysis to predict risk of early (0-5 years) and late (5-10 years) metastasis in the validation cohort of ER+/HER2- patients stratified by clinicopathological characteristics. Hazard ratios were estimated at each 10-unit increase with a Cox proportional hazards univariate (HR) and multivariable model (HR*), adjusted for Grade (G1, G2 and G3), Ki-67 (Ki-67<14% and Ki-67≥14%), tumour size (pT1 and pT2-3-4), number of positive lymph nodes (pN0, pN1-2-3 and pNX), and age at surgery (<50 and ≥50), as appropriate in each subgroup analysis. LVI, lymphovascular invasion. N, number of patients; HR, hazard ratio; CI, confidence interval.

Supplementary Table S17. Summary of the predictive power of the 2-class StemPrintER risk model in all patients and in different subgroups of ER+/HER2- patients of the validation set.

Time	Subgroup	Univariate analysis		Multivariable analysis	
		HR (95% CI)	p value	HR* (95% CI)	p value
Early metastasis (0-5 years)	**All**	4.68 (2.79–7.87)	<0.0001	2.48 (1.38–4.45)	0.0024
	Premenopausal	5.24 (2.14–12.81)	0.0003	3.12 (1.12–8.70)	0.0292
	Postmenopausal	4.33 (2.29–8.19)	<0.0001	2.19 (1.07–4.47)	0.0315
	Ductal	4.67 (2.57–8.49)	<0.0001	2.42 (1.24–4.73)	0.0094
	No Ductal	3.69 (1.17–11.62)	0.0258	2.46 (0.64–9.51)	0.19
	pT1	2.84 (1.34–6.00)	0.0065	2.34 (0.99–5.54)	0.0532
	pT2/3/4	3.70 (1.68–8.16)	0.0012	2.86 (1.23–6.66)	0.0148
	pN0	7.3 (2.48–21.47)	0.0003	4.09 (1.19–14.02)	0.025
	pN+	3.57 (1.94–6.59)	<0.0001	2.17 (1.10–4.27)	0.0247
	G1	2.49 (0.42–14.9)	0.32	1.13 (0.15–8.73)	0.91
	G2	2.78 (1.39–5.57)	0.0038	2.32 (1.10–4.91)	0.027
	G3	6.22 (1.51–25.67)	0.0115	4.39 (1.05–18.37)	0.0429
	LVI Absent	4.19 (2.17–8.10)	<0.0001	2.61 (1.24–5.49)	0.0115
	LVI Present	4.49 (1.89–10.69)	0.0007	2.24 (0.86–5.84)	0.1
	Ki-67<14%	3.15 (1.20–8.28)	0.0199	1.59 (0.50–5.07)	0.43
	Ki-67≥14%	5.22 (2.50–10.90)	<0.0001	3.05 (1.43–6.53)	0.004
Late metastasis (5-10 years)	**All**	3.35 (1.93–5.82)	<0.0001	1.90 (1.01–3.56)	0.047
	Premenopausal	3.51 (1.61–7.65)	0.0016	1.83 (0.72–4.64)	0.2
	Postmenopausal	3.21 (1.47–7.00)	0.0035	1.89 (0.79–4.50)	0.15
	Ductal	2.89 (1.56–5.38)	0.0008	1.9 (0.93–3.89)	0.0805
	No Ductal	5.17 (1.51–17.65)	0.0088	1.73 (0.43–6.88)	0.44
	pT1	3.30 (1.51–7.20)	0.0028	2.84 (1.20–6.71)	0.0176
	pT2/3/4	1.88 (0.85–4.18)	0.12	1.17 (0.47–2.94)	0.73
	pN0	2.05 (0.62–6.71)	0.24	1.14 (0.30–4.29)	0.85
	pN+	3.16 (1.67–5.96)	0.0004	2.16 (1.04–4.45)	0.0378
	G1	0.67 (0.08–5.57)	0.71	0.39 (0.04–3.52)	0.4
	G2	3.16 (1.42–7.04)	0.0048	2.45 (1.05–5.73)	0.0387
	G3	2.31 (0.69–7.73)	0.18	2.12 (0.48–9.46)	0.32
	LVI Absent	2.82 (1.38–5.77)	0.0045	1.73 (0.77–3.86)	0.18
	LVI Present	3.45 (1.40–8.51)	0.0072	1.93 (0.66–5.64)	0.23
	Ki-67<14%	1.42 (0.29–6.83)	0.66	1.18 (0.23–6.18)	0.84
	Ki-67≥14%	3.17 (1.59–6.34)	0.0011	2.19 (1.05–4.55)	0.0355

Legend to Supplementary Table S17. StemPrintER was used as a 2-class risk model in a univariate and multivariable analysis of risk of distant metastasis in the early (0-5 years) and late (5-10 years) time intervals. The hazard ratios (High *vs.* Low) were estimated with a Cox proportional hazards univariate (HR) and multivariable model (HR*), adjusted for Grade (G1, G2 and G3), Ki-67 (Ki-67<14% and Ki-67≥14%), tumour size (pT1 and pT2-3-4), number of positive lymph nodes (pN0, pN1-2-3 and pNX), and age at surgery (<50 and ≥50), as appropriate in each subgroup analysis. LVI, lymphovascular invasion. N, number of patients; HR, hazard ratio; CI, confidence interval.

Supplementary Table S18. Cumulative incidence of early and late distant metastasis events in all patients and in different ER+/HER2- patient subgroups of the training and validation set estimated using the 2-class StemPrintER risk model.

	Subgroup	Risk Class	<5-year CI-DM (95% CI)	5-10 year CI-DM (95% CI)
Training Set	All	High	13.7 (10.1–17.9)	9.7 (6.3–14)
		Low	2.0 (0.8–4.1)	2.6 (1.2–5.1)
Validation Set	All	High	12.3 (9.7–15.2)	10.1 (7.4–13.3)
		Low	2.8 (1.7v4.4)	3.2 (2.0–5.0)
	Premenopausal	High	10.4 (6.9–14.7)	11.5 (7.4–16.7)
		Low	2.1 (0.9–4.4)	3.5 (1.7–6.3)
	Postmenopausal	High	13.7 (10.2–17.6)	9.0 (5.8–13.1)
		Low	3.4 (1.9–5.7)	3.0 (1.5–5.4)
	Ductal	High	12.9 (10.1–16)	10 (7.1–13.4)
		Low	3.0 (1.7–4.9)	3.7 (2.1–5.9)
	No Ductal	High	8.9 (3.9–16.4)	10.9 (4.8–20)
		Low	2.5 (0.9–5.4)	2.3 (0.7–5.4)
	pT1	High	5.8 (3.6–8.8)	6.7 (4.1–10.3)
		Low	2.1 (1.1–3.6)	2.2 (1.1–3.8)
	pT2/3/4	High	20.2 (15.5–25.4)	15.3 (10.3–21.2)
		Low	6.2 (2.7–11.8)	8.5 (4.0–15.3)
	pN0	High	7.8 (4.9–11.7)	3.1 (1.3–6.2)
		Low	1.1 (0.4–2.7)	1.5 (0.6–3.4)
	pN+	High	16.1 (12.2–20.4)	16.9 (12.2–22.2)
		Low	5.0 (2.8–8.1)	5.8 (3.3–9.4)
	G1	High	3.5 (0.6–10.8)	2.0 (0.2–9.5)
		Low	1.4 (0.4–3.7)	3.1 (1.3–6.2)
	G2	High	9.1 (6–12.9)	8.8 (5.4–13.3)
		Low	3.5 (1.9–5.8)	3 (1.5–5.5)
	G3	High	19.2 (14.3–24.6)	14.1 (9.1–20.2)
		Low	3.5 (0.6–10.8)	6.0 (1.5–15.1)
	LVI Absent	High	9.8 (7.0–13.2)	7.5 (4.7–11)
		Low	2.5 (1.3–4.1)	2.8 (1.5–4.6)
	LVI Present	High	16.3 (11.7–21.6)	14.9 (9.7–21.2)
		Low	4.1 (1.7–8.3)	4.9 (2.0–9.7)
	Ki−67<14%	High	9.0 (4.0–16.7)	3.2 (0.6–10)
		Low	3.0 (1.5–5.2)	2.4 (1.1–4.6)
	Ki−67≥14%	High	12.8 (10–16)	11.3 (8.3–14.9)
		Low	2.7 (1.3–5.0)	3.8 (2.0–6.7)

Legend to Supplementary Table S18. The cumulative incidence of distant metastasis (CI-DM) was estimated in the early (0-5 years) and late (5-10 years) time interval post-surgery in all patients and in different patient subgroups of the ER+/HER2- training and validation sets using the StemPrintER 2-class risk model. pT, primary tumour size; pN, nodal status; LVI, lymphovascular invasion; CI, confidence interval.