ARTICLE   OPEN

pAKT pathway activation is associated with *PIK3CA* mutations and good prognosis in luminal breast cancer in contrast to p-mTOR pathway activation

Amir Sonnenblick [1], David Venet[2], Sylvain Brohée[2], Noam Pondé[2] and Christos Sotiriou[2]

Numerous studies have focused on the PI3K/AKT/mTOR pathway in estrogen receptor positive (ER) breast cancer (BC), as a linear signal transduction pathway and reported its association with worse clinical outcomes. We developed gene signatures that reflect the level of expression of phosphorylated-Serine473-AKT (pAKT) and phosphorylated-Serine2448-mTOR (p-mTOR) separately, capturing their corresponding level of pathway activation. Our analysis revealed that the pAKT pathway activation was associated with luminal A BC while the p-mTOR pathway activation was more associated with luminal B BC (Kruskal–Wallis test $p < 10^{-10}$). pAKT pathway activation was significantly associated with better outcomes (multivariable HR, 0.79; 95%CI, 0.74–0.85; $p = 2.5 \times 10^{-10}$) and *PIK3CA* mutations ($p = 0.0001$) whereas p-mTOR pathway activation showed worse outcomes (multivariable HR,1.1; 95% CI, 1.1–1.2; $p = 9.9 \times 10^{-4}$) and associated with *p53* mutations ($p = 0.04$). in conclusion, our data show that pAKT and p-mTOR pathway activation have differing impact on prognosis and suggest that they are not linearly connected in luminal breast cancers.

INTRODUCTION

The phosphatidylinositol 3-kinase (PI3K)/AKT/mTOR-signaling pathway mediates key cellular functions, including growth, proliferation, and survival and is frequently involved in carcinogenesis, tumor progression, and metastases.[1] Numerous studies have focused on the PI3K/AKT/mTOR pathway in estrogen receptor positive (ER-positive) breast cancer (BC) and have shown that *PIK3CA* mutations are frequent, that the PI3K/AKT/mTOR-signaling pathway is often dysregulated and that both correlate with worse clinical outcomes.[2–4] As a consequence, a large number of drugs targeting the various components of this pathway have been developed.[5] Everolimus (an mTOR inhibitor) is currently the only approved drug targeting mTOR based on the results of the BOLERO-2 trial.[6]

While AKT is activated by phospholipid binding and activation loop phosphorylation at Threonine308 by PDK1 and by phosphorylation within the carboxy terminus at Serine473, mTOR is phosphorylated at Serine2448 via the PI3K-signaling pathway.[7] AKT activates the mTOR complex 1 (mTORC1) which in addition to mTOR contains mLST8, PRAS40, and RAPTOR.[8] This activation involves phosphorylation of tuberous sclerosis complex 2 (TSC2), which blocks the ability of TSC2 to act as a GTPase-activating protein, thereby allowing accumulation of Rheb-GTP and mTORC1 activation. AKT can also activate mTORC1 by PRAS40 phosphorylation, thereby relieving the PRAS40-mediated inhibition of mTORC1.[9]

The PI3K/AKT/mTOR pathway is usually considered as a linear signal transduction pathway in BC, however in the ER-positive disease, we have previously shown that *PIK3CA* mutations were associated with relatively low mTORC1 functional output and with good outcomes in patients who received adjuvant tamoxifen monotherapy.[4] Therefore, to gain better insight into the relative contribution of each of the signaling pathways which lie downstream to PI3K (namely AKT and mTOR) to BC outcomes, we have developed a novel in silico approach which assessed the activation of each of these signaling pathways separately, by integrating reverse phase protein array (RPPA) and matched gene expression.

RESULTS

pAKT pathway activated and p-mTOR pathway activated ER-positive early BCs are associated with distinct and exclusive gene expression profiles

We first derived two distinct signatures whose expression levels could predict AKT and mTOR pathway activation through pAKT and p-mTOR RPPA levels by computing the differentially expressed genes between tumor samples with high and low RPPA levels of pAKT (respectively, activated and inactivated AKT pathway) and p-mTOR proteins (respectively, activated and inactivated mTOR pathway), using ER-positive tumors from the TCGA repository. It is important to note that the two signatures did not share any common genes (Fig. 1a). We next sought to assess their biological and clinical relevance in BC. Firstly, we compared both signatures to the reference classes of the Gene Ontology and the mSigDB signatures repositories using the Broad Institute site.[10] This showed that the pAKT signature was significantly enriched in genes up-regulated in less aggressive invasive BC tumors (e.g. grade 1 vs. grade 3[11]; fdr $= 2 \times 10^{-27}$). In contrast, the p-mTOR signature was enriched in genes expressed in mammary stem cells and more aggressive luminal B cancers[11,12] (fdr $= 2 \times 10^{-7}$, fdr $= 3 \times 10^{-5}$, respectively). A network clustering
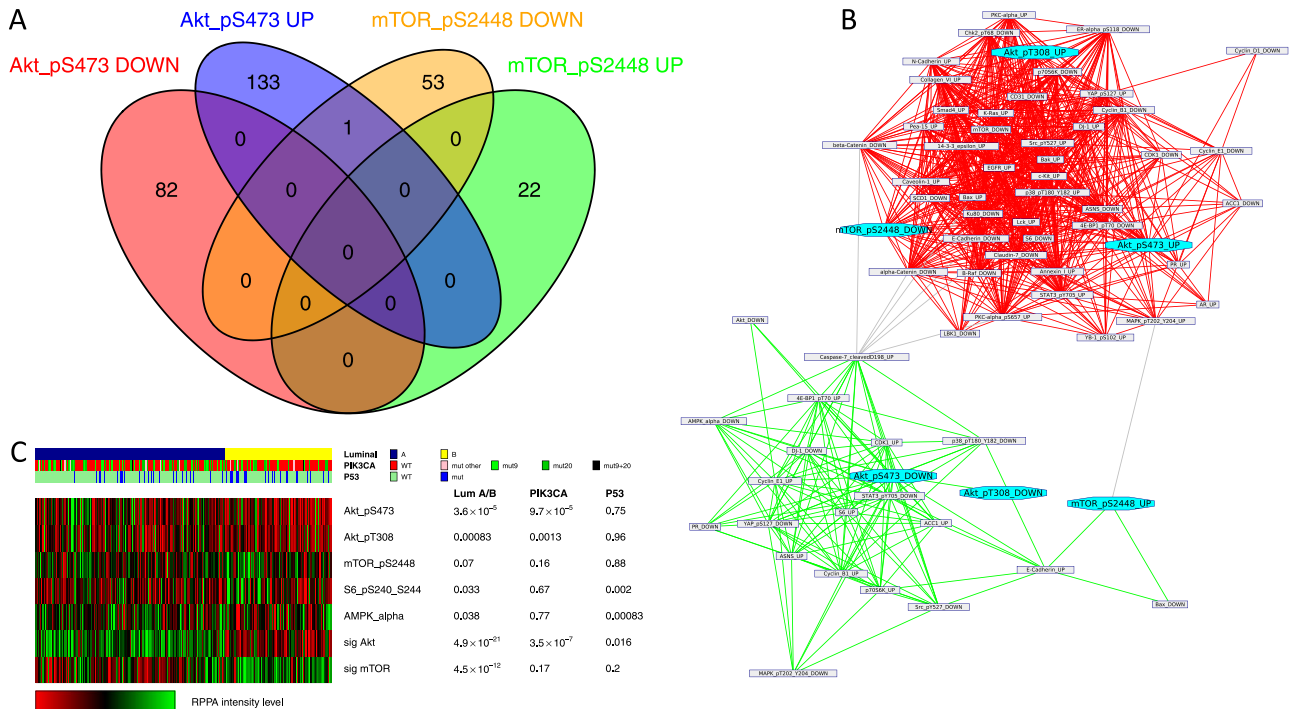
Fig. 1 pAKT and p-mTOR signatures derived from the TCGA. **a** Venn diagram shows no intersection between the pAKT and p-mTOR gene signatures. **b** Network representation of the gene signatures. Each node represents the genes up-regulated or down-regulated in the signature. Edges show signatures sharing a significant number of genes. Network clustering shows the tendency of these signatures to cluster together according to their proliferation status. **c** Integrated analysis of the *PIK3CA*/pAKT/m-TOR pathway in the TCGA. Luminal breast cancer subtypes differ by pAKT and p-mTOR activity. The panel includes a protein-based (RPPA) proteomic status. Tumors were ordered first by mRNA subtype (luminal A versus B). *P* values were calculated using the Mann–Whitney test

analysis using the pAKT and p-mTOR signatures, as well as other RPPA-derived signatures that demonstrated a significant intersection with them, identified two main sub-networks according mainly to their proliferation status, namely pAKT-high/p-mTOR-low and pAKT-low/p-mTOR high characterized with low and high proliferation levels, respectively (Fig. 1b). These observations were confirmed when analyzing the TCGA RPPA dataset (Figs. 1c and S1). AKT pathway was more often activated (elevated pAKT expression) in luminal A cancers whereas mTOR pathway was more often activated (elevated p-mTOR and pS6, an mTOR downstream target), in luminal B subtypes (Figs. 1c and S2). Next, we sought to determine how the pAKT and p-mTOR signatures correlate with other signatures and RPPA markers of the pathway. As shown in Fig. S3, the pAKT signature negatively correlates with downstream effectors of the pathway while the p-mTOR signature positively correlates with them.

Altogether, these results demonstrate that the pAKT and p-mTOR pathways, assessed through these RPPA-based gene expression signatures, have exclusive distribution according to luminal molecular subtypes and are not necessarily linearly connected.
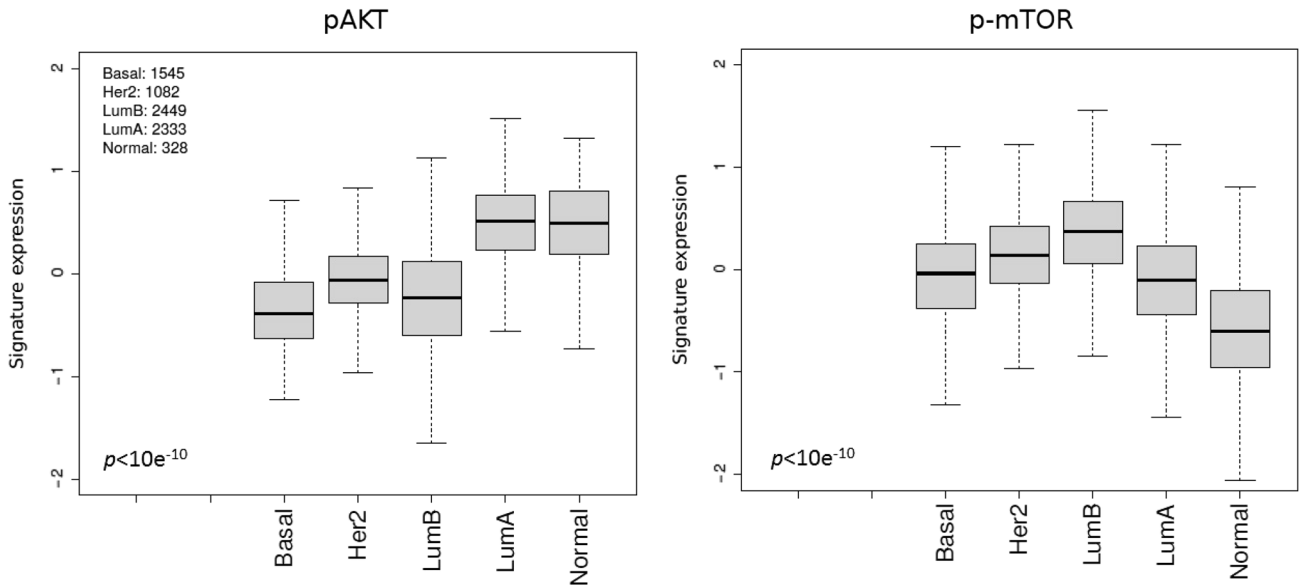
Association of the pAKT and p-mTOR pathway activation with clinical outcome in patients with ER-positive early BC

To ascertain the impact of each pathway on outcomes in ER-positive BC, we applied the pAKT and p-mTOR signatures on a dataset composed of 38 publicly available microarray datasets. We first assessed whether pAKT or p-mTOR pathway activation were associated with any particular luminal subtype. As expected, in the pooled set analysis pAKT pathway activation was significantly associated will luminal A cancers ($p < 10^{-10}$) whereas p-mTOR pathway activation was associated with luminal B cancers ($p <$
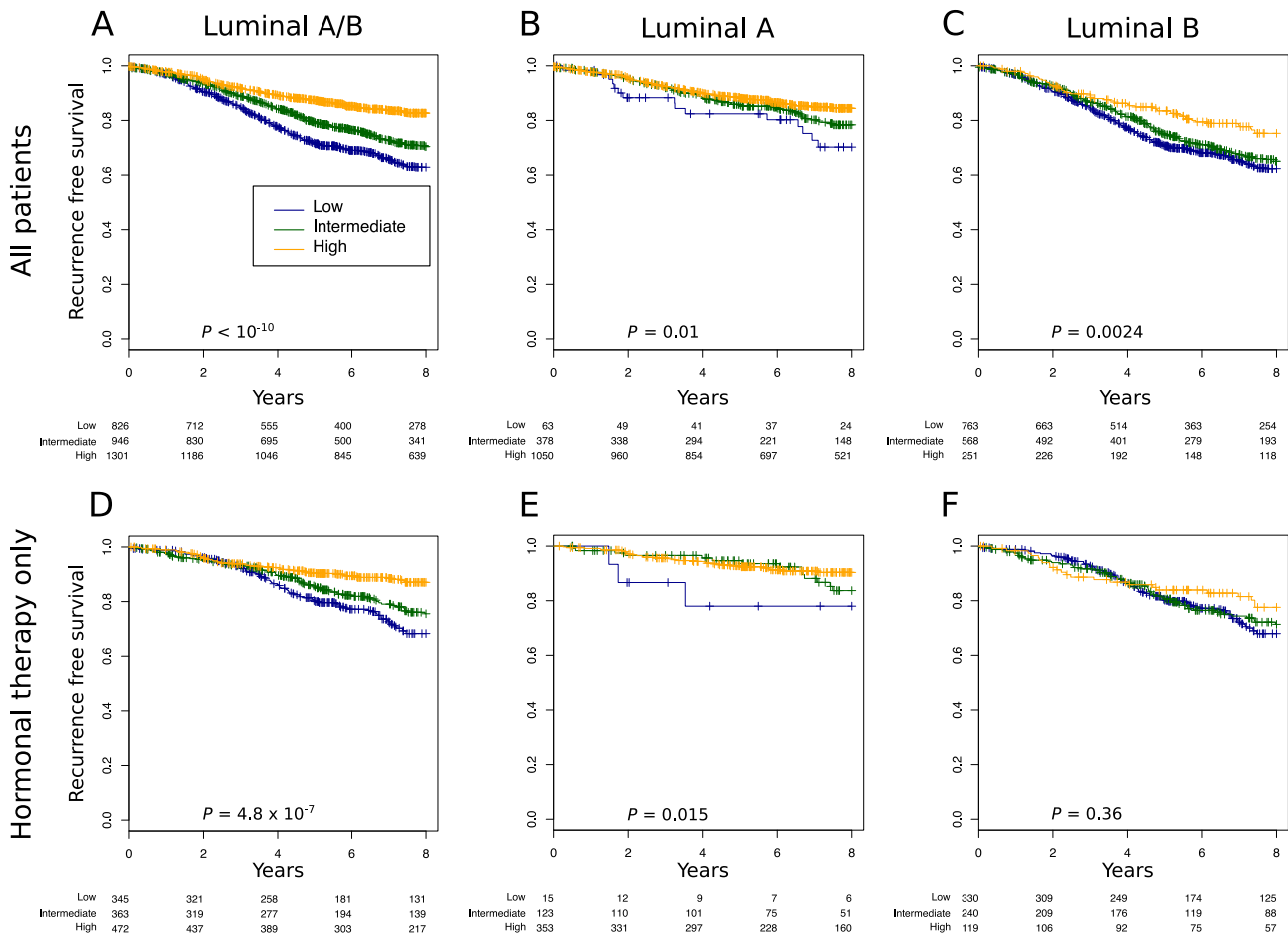
10^{-10}) (Fig. 2). We next assessed whether pAKT and p-mTOR pathway activation were correlated with outcomes (RFS) in ER-positive patients with relapse data available. As shown in Figs. 3 and 5, pAKT pathway activation was significantly associated with better outcomes in all luminal patients (multivariable HR, 0.79; 95% CI, 0.74–0.85; $p = 2.5 \times 10^{-10}$). Similar results were obtained with a dataset consisting of patients treated with endocrine therapy only (multivariable HR, 0.82; 95% CI, 0.73–0.93; $p = 0.002$). Indeed, patients with pAKT pathway activation had better outcomes irrespective of their specific subtype (luminal A multivariable HR, 0.85; 95% CI, 0.75–0.96; $p = 0.01$; luminal B HR, 0.91; 95% CI, 0.83–0.99; $p = 0.033$). In contrast, patients with p-mTOR pathway activation had significantly worse outcomes in all luminal patients (multivariable HR, 1.1; 95% CI, 1.1–1.2; $p = 9.9 \times 10^{-4}$) and this remained true when tested in the dataset consisting of patients treated with endocrine therapy only (multivariable HR, 1.2; 95% CI, 1.1–1.4; $p = 0.004$) (Figs. 4 and 5).

Next, we assessed the association between the *PIK3CA* and *P53* mutation status and pAKT and p-mTOR pathway activation in an independent set, namely the TCGA BC patients with RNA-sequence gene expression data that were not used to design the signatures. While the pAKT signature was associated with *PIK3CA* mutations ($p = 0.0001$), the p-mTOR signature was not ($p = 0.22$) (Fig. 6a, b). The opposite was true for *P53* mutations, which were positively correlated with p-mTOR pathway activation ($p = 0.04$), and negatively correlated with pAKT ($p = 0.0003$) (Fig. 6i, j). Analysis of the *PIK3CA* mutations by exon led to similar results (Fig. 6), although mutations outside of exons 9 and 20 seemed less associated with pAKT pathway activation.

Finally, in an effort to identify whether these signatures could predicts response to mTOR inhibitors, we evaluated another data set of neo-adjuvant patients treated with Everolimus.[13] Analysis of

pAKT

Basal: 1545
Her2: 1082
LumB: 2449
LumA: 2333
Normal: 328

$p < 10e^{-10}$

Signature expression

Basal Her2 LumB LumA Normal

p-mTOR

$p < 10e^{-10}$

Signature expression

Basal Her2 LumB LumA Normal

Fig. 2 pAKT (right) and p-mTOR (left) gene signatures expression in publicly available microarray datasets according to the PAM50 breast cancer subtype. Kruskal–Wallis p-value is shown

A Luminal A/B — All patients

Recurrence free survival

Low
Intermediate
High

$P < 10^{-10}$

	0	2	4	6	8
Low	826	712	555	400	278
Intermediate	946	830	695	500	341
High	1301	1186	1046	845	639

B Luminal A

$P = 0.01$

	0	2	4	6	8
Low	63	49	41	37	24
Intermediate	378	338	294	221	148
High	1050	960	854	697	521

C Luminal B

$P = 0.0024$

	0	2	4	6	8
Low	763	663	514	363	254
Intermediate	568	492	401	279	193
High	251	226	192	148	118

D — Hormonal therapy only

Recurrence free survival

$P = 4.8 \times 10^{-7}$

	0	2	4	6	8
Low	345	321	258	181	131
Intermediate	363	319	277	194	139
High	472	437	389	303	217

E

$P = 0.015$

	0	2	4	6	8
Low	15	12	9	7	6
Intermediate	123	110	101	75	51
High	353	331	297	228	160

F

$P = 0.36$

	0	2	4	6	8
Low	330	309	249	174	125
Intermediate	240	209	176	119	88
High	119	106	92	75	57

Fig. 3 High pAKT gene signature expression is associated with good prognosis in the luminal subtype. **a–f** We assessed the prognostic value of tertiles of pAKT gene signature expression in: **a** all luminal patients treated or not treated ($n = 3073$), **b** luminal A ($n = 1491$), **c** Luminal B ($n = 1582$), **d** all luminal treated with only hormonal therapy ($n = 1180$), **e** luminal A treated with only hormonal therapy ($n = 491$), and **f** luminal B treated with only hormonal therapy ($n = 689$). Significance (p-value) of differences in survival between patient groups defined by tertiles of pAKT signature expression is estimated by log-rank test. The analysis presented includes patients with lymph node-negative and lymph node-positive cancers
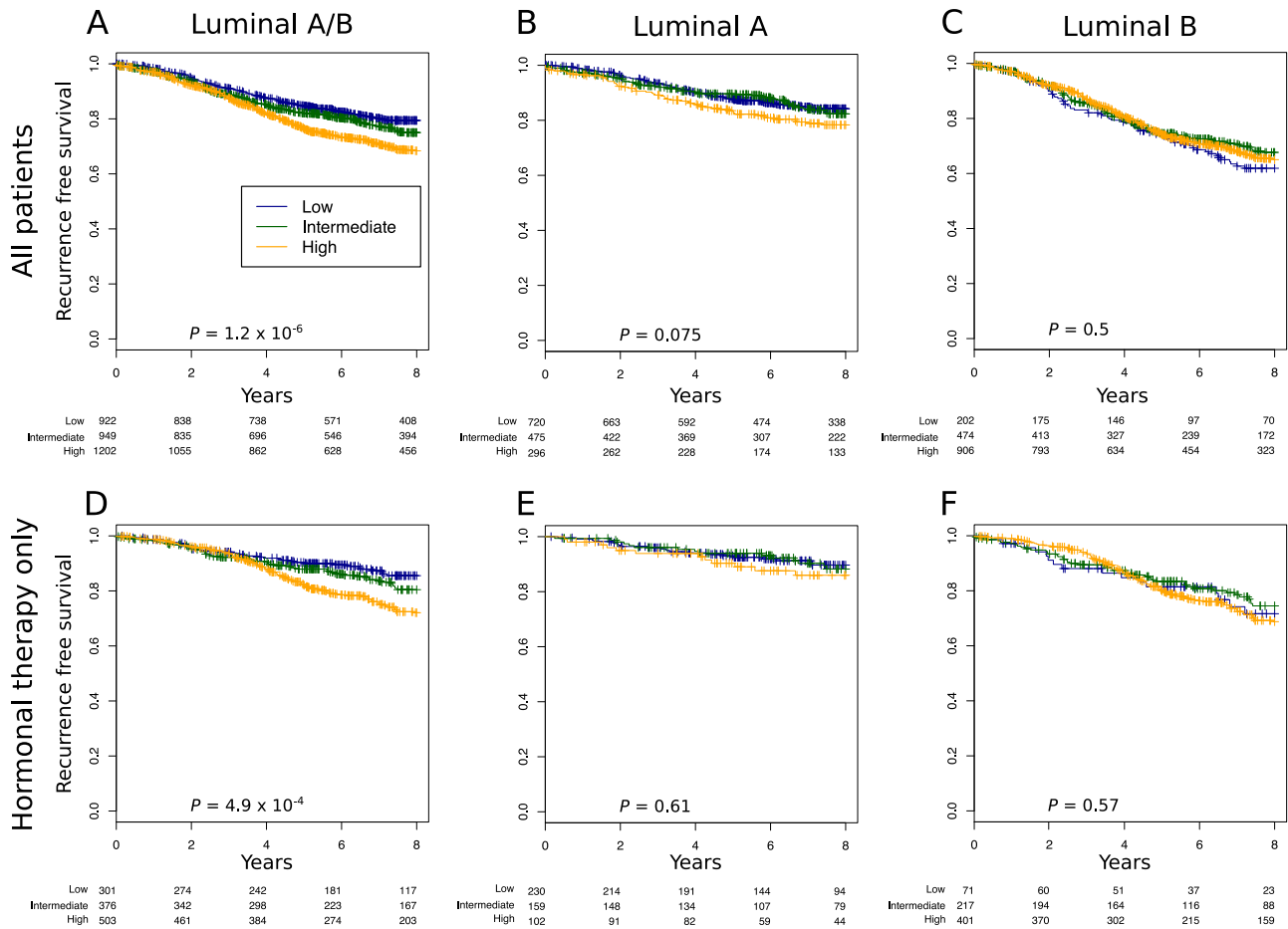
Fig. 4 High p-mTOR gene signature expression is associated with bad prognosis in the luminal subtype. **a–f** We assessed the prognostic value of tertiles of p-mTOR gene signature expression in: **a** all luminal patients treated or not treated ($n = 3073$), **b** luminal A ($n = 1491$), **c** Luminal B ($n = 1582$), **d** all luminal treated with only hormonal therapy ($n = 1180$), **e** luminal A treated with only hormonal therapy ($n = 491$), and **f** luminal B treated with only hormonal therapy ($n = 689$). Significance (p values) of differences in survival between patient groups defined by tertiles of p-mTOR signature expression is estimated by log-rank test. The analysis presented includes patients with lymph node-negative and lymph node-positive cancers

the correlations between the effectiveness of this treatment and the developed signatures suggests as expected that the pAKT signature is associated with less response to Everolimus ($r = 0.45$; $p = 0.031$, Fig. S4).

Overall, our data suggests that pAKT and p-mTOR pathway activation as assessed through the respective signatures, despite being major components of the same overarching pathway (PI3K), have distinctly different impacts on disease biology and consequently on outcomes in early disease.
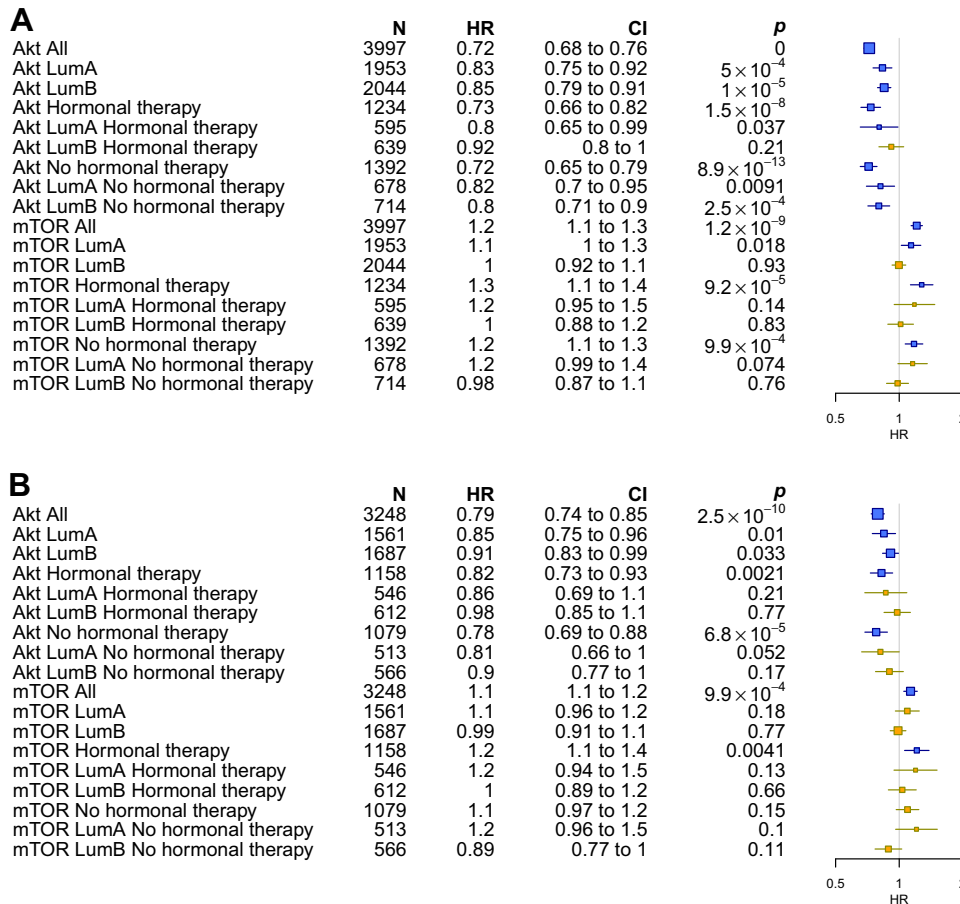
DISCUSSION

The goal of the present study was to better understand the distinct contribution to disease biology and clinical outcomes of signaling through the AKT and mTOR downstream pathways, which typically occur as part of PI3K pathway activation in luminal BCs. We found that pAKT and p-mTOR were differentially expressed according to luminal subtypes, implying different degrees of pathway activation, and that, more importantly, the pathways were not linearly connected. Additionally, we found that pAKT pathway activation was positively associated with *PIK3CA* mutations whereas the opposite was observed with p-mTOR pathway activation. In contrast, pAKT pathway activation was associated with good clinical outcome despite its known

tumorigenic effects. Between 30% and 40% of BCs, especially ER-positive tumors, have mutations in *PIK3CA*.[14] The vast majority of the *PIK3CA* mutations are missense mutations which are positioned in the helical domain (exon 9, mostly: E545K and E542K) and the kinase domain (exon 20, mostly H1047R) in hotspot clusters.[15] These mutations have direct effect on AKT phosphorylation. The effect of *PIK3CA* mutations/pAKT on prognosis is mixed in early BC.[16] We found that exons 9 and 20 mutations in *PIK3CA* were more associated with pAKT than mutations in other exons.

We previously reported that *PIK3CA* mutations were associated with improved outcome and low levels of signaling through the mTOR pathway in BC.[4,17] Several possible hypotheses were raised regarding the reasons for this. Some data available on PP2A and PML, both known to have an inhibitory effect on both AKT and mTOR,[18,19] have suggested that they may be upregulated in *PIK3CA*-activated tumors. Negative feedback regulation in PI3K-mediated cells through the insulin receptor substrate[20] and relatively weak pathway activation in *PIK3CA*-mutated cancers have also been suggested as possible explanations for low levels of signaling through mTOR in ER-positive BC.

According to our findings only pAKT pathway activation was found to be significantly different between the luminal subtypes (A and B) and *PIK3CA* wildtype versus mutant, whereas p-mTOR

A

	N	HR	CI	p
Akt All	3997	0.72	0.68 to 0.76	0
Akt LumA	1953	0.83	0.75 to 0.92	5×10^{-4}
Akt LumB	2044	0.85	0.79 to 0.91	1×10^{-5}
Akt Hormonal therapy	1234	0.73	0.66 to 0.82	1.5×10^{-8}
Akt LumA Hormonal therapy	595	0.8	0.65 to 0.99	0.037
Akt LumB Hormonal therapy	639	0.92	0.8 to 1	0.21
Akt No hormonal therapy	1392	0.72	0.65 to 0.79	8.9×10^{-13}
Akt LumA No hormonal therapy	678	0.82	0.7 to 0.95	0.0091
Akt LumB No hormonal therapy	714	0.8	0.71 to 0.9	2.5×10^{-4}
mTOR All	3997	1.2	1.1 to 1.3	1.2×10^{-9}
mTOR LumA	1953	1.1	1 to 1.3	0.018
mTOR LumB	2044	1	0.92 to 1.1	0.93
mTOR Hormonal therapy	1234	1.3	1.1 to 1.4	9.2×10^{-5}
mTOR LumA Hormonal therapy	595	1.2	0.95 to 1.5	0.14
mTOR LumB Hormonal therapy	639	1	0.88 to 1.2	0.83
mTOR No hormonal therapy	1392	1.2	1.1 to 1.3	9.9×10^{-4}
mTOR LumA No hormonal therapy	678	1.2	0.99 to 1.4	0.074
mTOR LumB No hormonal therapy	714	0.98	0.87 to 1.1	0.76

B

	N	HR	CI	p
Akt All	3248	0.79	0.74 to 0.85	2.5×10^{-10}
Akt LumA	1561	0.85	0.75 to 0.96	0.01
Akt LumB	1687	0.91	0.83 to 0.99	0.033
Akt Hormonal therapy	1158	0.82	0.73 to 0.93	0.0021
Akt LumA Hormonal therapy	546	0.86	0.69 to 1.1	0.21
Akt LumB Hormonal therapy	612	0.98	0.85 to 1.1	0.77
Akt No hormonal therapy	1079	0.78	0.69 to 0.88	6.8×10^{-5}
Akt LumA No hormonal therapy	513	0.81	0.66 to 1	0.052
Akt LumB No hormonal therapy	566	0.9	0.77 to 1	0.17
mTOR All	3248	1.1	1.1 to 1.2	9.9×10^{-4}
mTOR LumA	1561	1.1	0.96 to 1.2	0.18
mTOR LumB	1687	0.99	0.91 to 1.1	0.77
mTOR Hormonal therapy	1158	1.2	1.1 to 1.4	0.0041
mTOR LumA Hormonal therapy	546	1.2	0.94 to 1.5	0.13
mTOR LumB Hormonal therapy	612	1	0.89 to 1.2	0.66
mTOR No hormonal therapy	1079	1.1	0.97 to 1.2	0.15
mTOR LumA No hormonal therapy	513	1.2	0.96 to 1.5	0.1
mTOR LumB No hormonal therapy	566	0.89	0.77 to 1	0.11

Fig. 5 Forest plots showing the hazard ratios of the recurrence free survival of pAKT and p-mTOR gene signatures treated as a continuous variable using Cox univariate **a** and multivariable analysis **b**, in the pooled analysis. For multivariate analysis, we considered the following variables: age, tumor size, grade, and nodal status. Signatures with nominal significant effect ($p < 0.05$) are shown in blue

pathway activation was not significant for both. The inconclusive and relative activation of p-mTOR by the mutant *PIK3CA* may be also attributed to the different roles and activators of mTOR and the fact that mTOR is at the cross section of multiple signaling pathways. Several studies have clearly demonstrated that mTOR is a direct substrate for the AKT kinase and identified Serine2448 as the AKT target site in mTOR.[21] However, additional studies have demonstrated that rapamycin, an inhibitor of mTOR function, blocks serum-stimulated Serine2448 phosphorylation of mTOR in an AKT-independent manner and identified S6 kinase as a major effector of mTOR phosphorylation at Serine2448.[22] Indeed, our analysis of the TCGA data shows that the S6 kinase (downstream of mTOR) is associated with luminal B and *P53* mutations suggesting that while mTOR itself is at the cross section of conflicting pathways its downstream targets are not *PIK3CA* dependent. In addition, there are alternative kinases that can activate the mTOR pathway independently of AKT, such as RSK which leads to phosphorylation of TSC resulting in increased mTOR signaling and the PDK1–SGK1 axis that can sustain mTOR activity upon AKT suppression.[23–25]

The primary TCGA report, which investigated all BC subtypes,[26] confirmed a high frequency of *PIK3CA* mutations in luminal BC. Multiple platforms, which examined the relationship between *PIK3CA* mutation and protein expression, have demonstrated that pAKT and pS6 were not elevated in *PIK3CA*-mutated luminal cancers; instead, they were highly expressed in basal-like and HER2 subtypes. Our dataset, which is restricted to luminal cancers, demonstrated that pAKT pathway activation is associated with

luminal A/*PIK3CA* mutations and good prognosis, while p-mTOR/ pS6 is not, suggesting that the presence of HER2 and basal subtype in the primary analysis of the TCGA masked these observations.

In conclusion our data suggest that the AKT and mTOR pathways are not linearly connected in luminal BCs. pAKT pathway activation is associated with *PIK3CA* mutations, luminal A and good prognosis, while p-mTOR pathway activation is associated with luminal B, *P53* mutations, and bad prognosis. These results may have important clinical implications considering that in low p-mTOR BCs, treatment with mTOR inhibitors, such as everolimus, which is highly toxic, will possibly be of lower value since the pathway is not activated. Additionally, pAKT pathway activation, as measured through our gene signature, can add to presently used outcome prediction tools in both luminal A and luminal B tumors.

METHODS

Computation of RPPAs-based signatures

We downloaded clinic-pathological, normalized gene expression and RPPA data from the publicly available TCGA repository using its online bioinformatics tools[26] (Fig. S1 flow chart). ER-positive early BCs were analyzed based on the RPPA proteomic levels. 265 samples with available gene expression and RPPA data were considered as luminal (166 Luminal A and 99 Luminal B according to PAM50 computed on the cBioPortal website[27]). To identify the genes that were differentially expressed between the low and high expression groups, and to find the genes that

A. Sonnenblick et al.
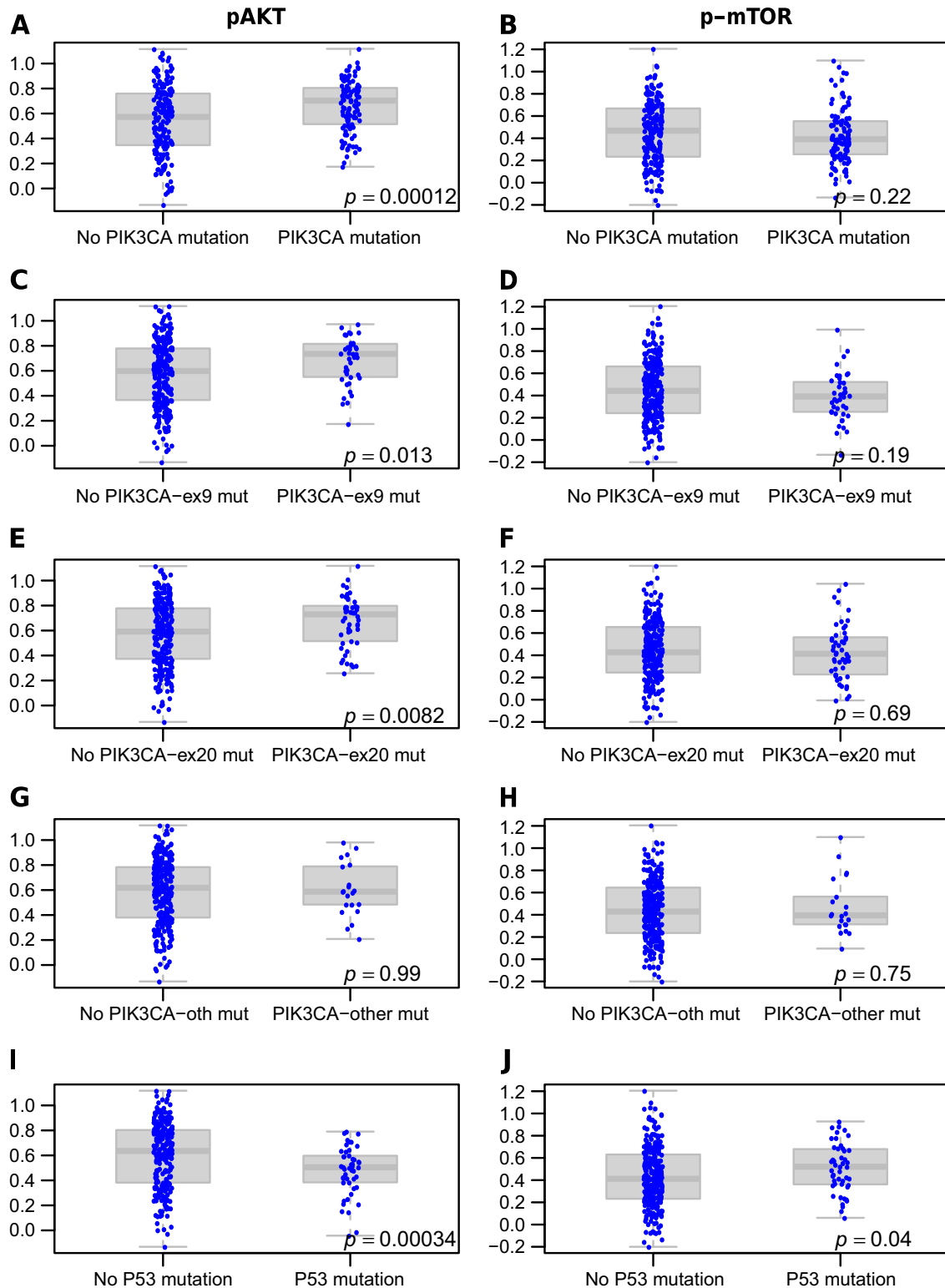
Fig. 6 Expression of pAKT (**a, c, e, g, i**) and p-mTOR (**b–d, f, h, j**) gene signatures levels in *PIK3CA* (all, exon9, exon20 or others) and *P53* mutated and wild type samples in an independent RNA sequencing set ($n = 309$)

would optimize the predictive power of our signatures, we used a machine learning approach as previously described.[28] After this process, we were left with 69 signatures (Supplementary Data) presenting a relevant AUC for proteomic status prediction. Among others, p-mTOR achieved an AUC of 0.71 ($p \sim 10^{-6}$) and pAKT an AUC of 0.77 ($p \sim 10^{-11}$) in both luminal A and B cancers.

Code availability

The expression levels of the signatures in the gene expression datasets were computed as previously described.[28] In brief, we evaluated using a nested 10-fold cross validation the maximal Benjamini–Hochberg false discovery rate and the minimal gene fold change that would optimize the ability of the differentially expressed genes to predict the high/low status

Table 1. The sources and locations for the 38 gene expression datasets analyzed

Dataset	No. of patients	Permanent identifier	References
NKI	337	10.1038/415530a	[31,32]
UCSF	162	GSE123833	[33,34]
STNO2	122	GSE4335	[35]
NCI	99	10.1073/pnas.1732912100	[36]
UNC4	337	GSE18229	[37]
CAL	118	E-TABM-158	[38]
MDA4	129 (65)‡	GSE123832	[39,40]
KOO	88	GSE123831	[41]
HLP	53	E-TABM-543	[42]
EXPO	353	GSE2109	[43]
VDX	344	GSE2034/GSE5327	[44,45]
MSK	99	GSE2603	[46]
UPP	251 (190)‡	GSE3494	[47]
STK	159	GSE1456	[48]
UNT	137 (92)‡	GSE2990	[36,49]
DUKE	171	GSE3143	[50]
TRANSBIG	198	GSE7390	[51]
DUKE2	160	GSE6961	[52]
MAINZ	200	GSE11121	[53]
LUND2	105	GSE5325	[54]
LUND	143	GSE5325	[55]
FNCLCC	150	GSE7017	[56]
EMC2	204	GSE12276	[57]
MUG	152	GSE10510	[58]
NCCS	183	GSE5364	[59]
MCCC	75	GSE19177	[60]
EORTC10994	49	GSE1561	[61]
DFHCC	115	GSE19615	[62]
DFHCC2	84 (75)‡	GSE18864	[63]
DFHCC3	40 (26)‡	GSE3744	[64]
DFHCC4	129	GSE5460	[65]
MAQC2	230	GSE20194	[66]
TAM	345 (242)‡	GSE6532/GSE9195	[67]
MDA5	298	GSE17705	[68]
VDX3	136	GSE12093	[69]
PNC	248	GSE20713	[70]
TCGA	517	https://tcga-data.nci.nih.gov/docs/publications/brca_2012/	[26]
METABRIC	1643	EGAS00000000083	[14]

‡Duplicated patients were removed from few datasets for the estimation of concordance and prognostic value

of the RPPA in luminal A and B patients together and separately. While the parameters were selected in a 10-fold cross validation, the procedure itself was assessed using a nested cross validation. All analyses were performed using the genefu package of the R (v3.2)/bioconductor (v1.18) statistical suite.

Network analysis and clustering
Network analysis and clustering was performed as previously described.[28] The MCL graph clustering algorithm[29] was applied. Visualization of the network has been rendered using the yED software.

Gene-expression data and statistical analyses
We analyzed 38 gene expression datasets totaling more than 7000 tumors (detailed in Table 1). To ensure comparability of expression values across multiple data sets, a 0.95 quantile normalization was performed. Differences in expression of pAKT and p-mTOR signatures according to subtype were examined using the Kruskal–Wallis test. Survival outcome data are presented as recurrence free survival (RFS). Survival plots according to the pAKT and p-mTOR signatures tertiles were drawn using the Kaplan–Meier method. Association of the signatures (i.e. pathway activation) with good or bad outcomes were computed using uni-variate or multi-variate Cox regression analyses and data were presented as forest plots. For multivariate analysis, we considered the following variables: age, tumor size, grade, and nodal status. To assess the correlation between the *PIK3CA* mutation status and AKT and p-mTOR gene pathway activation, as represented by the gene signature scores, we analyzed the TCGA cohort of RNA sequenced data that was not used for the computation of the signatures (309 samples), and for which both mutational and gene expression data were available. Each sample was considered as mutated or not (so a sample with four mutations was considered just like a sample with one mutation). All mutations were taken into account. *PIK3CA* mutations were also analyzed by specific exons (exons 9, 20, and all others grouped together).

Reporting summary
Further information on experimental design is available in the Nature Research Reporting Summary linked to this article.

DATA AVAILABILITY
The sources and locations for the 38 gene expression datasets analyzed during the current study are available in Table 1 and the figshare repository https://doi.org/10.6084/m9.figshare.7461776.[30]

AUTHOR CONTRIBUTIONS
A.S. conceived of the study and participated in its design and coordination, analyzed and interpreted the data and wrote the manuscript. S.B. analyzed and interpreted the data, performed the statistical analysis, and wrote the manuscript. N.P. coordinated the analysis and drafted the manuscript. D.V. carried out the statistical analysis analyzed and interpreted the data and drafted the manuscript. C.S. participated in the study design and coordination, analyzed, interpreted, the data, and wrote the manuscript. All authors read and approved the final manuscript.

ADDITIONAL INFORMATION
Supplementary information accompanies the paper on the *npj Breast Cancer* website (https://doi.org/10.1038/s41523-019-0102-1).

Competing interests: The authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

REFERENCES
1. Hennessy, B. T., Smith, D. L., Ram, P. T., Lu, Y. & Mills, G. B. Exploiting the PI3K/AKT pathway for cancer drug discovery. *Nat. Rev. Drug Discov.* **4**, 988–1004 (2005).
2. Koboldt, D. C. et al. Comprehensive molecular portraits of human breast tumours. *Nature* **490**, 61–70 (2012).
3. Campbell, I. G. Mutation of the PIK3CA gene in ovarian and breast cancer. *Cancer Res.* **64**, 7678–7681 (2004).
4. Loi, S. et al. PIK3CA mutations associated with gene signature of low mTORC1 signaling and better outcomes in estrogen receptor-positive breast cancer. *Proc. Natl Acad. Sci. USA* **107**, 10208–10213 (2010).

5. Dienstmann, R., Rodon, J., Serra, V. & Tabernero, J. Picking the point of inhibition: a comparative review of PI3K/AKT/mTOR pathway inhibitors. *Mol. Cancer Ther.* **13**, 1021–1031 (2014).

6. Baselga, J. et al. Everolimus in postmenopausal hormone-receptor–positive advanced breast cancer. *N. Engl. J. Med.* **366**, 520–529 (2012).

7. Navé, B. T., Ouwens, M., Withers, D. J., Alessi, D. R. & Shepherd, P. R. Mammalian target of rapamycin is a direct target for protein kinase B: identification of a convergence point for opposing effects of insulin and amino-acid deficiency on protein translation. *Biochem. J.* **344**, 427–431 (1999).

8. Guertin, D. A. & Sabatini, D. M. Defining the role of mTOR in cancer. *Cancer Cell* **12**, 9–22 (2007).

9. Manning, B. D. & Cantley, L. C. AKT/PKB signaling: navigating downstream. *Cell* **129**, 1261–1274 (2007).

10. Subramanian, A. et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl Acad. Sci. USA* **102**, 15545–15550 (2005).

11. Sotiriou, C. et al. Gene expression profiling in breast cancer: understanding the molecular basis of histologic grade to improve prognosis. *J. Natl. Cancer Inst.* **98**, 262–272 (2006).

12. Lim, E. et al. Transcriptome analyses of mouse and human mammary cell sub-populations reveal multiple conserved genes and pathways. *Breast Cancer Res.* **12**, R21 (2010).

13. Sabine, V. S. et al. Gene expression profiling of response to mTOR inhibitor everolimus in pre-operatively treated post-menopausal women with oestrogen receptor-positive breast cancer. *Breast Cancer Res. Treat.* **122**, 419–428 (2010).

14. Curtis, C. et al. The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. *Nature* **486**, 346–352 (2012).

15. Samuels, Y. et al. High frequency of mutations of the PIK3CA gene in human cancers. *Science* **304**, 554 (2004).

16. Yang, S. X., Polley, E. & Lipkowitz, S. New insights on PI3K/AKT pathway altera-tions and clinical outcomes in breast cancer. *Cancer Treat. Rev.* **45**, 87–96 (2016).

17. Kalinsky, K. et al. PIK3CA mutation associates with improved outcome in breast cancer. *Clin. Cancer Res.* **15**, 5049–5059 (2009).

18. Trotman, L. C. et al. Identification of a tumour suppressor network opposing nuclear Akt function. *Nature* **441**, 523–527 (2006).

19. Bernardi, R. et al. PML inhibits HIF-1alpha translation and neoangiogenesis through repression of mTOR. *Nature* **442**, 779–785 (2006).

20. Paz, K. et al. A molecular basis for insulin resistance. Elevated serine/threonine phosphorylation of IRS-1 and IRS-2 inhibits their binding to the juxtamembrane region of the insulin receptor and impairs their ability to undergo insulin-induced tyrosine phosphorylation. *J. Biol. Chem.* **272**, 29911–29918 (1997).

21. Sekulić, A. et al. A direct linkage between the phosphoinositide 3-kinase-AKT signaling pathway and the mammalian target of rapamycin in mitogen-stimulated and transformed cells. *Cancer Res.* **60**, 3504 (2000).

22. Chiang, G. G. & Abraham, R. T. Phosphorylation of mammalian target of rapa-mycin (mTOR) at Ser-2448 is mediated by p70S6 kinase. *J. Biol. Chem.* **280**, 25485–25490 (2005).

23. Roux, P. P., Ballif, B. A., Anjum, R., Gygi, S. P. & Blenis, J. Tumor-promoting phorbol esters and activated Ras inactivate the tuberous sclerosis tumor suppressor complex via p90 ribosomal S6 kinase. *Proc. Natl Acad. Sci. USA* **101**, 13489–13494 (2004).

24. Castel, P. et al. PDK1-SGK1 signaling sustains AKT-independent mTORC1 activa-tion and confers resistance to PI3Kα inhibition. *Cancer Cell* **30**, 229–242 (2016).

25. Ma, L., Chen, Z., Erdjument-Bromage, H., Tempst, P. & Pandolfi, P. P. Phosphor-ylation and functional inactivation of TSC2 by Erk implications for tuberous sclerosis and cancer pathogenesis. *Cell* **121**, 179–193 (2005).

26. Cancer Genome Atlas Network. Comprehensive molecular portraits of human breast tumours. *Nature* **490**, 61–70 (2012).

27. Gao, J. et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci. Signal.* **6**, pl1 (2013).

28. Sonnenblick, A. et al. p-STAT3 in luminal breast cancer: integrated RNA-protein pooled analysis and results from the BIG 2-98 phase III trial. *Int. J. Oncol.* **52**, 424–432 (2018).

29. Enright, A. J., Van Dongen, S. & Ouzounis, C. A. An efficient algorithm for large-scale detection of protein families. *Nucleic Acids Res.* **30**, 1575–1584 (2002).

30. Sylvain, S. B., Sonnenblick, A., & Venet, D. Gene expression data sources for in silico approach to assessing activation of AKT/mTOR signalling pathway signal-ling pathways ER-positive early BC. figshare. Fileset https://doi.org/10.6084/m9.figshare.7461776 (2018).

31. van de Vijver, M. J. et al. A gene-expression signature as a predictor of survival in breast cancer. *N. Engl. J. Med.* **347**, 1999–2009 (2002).

32. van't Veer, L. J. et al. Gene expression profiling predicts clinical outcome of breast cancer. *Nature* **415**, 530–536 (2002).

33. Korkola, J. E. et al. Identification of a robust gene signature that predicts breast cancer outcome in independent data sets. *BMC Cancer* **7**, 61 (2007).

34. Korkola, J. E. et al. Differentiation of lobular versus ductal breast carcinomas by expression microarray analysis. *Cancer Res.* **63**, 7167–7175 (2003).

35. Sorlie, T. et al. Repeated observation of breast tumor subtypes in independent gene expression data sets. *Proc. Natl Acad. Sci. USA* **100**, 8418–8423 (2003).

36. Sotiriou, C. et al. Breast cancer classification and prognosis based on gene expression profiles from a population-based study. *Proc. Natl Acad. Sci. USA* **100**, 10393–10398 (2003).

37. Prat, A. et al. Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer. *Breast Cancer Res.* **12**, R68 (2010).

38. Chin, K. et al. Genomic and transcriptional aberrations linked to breast cancer pathophysiologies. *Cancer Cell* **10**, 529–541 (2006).

39. Hess, K. R. et al. Pharmacogenomic predictor of sensitivity to preoperative che-motherapy with paclitaxel and fluorouracil, doxorubicin, and cyclophosphamide in breast cancer. *J. Clin. Oncol.* **24**, 4236–4244 (2006).

40. Liedtke, C. et al. Response to neoadjuvant therapy and long-term survival in patients with triple-negative breast cancer. *J. Clin. Oncol.* **26**, 1275–1281 (2008).

41. Huang, E. et al. Gene expression predictors of breast cancer outcomes. *Lancet Lond. Engl.* **361**, 1590–1596 (2003).

42. Natrajan, R. et al. An integrative genomic and transcriptomic analysis reveals molecular pathways and networks regulated by copy number aberrations in basal-like, HER2 and luminal cancers. *Breast Cancer Res. Treat.* **121**, 575–589 (2010).

43. Bittner M. (2005) *Expression Project for Oncology (expO)*. www.intgen.org/expo. Accessed 3 Nov 2010.

44. Wang, Y. et al. Gene-expression profiles to predict distant metastasis of lymph-node-negative primary breast cancer. *Lancet Lond. Engl.* **365**, 671–679 (2005).

45. Minn, A. J. et al. Lung metastasis genes couple breast tumor size and metastatic spread. *Proc. Natl Acad. Sci. USA* **104**, 6740–6745 (2007).

46. Minn, A. J. et al. Genes that mediate breast cancer metastasis to lung. *Nature* **436**, 518–524 (2005).

47. Miller, L. D. et al. 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 (2005).

48. Pawitan, Y. et al. Gene expression profiling spares early breast cancer patients from adjuvant therapy: derived and validated in two population-based cohorts. *Breast Cancer Res.* **7**, R953–R964 (2005).

49. Loi, S. et al. Definition of clinically distinct molecular subtypes in estrogen receptor-positive breast carcinomas through genomic grade. *J. Clin. Oncol.* **25**, 1239–1246 (2007).

50. Bild, A. H. et al. Oncogenic pathway signatures in human cancers as a guide to targeted therapies. *Nature* **439**, 353–357 (2006).

51. Desmedt, C. et al. Strong time dependence of the 76-gene prognostic signature for node-negative breast cancer patients in the TRANSBIG multicenter inde-pendent validation series. *Clin. Cancer Res.* **13**, 3207–3214 (2007).

52. Bonnefoi, H. et al. Validation of gene signatures that predict the response of breast cancer to neoadjuvant chemotherapy: a substudy of the EORTC 10994/BIG 00-01 clinical trial. *Lancet Oncol.* **8**, 1071–1078 (2007).

53. Schmidt, M. et al. The humoral immune system has a key prognostic impact in node-negative breast cancer. *Cancer Res.* **68**, 5405–5413 (2008).

54. Saal, L. H. et al. Poor prognosis in carcinoma is associated with a gene expression signature of aberrant PTEN tumor suppressor pathway activity. *Proc. Natl Acad. Sci. USA* **104**, 7564–7569 (2007).

55. Niméus-Malmström, E. et al. Gene expression profiling in primary breast cancer distinguishes patients developing local recurrence after breast-conservation surgery, with or without postoperative radiotherapy. *Breast Cancer Res.* **10**, R34 (2008).

56. Campone, M. et al. Prediction of metastatic relapse in node-positive breast cancer: establishment of a clinicogenomic model after FEC100 adjuvant regimen. *Breast Cancer Res. Treat.* **109**, 491–501 (2008).

57. Bos, P. D. et al. Genes that mediate breast cancer metastasis to the brain. *Nature* **459**, 1005–1009 (2009).

58. Calabrò, A. et al. Effects of infiltrating lymphocytes and estrogen receptor on gene expression and prognosis in breast cancer. *Breast Cancer Res. Treat.* **116**, 69–77 (2009).

59. Yu, K. et al. A precisely regulated gene expression cassette potently modulates metastasis and survival in multiple solid cancers. *PLoS Genet.* **4**, e1000129 (2008).

60. Waddell, N. et al. Subtypes of familial breast tumours revealed by expression and copy number profiling. *Breast Cancer Res. Treat.* **123**, 661–677 (2010).

61. Farmer, P. et al. Identification of molecular apocrine breast tumours by micro-array analysis. *Oncogene* **24**, 4660–4671 (2005).

62. Li, Q. et al. Minimising immunohistochemical false negative ER classification using a complementary 23 gene expression signature of ER status. *PloS ONE* **5**, e15031 (2010).

63. Silver, D. P. et al. Efficacy of neoadjuvant Cisplatin in triple-negative breast cancer. *J. Clin. Oncol.* **28**, 1145–1153 (2010).

64. Richardson, A. L. et al. X chromosomal abnormalities in basal-like human breast cancer. *Cancer Cell* **9**, 121–132 (2006).

65. Lu, X. et al. Predicting features of breast cancer with gene expression patterns. *Breast Cancer Res. Treat.* **108**, 191–201 (2008).

66. Popovici, V. et al. Effect of training-sample size and classification difficulty on the accuracy of genomic predictors. *Breast Cancer Res.* **12**, R5 (2010).

67. Loi, S. et al. Predicting prognosis using molecular profiling in estrogen receptor-positive breast cancer treated with tamoxifen. *BMC Genom.* **9**, 239 (2008).

68. Symmans, W. F. et al. Genomic index of sensitivity to endocrine therapy for breast cancer. *J. Clin. Oncol.* **28**, 4111–4119 (2010).

69. Zhang, Y. et al. The 76-gene signature defines high-risk patients that benefit from adjuvant tamoxifen therapy. *Breast Cancer Res. Treat.* **116**, 303–309 (2009).

70. Dedeurwaerder, S. et al. DNA methylation profiling reveals a predominant immune component in breast cancers. *EMBO Mol. Med.* **3**, 726–741 (2011).