IRIS AperTO

UNIVERSITÀ DEGLI STUDI DI TORINO

AperTO - Archivio Istituzionale Open Access dell'Università di Torino

Integrative molecular and functional profiling of ERBB2-amplified breast cancers identifies new genetic dependencies.

This is the author's manuscript

Original Citation:

Availability:

This version is available http://hdl.handle.net/2318/128908 since

Published version:

DOI:10.1038/onc.2012.625

Terms of use:

Open Access

Anyone can freely access the full text of works made available as "Open Access". Works made available under a Creative Commons license can be used according to the terms and conditions of said license. Use of all other works requires consent of the right holder (author or publisher) if not exempted from copyright protection by the applicable law.

(Article begins on next page)

12 October 2024

UNIVERSITÀ DEGLI STUDI DI TORINO

This is an author version of the contribution published on:
Questa è la versione dell'autore dell'opera:
Oncogene. 2014 Jan 30;33(5):619-31. doi: 10.1038/onc.2012.625

The definitive version is available at:
La versione definitiva è disponibile alla URL:
http://www.nature.com/onc/journal/v33/n5/full/onc2012625a.html

Integrative molecular and functional profiling of ERBB2-amplified breast cancers identifies new genetic dependencies

K-K Shiu[1], D Wetterskog[1], A Mackay[1], R Natrajan[1], M Lambros[1], D Sims[1], I Bajrami[1], R Brough[1], J Frankum[1], R Sharpe[1], C Marchio[2],

H Horlings[3], F Reyal[3], M van der Vijver[3], N Turner[1], JS Reis-Filho[1], CJ Lord[1] and A Ashworth[1]

Overexpression of the receptor tyrosine kinase ERBB2 (also known as HER2) occurs in around 15% of breast cancers and is driven by amplification of the ERBB2 gene. ERBB2 amplification is a marker of poor prognosis, and although anti-ERBB2-targeted therapies have shown significant clinical benefit, de novo and acquired resistance remains an important problem. Genomic profiling has demonstrated that ERBB2þve breast cancers are distinguished from ERþve and 'triple-negative' breast cancers by harbouring not only the ERBB2 amplification on 17q12, but also a number of co-amplified genes on 17q12 and amplification events on other chromosomes. Some of these genes may have important roles in influencing clinical outcome, and could represent genetic

dependencies in ERBB2þve cancers and therefore potential therapeutic targets. Here, we describe an integrated genomic, gene expression and functional analysis to determine whether the genes present within amplicons are critical for the survival of ERBB2þve breast tumour cells. We show that only a fraction of the ERBB2-amplified breast tumour lines are truly addicted to the ERBB2 oncogene at the mRNA level and display a heterogeneous set of additional genetic dependencies. These include an addiction to the transcription factor gene TFAP2C when it is amplified and overexpressed, suggesting that TFAP2C represents a genetic dependency in some ERBB2þve breast cancer cells.

Keywords: HER2-positive breast cancer; microarrays; gene expression; comparative genomic hybridization; siRNA screens; TFAP2C

INTRODUCTION

Amplification and overexpression of the gene encoding the receptor tyrosine kinase ERBB2 is present in around 15% of all breast cancers, and is associated with aggressive disease and poor clinical outcome. Anti-ERBB2-targeted therapies such as trastuzumab and lapatinib have been shown to significantly increase the survival of patients with ERBB2þve cancer,1,2 but de novo resistance is common with response rates of B30% in the metastatic setting.3 Therefore, the positive predictive value of ERBB2 amplification/ overexpression for trastuzumab sensitivity is relatively low.

Furthermore, acquired resistance to trastuzumab inevitably occurs in the metastatic setting as represented by a progression-free survival benefit of only 2.2 months when compared with standard chemotherapy.1 In the adjuvant setting, relapse-free survival is significantly prolonged with trastuzumab–chemotherapy combinations,

4,5 although overall survival benefits may be somewhat lower than first hoped6,7 Therefore, the identification of additional drivers of ERBB2þve cancer is required, and better prognostic and predictive biomarkers are needed.

At the genomic level, ERBB2-amplified breast cancers display complex genome profiles and a number of discrete gene

amplification events. For example, amplifications of genetic material at chromosomes 1q, 8q, 17q and 20q are relatively

common events in ERBB2-amplified breast tumours.8–10 However, with the exception of ERBB2 itself, the biological significance of these additional amplification events is not yet clear. Some of these amplicons most likely encompass bona fide oncogenes and candidate therapeutic targets.11,12 Here, we describe an integrated genetic, transcriptomic and functional analysis of a wide panel of genes amplified and overexpressed in ERBB2þve breast tumours and the identification of additional genetic dependencies in ERBB2þve breast cancer models.

RESULTS

ERBB2-amplified breast cancers are heterogeneous at the genomic level

To investigate the significance of amplification events in ERBB2þ ve breast cancer we used an integrated approach. In summary, this involved (i) generating genomic and transcriptomic profiles of a well-curated panel of primary breast cancers harbouring ERBB2 gene amplification; (ii) integrating these data to identify effects where gene amplification corresponded with high-level expression of the same gene; and (iii) the use of high throughput RNA interference screening to functionally assess whether genes that are amplified and overexpressed are also essential for tumour cell

survival. To analyse the genomic composition of ERBB2þve breast tumours, we used a high-resolution Array-based Comparative Genome Hybridization (aCGH) platform. This platform has a resolution of 50 kb and has previously been used to identify subsequently validated gene copy number aberrations in other human tumours.13–15 Using this platform, we profiled DNA extracted from 58 primary, fresh frozen, ERBB2-amplified breast cancers. To minimise the confounding effects of stromal material,16 all tumours were either determined by light microscopy to contain over 70% tumour content or microdissected to remove stromal contamination, ensuring at least 70% of tumour cell content in the final dissected material.

The main clinicopathological characteristics of these tumours are described in Supplementary Table S1. All tumours were from patients who had not received neoadjuvant therapy. ERBB2 overexpression in each tumour had been previously defined by immunohistochemistry (3þ) and ERBB2 amplification confirmed by chromogenic in situ hybridization (CISH) and/or aCGH to a 381 kb smallest region of amplification on 17q12; 34.98–35.36 Mb

(Supplementary Figure 1).

The majority of tumours exhibited a ‘firestorm’ genomic pattern17 characterized by multiple discrete amplifications

(Figure 1ai, data available on the ROCK database18 (www.rock.icr.ac.uk)), supporting previous observations.10,19 We also noted a small number of tumours (7/58, 12%) that displayed a

‘sawtooth’ pattern17 (Figure 1aii), which is reported to be more commonly associated with basal-like/triple-negative breast cancers.20

The number of amplifications seen in each tumour was variable (mean¼13 s.d.þ/_6.6, range 1–34). In total, 122 discrete and recurrent loci (nX2) of high-level gains and/or amplicons were identified (Figure 1bi). Consistent with previous studies,10,19 the most prevalent regions of amplification included 8q24 (8.8% of all tumours), 11q13 (11%) and 20q13 (13%). In our analysis we also noted moderate differences between the ERBB2þ/ERþ and ERBB2þ/ER_ tumour subgroups. Although only 0.87% of the genome was differentially gained, lost or amplified, loss of 14q32.2 was significantly more prevalent (Po0.05 Fisher’s exact test adjusted for multiple comparisons in the ERBB2þ/ER_ subgroup when compared with the ERBB2þ/ERþ subgroup. Amplification of 17q12 (Start; End 32.82:33.21 Mb) was significantly more prevalent in ERBB2þ/ERþve tumours (Po0.05) as shown in Supplementary Table S2.

This genomic profiling confirmed previous suggestions10,19 that, rather than being homogeneous at the genomic level, ERBB2-amplified breast tumours exhibit a significant level of genomic heterogeneity, though not greatly influenced by oestrogen receptor (ER) status, perhaps explaining the variable clinical behaviour of ERBB2þve disease.

Integration of genomic and transcriptomic data from ERBB2-amplified tumours

On the basis of the ERBB2 paradigm, where gene addiction effects may be represented by gene amplification that drives high-level expression, our aim was to identify similar ‘driver’ genes within the multiple amplicons we had identified.21 RNA of sufficient quality was available from most (76%, 45/58) of the tumours, which we had genomically characterized and so these were profiled using gene expression microarrays as previously described.14,22

We used two parallel methods to select genes within amplicons to functionally validate: (i) a Pearson’s correlation to determine the genes whose expression was significantly associated with changes in gene copy number and (ii) a Mann–Whitney U-test to assess correlation between transcript levels and amplification status as previously described.23 Genes identified as having significant correlation between copy number/amplification and expression using either statistical method totalled 511 (Figures 2a and b and Supplementary Table S3). In the 45 tumours, 93 unique recurrent (nX2) amplicons were identified (Figure 2c). The majority of genes (323, 63%) were identified using both methods included MTDH, PPM1D, RPS6KB1 and AURKA. The oncogenic nature of many of these genes has been previously demonstrated,13,24,25 suggesting that using this approach may constitute a valid approach for the identification of additional candidate drivers in ERBB2-amplified tumours. A total of 168 genes exclusively identified using Pearson’s correlation (168, 33%) included ERBB2, GRB7 and STARD3 on the ERBB2 amplicon, and MYC and ZNF217 on 8q24 and 20q13, respectively. An additional 20 genes (4%) were exclusively identified using the Mann–Whitney U-test and included CCND1, RAB34 and ORAOV2.

Genomic and transcriptomic profiling of an ERBB2-amplified breast tumour cell line panel

We aimed to identify amplified and overexpressed genes in ERBB2-amplified tumours and functionally assess the possibility that these same genes were also essential for tumour cell survival and, therefore, be candidate therapeutic targets. To define suitable models for high throughput short

interfering RNA (siRNA) screening, we compiled a panel of 14 ERBB2-amplified and 9 non-ERBB2-amplified breast cancer cell lines and characterized this panel by aCGH and expression arrays in a similar fashion to the tumour collection.18 As observed in the tumours, the 14 ERBB2-amplified cell lines exhibited multiple recurrent amplifications (n¼67), although common amplicons on 1q, 8q, 11q, 17q and 20q were again noted (Figure 1bii) and all displayed a 'firestorm' pattern. In addition, the ERBB2, ER and progesterone receptor status of these cell lines was defined by CISH and immunohistochemistry on a customised cell line microarray (Figures 3a–c). Finally, we also used viability profiling to define the sensitivity of each cell line to the small molecule ERBB2 inhibitor, lapatinib Figure 3d).

Functional profiling in ERBB2-amplified breast tumour models

Analysis of the aCGH and expression profiles of tumour cell lines suggested that 72% (369/511) of the amplified/overexpressed genes identified in the primary tumour analysis were also amplified in at least two cell lines (Figure 2c and Supplementary Table S3). To establish the functional significance of amplified/overexpression effects in ERBB2-amplified tumours, we optimised high throughput RNA interference screening procedures in a panel of ERBB2-amplified tumour cell lines (see Methods).

In addition to several ERBB2-amplified tumour lines we included a number of non-ERBB2-amplified breast tumour models as controls. In brief, cell lines were reverse transfected with siRNA targeting each gene in a 96-well plate format and cell viability determined 7 days later using Cell TitreGlo (Promega, Madison, WI, USA), which measures cellular ATP. To identify optimal screening conditions, cell number, siRNA concentration and duration of the screen were optimised to obtain (i) o10% reduction in viability in cells transfected with a control, non-targeting siRNA when compared with mock (no siRNA)-transfected cells; (ii) high efficiency of transfection, as defined by 470% loss of viability in cells transfected with siRNA targeting the essential kinase PLK1; and (iii) o90% cell confluency at 7 days post transfection, so as to maximize the possibility of identifying loss of viability effects. As far as was possible, cell lines with similar doubling times (maximum 48 h) were selected for screening so as to minimise the impact of differential proliferation rates across the screen as a whole. Cell lines excluded from further siRNA screening as they failed growth time criteria were SUM190, UACC812 and UACC893.

Following optimisation, each cell line was screened in triplicate using a siRNA library targeting the 369 genes previously identified as amplified/overexpressed in primary ERBB2þve breast cancers and also amplified in at least two ERBB2þve cell lines. The siRNA screening library was composed of one SMARTPool (Dharmacon, Lafayette, CO, USA) per well, where each SMARTPool consisted of four different siRNA species, all targeting the same gene. Cell viability data from each replicate screen was standardized using a Z-score/median absolute deviation approach where negative Z-scores represented cell inhibitory effects caused by siRNA transfection.26 The quality of each screen was assessed by the performance of non-targeting siRNA (siCON) and siRNA targeting PLK1 with the subsequent calculation of a dynamic range z027 for each screen (Figures 4a–c). The extent of replication between triplicate screens was assessed by Spearman's correlation coefficients (r2), and in each case we observed r2 of 40.7 for each replica comparison, suggesting highly reproducible data. In addition, unsupervised clustering of Z-score data also showed that experimental triplicates for each cell line clustered together with 100% concordance (Figure 4d). In

total, these quality metrics suggested that the effects observed were reproducible and robust, and that we were able to confidently compare viability effects across the panel of cell lines (Supplementary Table S5).

Ten ERBB2-amplified cell lines were used in the initial siRNA screen and six non-ERBB2-amplified cell lines were used as controls. Following robust quality control, two ERBB2-amplified cell lines—MDAMB453 and SUM225—were excluded from the analysis as the separation of negative and positive controls as assessed by z0 was inadequate. A Z-score ofp_2 was used as the threshold for defining significant cell inhibitory effects in each screen, where Zp_2 is approximately equal to a chance probability, Q, ofo0.02 (1 in 44) in each screen and Q¼(0.02)3 for effects that occurred in all three replica screens. Using this threshold, 128 siRNAs inhibited growth of at least one cell line and 72 siRNAs inhibited at least two cell lines. There were 20 siRNAs that inhibited at least half of the panel. The top hits ranked by Z-score for each cell line are shown in Supplementary Table S6. As a general observation, we did note that the functional viability profiles of the ERBB2-amplified cell lines were relatively diverse and only 20 siRNAs elicited inhibition in at least half the panel of 14 cell lines. This was perhaps expected, given the relative genetic heterogeneity of ERBB2-amplified primary tumours and cell lines.

Correlation of functional viability profiles with genomic and transcriptomic profiles As proof of principle that our approach could identify genetic dependencies in ERBB2-amplified breast cancer, we integrated genomic, transcriptomic and functional data for the ERBB2 gene itself. Although all ERBB2þve cell lines showed ERBB2 amplification and overexpression, only the viability of three (BT474, SKBR3 out type II or b errors for the identification of genes that were consistently overexpressed when amplified. To avoid the exclusion of potential drivers solely on statistical grounds, we also searched for recurrently amplified genes whose siRNA silencing caused significant loss of cell viability in at least two cell lines. These genes are shown in Table 1. Of these genes, TFAP2C was amplified and overexpressed in BT474 and SKBR3 (Supplementary Figure 4a). TFAP2C amplification and overexpression in these models was confirmed using a CISH probe and a specific antibody to TFAP2C on our cell line microarray (Figure 6c), as well as by western blot analysis (Supplementary Figure 4b). We also observed a significant correlation between TFAP2C copy number gain and amplification/overexpression in our tumour cohort (Supplementary Figure 4c). Furthermore, in our entire 58 ERBB2þ ve primary breast cancer data set, three tumours exhibited TFAP2C amplification and 26 exhibited TFAP2C copy number gain (www.rock.icr.ac.uk). To eliminate possible off target effects of siRNA, we then repeated the screen with three independent siRNAs to TFAP2C across nine ERBB2-amplified cell lines. We found that these three siRNAs and the original siRNA smart pool significantly inhibited two cell lines BT474 and SKBR3 (Figures 6a and b). To validate this finding, we transfected each siRNA duplex into BT474 and SKBR3 cells to establish the gene-silencing effect.

Three siRNAs significantly reduced TFAP2C protein levels 72 h following transfection (Figure 6d). In addition, silencing of TFAP2C caused significant apoptosis as shown by an increase in PARP cleavage (Figure 6d) and caspase activity (Supplementary Figure 5). Taken together, this suggested that TFAP2C when amplified and overexpressed is a potential oncogenic driver when co-amplified with ERBB2, and that its silencing causes cell death via an apoptotic pathway.

We also assessed the possibility that the effect of reducing TFAP2C could be caused by affecting ERBB2 expression as previous studies have shown that TFAP2C can bind to promoter regions upstream and downstream of ERBB230,31 and affect ERBB2 transcriptional activity.32 Therefore, we transfected three independent siRNAs that targeted TFAP2C into BT474 and SKBR3 cells. This resulted in decreased ERBB2 protein expression after 72 h (Figure 6d). However, the expression of ERBB2 cDNA via a heterologous promoter did not ostensibly alter the cell inhibitory or proapoptotic effects of TFAP2C siRNA (Supplementary Figures 5 and 6), suggesting that the inhibitory effect of TFAP2C silencing may not entirely be mediated by an effect on ERBB2 levels.

DISCUSSION

In this study, we have assessed the multiple genetic dependencies of ERBB2-amplified breast cancers, using a combination of genomic profiling and parallel functional analysis in multiple breast cancer cell lines. The integration of genomic and transcriptomic profiling in our study highlights the genomic heterogeneity in this breast cancer subtype. Although this genomic heterogeneity might underpin the variable clinical outcomes for patients with this aggressive disease, we reasoned that focussing upon those genes that are recurrently amplified and overexpressed in ERBB2þve breast tumours might be a route to identify additional genes that are critical for the survival of ERBB2þve breast tumour cells. Using this approach of integrated molecular and functional profiling we identified a number of genetic dependencies, most notably an effect involving the transcription factor gene TFAP2C. In models where TFAP2C is amplified and overexpressed, TFAP2C silencing by siRNA causes inhibition of cell growth, most likely by inducing apoptosis (Figure 6d). The sensitivity of TFAP2C-amplified/overexpressing models to siRNA silencing of this gene and the frequency of TFAP2C amplification/overexpression in ERBB2þve breast cancers we only profiled biopsies with 470% tumour content in order to avoid biases associated with stromal cell/non-neoplastic cell contamination. This approach could have increased the level of sensitivity in terms of detecting recurrent amplicons; (ii) although ERBB2-amplified tumours profiled in this study had a predominantly ‘firestorm’ genomic pattern, in keeping with previous reports,17 we observed that those ERBB2-amplified tumours with a ‘sawtooth’ genomic profile were predominantly ER_ve/ERBB2þve. Sawtooth patterns have been reported to be associated with a basal-like/’triple-negative’ phenotype and underlying DNA repair deficiencies13,23 rather than ERBB2-amplified tumours. Furthermore, we are currently the only study that has used the genomic and transcriptional aberrations observed in this breast cancer subtype and integrated this with functional profiling via high throughput siRNA screens. These findings emphasize how the genomically heterogeneous ERBB2-amplified breast cancers are even within this phenotypic subgroup, and may explain why patients diagnosed with ERBB2þve breast cancers have differing clinical outcomes despite receiving similar therapies.6,7

This is the first study to identify TFAP2C (AP2g) as being essential for cell survival in a subgroup of ERBB2-amplified breast tumour models defined by TFAP2C amplification and overexpression. Previously, a number of observations have linked TFAP2C expression to breast cancer. TFAP2C is a member of a family of five activating enhancer transcription factors (TFAP2A-G) that are essential for normal embryogenesis, as well as being linked to cell proliferation, dfferentiation and cell survival.34 TFAP2C is located on 20q13, a region which is copy number gained in B18% of all

breast cancers.35 Retrospective studies have correlated high-level TFAP2C expression with high-grade breast cancers,34 poorer prognosis in patients with invasive breast cancers and resistance to hormone therapy.36,37 This may be partly explained by the fact that it has been shown to have roles in upregulating ER-mediated gene transcription by mediating long range chromatin interactions38 and to act as repressor of transcription by inhibiting CDKN1A, a growth inhibitor.39 Here we expand these observations and demonstrate that TFAP2C amplification is significantly enriched in patients with ERBB2þve breast cancer (29/58, 50%) and that amplification also correlates with TFAP2C mRNA overexpression (Supplementary Figure 4a). TFAP2C has been shown to bind at multiple promoter sites proximal and distal of the ERBB2 gene,30,31,40 suggesting a mechanism by which TFAP2C could modulate ERBB2 signalling. Indeed, we show here that in ERBB2þve models that are inhibited by TFPA2C siRNA, ERBB2 protein levels are suppressed by TFAP2C silencing (Figure 6d). However, as transcription factors are rarely specific for single genes, we cannot discount the possibility that the dependency upon TFAP2C that we observe is also mediated by modulating the expression of a number of genes.34,38 We do note that the expression of ERBB2 cDNA from a heterologous promoter did not reverse the cell inhibitory effect of TFAP2C siRNA (Supplementary Figure 6). Notwithstanding the difficulties in replicating physiological expression of genes with cDNA expression constructs, it seems possible that this result is in keeping with the hypothesis that the dependency on TFAP2C is explained by the activity of this transcription factor on a number of genes in addition to ERBB2.

In addition to identifying TFAP2C as a genetic dependency in ERBB2þve breast cancer, a comparative analysis of anti-ERBB drug sensitivity and ERBB2 siRNA sensitivity in ERBB2þve breast cancer models confirms that only a small fraction of the ERBB2-amplified breast cancer lines are truly addicted to the ERBB2 oncogene.41–43 For example, we have shown that ERBB2þve tumour cell line models that are resistant to multiple ERBB2 and pan-ERB-targeted agents are also resistant to ERBB2 silencing by siRNA (Figure 5c). Although multiple potential mechanisms of acquired and de novo resistance to ERBB2-targeted therapies have been previously proposed,42,44–47 few have been validated in prospective clinical trials.47 Our data indicate that one fundamental reason for de novo resistance to ERBB2-targeted agents is the loss of dependency on the ERBB2 pathway, as opposed to a pharmacological mechanism, such as increased drug efflux or metabolism that could impair the ability of a drug to inhibit its target. The clinical implication of these data would be that for some patients, the failure to respond to an ERBB2-targeted agent such as lapatinib would also predict a poor response to other ERBB2-targeted therapies. We also demonstrate that ERBB3 represents a genetic dependency in a subset of ERBB2-amplified models, supporting the focus upon this ERBB2 dimer partner as an additional therapeutic target.28,48 Our comparative anti-ERB drug/ERBB2 siRNA data indicate, however, that such therapy may only be particularly effective in ERBB2þve cancers that remain addicted to ERBB2. Our study also highlights the fact that, in general, the functional dependencies in ERBB2þve breast tumour models are relatively diverse and that many of the dependencies identified are isolated to a relatively small number of models. It seems unlikely that this is an artefact of the RNAi screening approach used (for example, the effect of ERBB2 siRNA mirrors the effect of anti-ERB drugs, giving some confidence in the validity of the RNAi functional profiling) but perhaps is more reflective of the inherent genetic and functional heterogeneity of both the models and the tumours profiled.

In conclusion, we propose that TFAP2C amplification and overexpression represents a genetic dependency in ERBB2þve breast cancer. Integration of molecular and functional profiling in ERBB2þve breast cancer identifies genes that warrant further study. One particular area of focus in the future will be to correlate the expression and copy number genes such as TFAP2C in tumour biopsy material obtained from large neoadjuvant-targeted therapy trials involving anti-ERBB2-targeted agents as these become available.

MATERIALS AND METHODS

Patient tumour samples

Fifty-eight fresh frozen ERBB2þve breast cancer samples were obtained after approval by local Ethics Committees from the authors' institutions. All tumours were either ERBB2 2þ or 3þ on immunohistochemistry and were confirmed to be amplified with CISH and/or aCGH. All patients had stage I–III disease, and samples received were from the primary surgery. None of the tumours had received neoadjuvant therapy. Their clinicopathological characteristics are shown in Supplementary Table S1.

All tumours possessed 470% neoplastic cell content and had matching RNA of sufficient quality for expression arrays. DNA and RNA were extracted as previously described.13,49,50 ER and progesterone receptor status was determined by immunohistochemical analysis using the 6F11 (1:150) and PgR636 (1:200) antibodies, respectively, as previously described.13,22 The Allred scoring system was employed and tumours were considered positive when the score wasX3. HER2 gene amplification was defined based on CISH analysis using an US Food and Drug Administrator approved probe (SpotLight HER2 amplification probe, Invitrogen, Carlsbad, CA, USA) and/ or by inspection of the results of aCGH analysis, as previously described.9

Cancer cell lines

Fourteen commercially available cell lines known to harbour ERBB2 gene amplification and 9 non-ERBB2-amplified controls representing the other molecular subtypes of breast cancer were selected. They were purchased from ATCC (Manassas, VI, USA) unless otherwise stated, and were maintained as per source recommendations. These included CAL51, CAMA1, HCC1143, MCF7, MDAMB468, MDAMB231, MCF12A, SUM149 (Asterand), T47D, BT474, HCC202, HCC1569, HCC1954, MDA-MB-361, MDAMB-453, SKBR3, UACC-812, UACC-893, ZR-75.30, SUM190PT (Asterand) SUM225CWN (Asterand), JIMT-1 (DSMZ) and VP229 (EACC). DNA was extracted and purified using DNeasy (Qiagen Ltd, Crawley, West Sussex, UK) total DNA purification protocol for cultured animal cells, and quantified using Pico Green. DNA was then stored at _20 1C before use. Total RNA was extracted in a RNase-free environment using TRIzol (Invitrogen) standard manufacturer's protocol (Cat No 15596-018). RNA pellets were airdried and resuspended in a total volume of 50 ml nuclease-free water (Ambion, Paisley, UK). RNA quality and quantity were assessed by using an Eukaryote Total RNA Nano Series II Chip with an Agilent 2100 Bioanalyser (Winnersh, UK) according to the manufacturer's instructions, then stored at _80 1C before use. We used short tandem repeat profiling to verify authenticity of the panel of cell lines. We simultaneously amplified eight short tandem repeat loci, in a multiplex PCR reaction (Promega PowerPlex 1.2 System) and used the ATCC and DMSZ databases for comparison, where

possible. In addition, gain and loss of expression of particular subtype markers, described by Hollestelle et al.51 were confirmed by western blot as well as microarray analysis.

Microarray comparative genomic hybridization and gene expression profiling

All cell line and tumours were subjected to aCGH (Breakthrough Breast Cancer Research Centre) 32K BAC array platform and profiled as previously described.23 All cell lines and 45 tumours were subjected to mRNA gene expression profiling. In brief expression profiling data of cases from the NKI were generated as previously described22 using the Human Genome Oligo Set Version 3.0 arrays, which contains 34 580 probes representing 24 650 genes. These arrays were obtained from the central microarray facility at the NKI. Detailed information on the array platform and hybridization protocols can be found at http://microarrays.nki.nl. The RNA from the 14 Centro Nacional de Investigaciones Oncolo´ gicas (CNIO) tumours and all cell lines was processed using the Illumina Human Ref6 Chip v2, as per the manufacturer's protocol (www.illumina.com). Fluorescent intensities were normalized through Bead Studio, and processed using Lumi package in R as previously described.20 Before data analysis the gene lists for the NKI and CNIO cohorts were mapped to the human genome (ensembl HG18 assembly 49). Genes with missing values or ambiguous mapping information were excluded. Probes with 475% flagged values were removed. This left 17 490 and 12 235 unambiguously mapped unique genes with known Ensembl identities for analysis in the NKI and CNIO cohorts, respectively. Correlation of gene expression and aCGH data are described previously.23,52 Data acquisition and analysis were fully Minimal Information About a Microarray Experiment (MIAME) compliant. For the CNIO cohort, gene expression data are publicly available at ArrayExpress http://www.ebi.ac.uk/microarray-as/ae/ (accession number: E-TABM-543). For details of the NKI cohort, see Turner et al.23

Cell lines, compounds and siRNA

Cell lines were grown and transfected with SMARTpool siRNAs using Dharmafect 3 (DF3), Dharmafect 4 (DF4) (Dharmacon), Oligofectamine, Lipofectamine 2000 or RNAiMAX (Invitrogen). The custom library was obtained in four 96-well plates from Dharmacon. Each well in this library contained a SMARTpool of four distinct siRNA species targeting different sequences of the target transcript. Each plate included four wells each with Mock (transfection lipid only), siControl Pool1, siControl Pool2 and PLK1, Dharmacon as negative and positive controls.

Correlation of siRNA Z-score with gene expression and aCGH data

The correlation between siRNA Z-score and normalized gene expression was examined for genes where siRNA caused significant loss of viability (Zo_2). Z-score was compared with normalized gene expression using Pearson correlation coefficient. A gene was taken as being significantly correlated if the Pearson correlation coefficient was significantly different to the null hypothesis (Po0.05), the correlation was inverse, and the variation in gene expression between cells lines were significantly different as assessed by one-way analysis of variance. The number of cases harbouring amplifications, gains losses and deletions which also corresponded with Z-score hits in the same cell line was also calculated as previously described.52

Immunoblotting

Protein lysates were prepared using RIPA lysis buffer (50 nM Tris pH 8.0, 150mM NaCl, 0.1% SDS, 0.1% deoxycholic acid (DOC), 1% TritonX-100, 50mM NaF, 1mM Na3VO4and protease inhibitors). Protein concentrations were measured using BioRad Protein Assay Reagent (BioRad, Hemel Hempstead, UK). Antibodies targeting the following were used as per manufacturer's instructions: Actin, TFAP2C sc-12763 (Santa Cruz Biotechnology, Santa Cruz, CA, USA), Cleaved PARP 19F4, ERBB2 29D8, ERBB3 1B2 (Cell Signalling, Danvers, MA, USA). All secondary antibodies used for western blot analysis were horseradish peroxidase-conjugated.

Protein bands were visualised using ECL (GE Healthcare, Cardiff, UK) and MR or XAR film (Kodak).

Validation of gene silencing by siRNA Validation of RNAi gene silencing was determined by western blotting and by viability assays of silencing effects with individual siRNA duplexes (siGENOME, Dharmacon). Protein lysates were collected 48–72 h following transfection of individual siRNA for western blot analysis.

Individual siRNA sequences that were on target for TFAP2C were 50-GUAAACCAGUGGCAGAAUA-30, 50-CGACAUGCCUCACCAGAUG-30, 50-ACACUGGAGUCGCCGAAUA-30 , for ERBB2 were; 50-GGACGAAUUCUGCACAAUG-30 , 50-CUACAACACAGACACGUUU-30, 50-AGACGAAGCAUACGUGAUG-30, for ERBB3 were: 50-GCAGUGGAUUCGAGAAGUG-30, 50-AGAUUGUGCUCACGGGACA-30, 50-GUGGAUUCGAGAAGUGACA-30 and 50-GCGAUGCUGAGAACCAAUA-30 . All siRNA identities are available via the Dharmacon website (http://www.dharmacon.com).

Survival assays

For measurement of sensitivity to lapatinib, canertinib and BIBW2992 treatment, cells were plated in 96-well plates and exposed to the drug at the indicated concentrations. Cells were dosed at 24 and 96 h. After 7 days, cell viability was measured using CellTiter Glo Luminescent Cell Viability Assay (Promega). Surviving fractions were calculated and drug sensitivity curves plotted as previously described using Prism.53

Caspase activity assay

The ApoTox Glo triplex assay (Promega) was used to assess caspase 3/7 activity and viability as per the manufacturer's instructions. Cells were reverse transfected with siRNA (Dharmacon) using Lipofectamine 2000 (Invitrogen). Cells were plated in 96-well plates at a density of 20 000 cells per well. Twenty-four, and 96 h post transfection, caspase 3/7 activity and viability effects were assessed, respectively.

Immunohistochemistry

Cell line microarray blocks containing three representative replicate cores from each cell line were prepared, cut and subjected to immunohistochemical analysis with antibodies against ER, progesterone receptor, ERBB2, EGFR, and TFAP2C. All markers were scored by JR-F blinded to the identity of the samples.

Chromogenic in situ hybridization

The CISH probe for TFAP2C was constructed and optimized on the cell line microarray as previously described.54 RPCI tiling path clones were purchased from Source Bioscience: RP5-897D18, RP5-843L14 and RP4-539E24.

Statistical analysis

Statistical analysis of aCGH and gene expression data was performed with Excel and GraphPad Prism version 5.01 (GraphPad Software, La Jolla, CA, USA).

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

This work was supported by grants from Cancer Research UK, AACR as part of the Stand Up To Cancer Breast Cancer Dream Team Initiative, and Breakthrough Breast Cancer. Dr Kai-Keen Shiu is an Avon Clinical Fellow.

REFERENCES

1 Slamon DJ, Leyland-Jones B, Shak S, Fuchs H, Paton V, Bajamonde A et al. Use of chemotherapy plus a monoclonal antibody against HER2 for metastatic breast cancer that overexpresses HER2. N Engl J Med 2001; 344: 783–792.

2 Geyer CE, Forster J, Lindquist D, Chan S, Romieu CG, Pienkowski T et al. Lapatinib plus capecitabine for HER2-positive advanced breast cancer. N Engl J Med 2006; 355: 2733–2743.

3 Vogel C, Cobleigh MA, Tripathy D, Gutheil JC, Harris LN, Fehrenbacher L et al. First-line, single-agent Herceptin(R) (trastuzumab) in metastatic breast cancer. A preliminary report. Eur J Cancer 2001; 37(Suppl 1): 25–29.

4 Romond EH, Perez EA, Bryant J, Suman VJ, Geyer Jr CE, Davidson NE et al. Trastuzumab plus adjuvant chemotherapy for operable HER2-positive breast cancer. N Engl J Med 2005; 353: 1673–1684.

5 Smith I, Procter M, Gelber RD, Guillaume S, Feyereislova A, Dowsett M et al. 2-year follow-up of trastuzumab after adjuvant chemotherapy in HER2-positive breast cancer: a randomised controlled trial. Lancet 2007; 369: 29–36.

6 Gianni L, Dafni U, Gelber RD, Azambuja E, Muehlbauer S, Goldhirsch A et al. Treatment with trastuzumab for 1 year after adjuvant chemotherapy in patients with HER2-positive early breast cancer: a 4-year follow-up of a randomised controlled trial. Lancet Oncol 2011; 12: 236–244.

7 Perez EA, Romond EH, Suman VJ, Jeong JH, Davidson NE, Geyer Jr CE et al. Four-year follow-up of trastuzumab plus adjuvant chemotherapy for operable human epidermal growth factor receptor 2-positive breast cancer: joint analysis of data from NCCTG N9831 and NSABP B-31. J Clin Oncol 2011; 29: 3366–3373.

8 Chin K, DeVries S, Fridlyand J, Spellman PT, Roydasgupta R, Kuo WL et al. Genomic and transcriptional aberrations linked to breast cancer pathophysiologies. Cancer Cell 2006; 10: 529–541.

9 Arriola E, Marchio C, Tan DS, Drury SC, Lambros MB, Natrajan R et al. Genomic analysis of the HER2/TOP2A amplicon in breast cancer and breast cancer cell lines. Lab Invest 2008; 88: 491–503.

10 Staaf J, Jonsson G, Ringner M, Vallon-Christersson J, Grabau D, Arason A et al. High-resolution genomic and expression analyses of copy number alterations in HER2-amplified breast cancer. Breast Cancer Res 2010; 12: R25.

11 Shiu KK, Natrajan R, Geyer FC, Ashworth A, Reis-Filho JS. DNA amplifications in breast cancer: genotypic-phenotypic correlations. Future Oncol 2010; 6: 967–984.

12 Santarius T, Shipley J, Brewer D, Stratton MR, Cooper CS. A census of amplified and overexpressed human cancer genes. Nat Rev Cancer 2010; 10: 59–64.

13 Natrajan R, Lambros MB, Rodriguez-Pinilla SM, Moreno-Bueno G, Tan DS, Marchio C et al. Tiling path genomic profiling of grade 3 invasive ductal breast cancers. Clin Cancer Res 2009; 15: 2711–2722.

14 Mackay A, Tamber N, Fenwick K, Iravani M, Grigoriadis A, Dexter T et al. A highresolution integrated analysis of genetic and expression profiles of breast cancer cell lines. Breast Cancer Res Treat 2009; 118: 481–498.

15 Gunnarsson R, Staaf J, Jansson M, Ottesen AM, Goransson H, Liljedahl U et al. Screening for copy-number alterations and loss of heterozygosity in chronic lymphocytic leukemia--a comparative study of four differently designed, high resolution microarray platforms. Genes Chromosomes Cancer 2008; 47: 697–711.

16 Cleator SJ, Powles TJ, Dexter T, Fulford L, Mackay A, Smith IE et al. The effect of the stromal component of breast tumours on prediction of clinical outcome using gene expression microarray analysis. Breast Cancer Res 2006; 8: R32.

17 Hicks J, Krasnitz A, Lakshmi B, Navin NE, Riggs M, Leibu E et al. Novel patterns of genome rearrangement and their association with survival in breast cancer. Genome Res 2006; 16: 1465–1479.

18 Sims D, Bursteinas B, Gao Q, Jain E, MacKay A, Mitsopoulos C et al. ROCK: a breast cancer functional genomics resource. Breast Cancer Res Treat 2010; 124: 567–572.

19 Sircoulomb F, Bekhouche I, Finetti P, Adelaide J, Ben Hamida A, Bonansea J et al. Genome profiling of ERBB2-amplified breast cancers. BMC Cancer 2010; 10: 539.

20 Natrajan R, Weigelt B, Mackay A, Geyer FC, Grigoriadis A, Tan DS 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 2010; 121: 575–589.

21 Tan DS, Reis-Filho JS. Comparative genomic hybridisation arrays: high-throughput tools to determine targeted therapy in breast cancer. Pathobiology 2008; 75: 63–74.

22 Kreike B, van Kouwenhove M, Horlings H, Weigelt B, Peterse H, Bartelink H et al. Gene expression profiling and histopathological characterization of triplenegative/basal-like breast carcinomas. Breast Cancer Res 2007; 9: R65.

23 Turner N, Lambros MB, Horlings HM, Pearson A, Sharpe R, Natrajan R et al. Integrative molecular profiling of triple negative breast cancers identifies amplicon drivers and potential therapeutic targets. Oncogene 2010; 29: 2013–2023.

24 Santarius T, Shipley J, Brewer D, Stratton MR, Cooper CS. A census of amplified and overexpressed human cancer genes. Nat Rev 2009; 10: 59–64.

25 Kao J, Pollack JR. RNA interference-based functional dissection of the 17q12 amplicon in breast cancer reveals contribution of coamplified genes. Genes Chromosomes Cancer 2006; 45: 761–769.

26 Boutros M, Bras LP, Huber W. Analysis of cell-based RNAi screens. Genome Biol 2006; 7: R66.

27 Zhang JH, Chung TD, Oldenburg KR. A simple statistical parameter for use in evaluation and validation of high throughput screening assays. J Biomol Screen 1999; 4: 67–73.

28 Lee-Hoeflich ST, Crocker L, Yao E, Pham T, Munroe X, Hoeflich KP et al. A central role for HER3 in HER2-amplified breast cancer: implications for targeted therapy. Cancer Res 2008; 68: 5878–5887.

29 Schoeberl B, Pace EA, Fitzgerald JB, Harms BD, Xu L, Nie L et al. Therapeutically targeting ErbB3: a key node in ligand-induced activation of the ErbB receptor-PI3K axis. Sci Signal 2009; 2: ra31.

30 Bosher JM, Williams T, Hurst HC. The developmentally regulated transcription factor AP-2 is involved in c-erbB-2 overexpression in human mammary carcinoma. Proc Natl Acad Sci USA 1995; 92: 744–747.

31 Vernimmen D, Begon D, Salvador C, Gofflot S, Grooteclaes M, Winkler R. Identification of HTF (HER2 transcription factor) as an AP-2 (activator protein-2) transcription factor and contribution of the HTF binding site to ERBB2 gene overexpression. Biochem J 2003; 370(Pt 1): 323–329.

32 Ailan H, Xiangwen X, Daolong R, Lu G, Xiaofeng D, Xi Q et al. Identification of target genes of transcription factor activator protein 2 gamma in breast cancer cells. BMC Cancer 2009; 9: 279.

33 Chin SF, Wang Y, Thorne NP, Teschendorff AE, Pinder SE, Vias M et al. Using array comparative genomic hybridization to define molecular portraits of primary breast cancers. Oncogene 2007; 26: 1959–1970.

34 Turner BC, Zhang J, Gumbs AA, Maher MG, Kaplan L, Carter D et al. Expression of AP-2 transcription factors in human breast cancer correlates with the regulation of multiple growth factor signalling pathways. Cancer Res 1998; 58: 5466–5472.

35 Zhao C, Yasui K, Lee CJ, Kurioka H, Hosokawa Y, Oka T et al. Elevated expression levels of NCOA3, TOP1, and TFAP2C in breast tumors as predictors of poor prognosis. Cancer 2003; 98: 18–23.

36 Gee JM, Eloranta JJ, Ibbitt JC, Robertson JF, Ellis IO, Williams T et al. Overexpression of TFAP2C in invasive breast cancer correlates with a poorer response to anti-hormone therapy and reduced patient survival. J Pathol 2009; 217: 32–41.

37 Guler G, Iliopoulos D, Guler N, Himmetoglu C, Hayran M, Huebner K. Wwox and Ap2gamma expression levels predict tamoxifen response. Clin Cancer Res 2007; 13: 6115–6121.

38 Woodfield GW, Chen Y, Bair TB, Domann FE, Weigel RJ. Identification of primary gene targets of TFAP2C in hormone responsive breast carcinoma cells. Genes Chromosomes Cancer 2010; 49: 948–962.

39 Williams CM, Scibetta AG, Friedrich JK, Canosa M, Berlato C, Moss CH et al. AP-2gamma promotes proliferation in breast tumour cells by direct repression of the CDKN1A gene. EMBO J 2009; 28: 3591–3601.

40 Delacroix L, Begon D, Chatel G, Jackers P, Winkler R. Distal ERBB2 promoter fragment displays specific transcriptional and nuclear binding activities in ERBB2 overexpressing breast cancer cells. DNA Cell Biol 2005; 24:582–594.

41 Weinstein IB. Cancer. Addiction to oncogenes--the Achilles heal of cancer. Science 2002; 297: 63–64.

42 Scaltriti M, Eichhorn PJ, Cortes J, Prudkin L, Aura C, Jimenez J et al. Cyclin E amplification/overexpression is a mechanism of trastuzumab resistance in HER2þ breast cancer patients. Proc Natl Acad Sci USA 2011; 108:3761–3766.

43 O'Brien NA, Browne BC, Chow L, Wang Y, Ginther C, Arboleda J et al. Activated phosphoinositide 3-kinase/AKT signaling confers resistance to trastuzumab but not lapatinib. Mol Cancer Ther 2010; 9: 1489–1502.

44 Scaltriti M, Rojo F, Ocana A, Anido J, Guzman M, Cortes J et al. Expression of p95HER2, a truncated form of the HER2 receptor, and response to anti-HER2 therapies in breast cancer. J Natl Cancer Inst 2007; 99:628–638.

45 Eichhorn PJ, Gili M, Scaltriti M, Serra V, Guzman M, Nijkamp W et al. Phosphatidylinositol 3-kinase hyperactivation results in lapatinib resistance that is reversed by the mTOR/phosphatidylinositol 3-kinase inhibitor NVP-BEZ235. Cancer Res 2008; 68: 9221–9230.

46 Harris LN, You F, Schnitt SJ, Witkiewicz A, Lu X, Sgroi D et al. Predictors of resistance to preoperative trastuzumab and vinorelbine for HER2-positive early breast cancer. Clin Cancer Res 2007; 13: 1198–1207.

47 Dave B, Migliaccio I, Gutierrez MC, Wu MF, Chamness GC, Wong H et al. Loss of phosphatase and tensin homolog or phosphoinositol-3 kinase activation and response to trastuzumab or lapatinib in human epidermal growth factor receptor 2-overexpressing locally advanced breast cancers. J Clin Oncol 2011; 29: 166–173.

48 Baselga J, Gelmon KA, Verma S, Wardley A, Conte P, Miles D et al. Phase II trial of pertuzumab and trastuzumab in patients with human epidermal growth factor receptor 2-positive metastatic breast cancer that progressed during prior trastuzumab therapy. J Clin Oncol 2010; 28: 1138–1144.

49 Marchio C, Natrajan R, Shiu KK, Lambros MB, Rodriguez-Pinilla SM, Tan DS et al. The genomic profile of HER2-amplified breast cancers: the influence of ER status. J Pathol 2008; 216: 399–407.

50 Pietersen AM, Horlings HM, Hauptmann M, Langerod A, Ajouaou A, Cornelissen-Steijger P et al. EZH2 and BMI1 inversely correlate with prognosis and TP53 mutation in breast cancer. Breast Cancer Res 2008; 10: R109.

51 Hollestelle A, Nagel JH, Smid M, Lam S, Elstrodt F, Wasielewski M et al. Distinctgene mutation profiles among luminal-type and basal-type breast cancer cell lines. Breast Cancer Res Treat 2010; 121: 53–64.

52 Brough R, Frankum JR, Sims D, Mackay A, Mendes-Pereira AM, Bajrami I et al. Functional viability profiles of breast cancer. Cancer Discov 2011; 1: 260–273.

53 Farmer H, McCabe N, Lord CJ, Tutt AN, Johnson DA, Richardson TB et al. Targeting the DNA repair defect in BRCA mutant cells as a therapeutic strategy. Nature 2005; 434: 917–921.

54 Lambros MB, Simpson PT, Jones C, Natrajan R, Westbury C, Steele D et al. Unlocking pathology archives for molecular genetic studies: a reliable method to generate probes for chromogenic and fluorescent in situ hybridization. Lab Invest 2006; 86: 398–408.

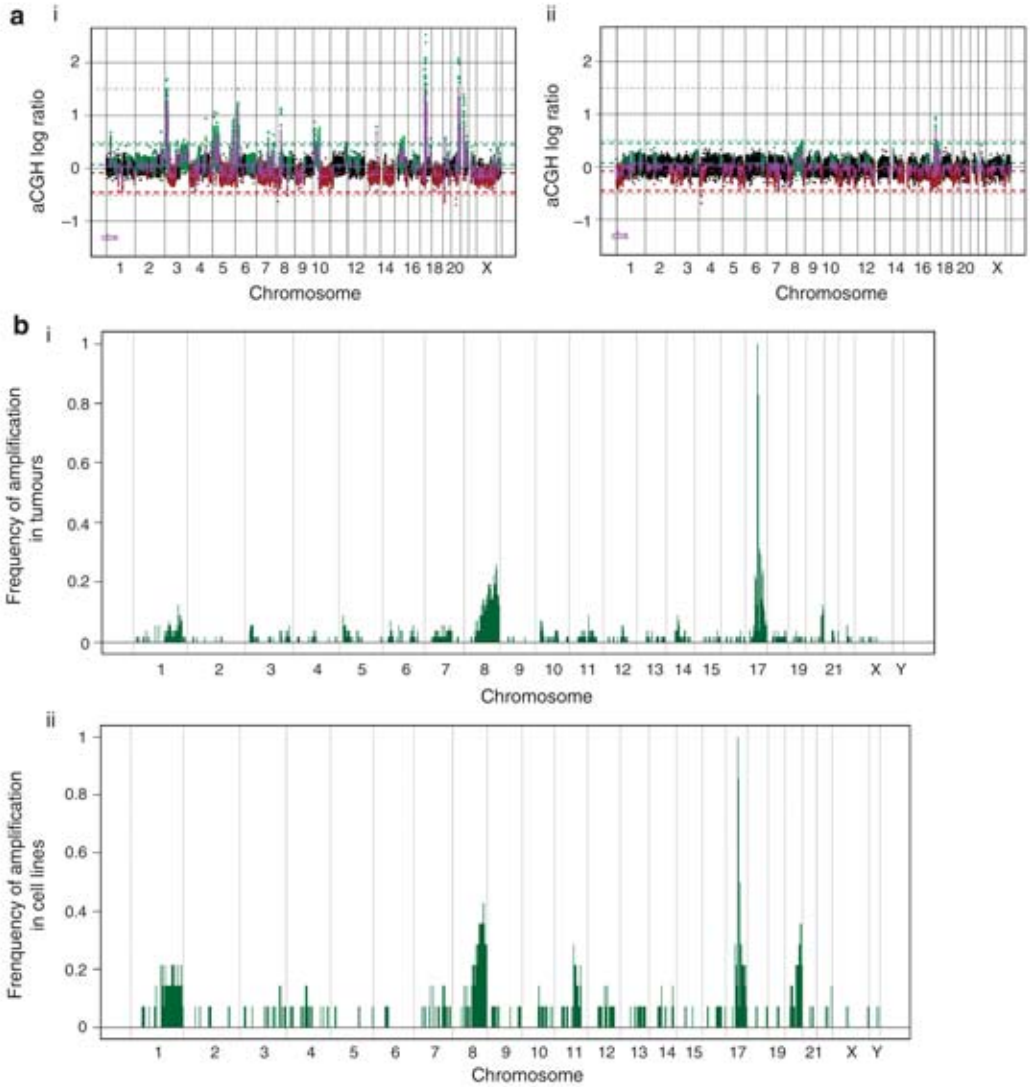

Figure 1. *ERBB2* amplified tumours have heterogeneous genome profiles and harbour multiple amplicons. (**a**) Representative genome plots of two tumours using 32K bacterial artificial chromosome (BAC) aCGH platform (i) 'Firestorm profile' with 13 amplicons and (ii) 'Sawtooth profile' with one amplicon and multiple intrachromosomal gains and losses. The x axis identifies chromosome position, y axis the circular binary segmentation (cbs) smoothed $\log_2$ ratio for each BAC clone. A threshold of of 0.45 was set for amplification as previously described.[23] (**b**) Karyogram showing (i) 122 recurrent amplified loci in the 58 primary *ERBB2*-amplified breast cancers and (ii) 67 recurrently amplified loci in 14 *ERBB2* amplified breast cancer cell lines. Proportion of tumours in which each clone is amplified (green bars) is plotted (y axis) for each BAC clone according to genomic location (x axis).

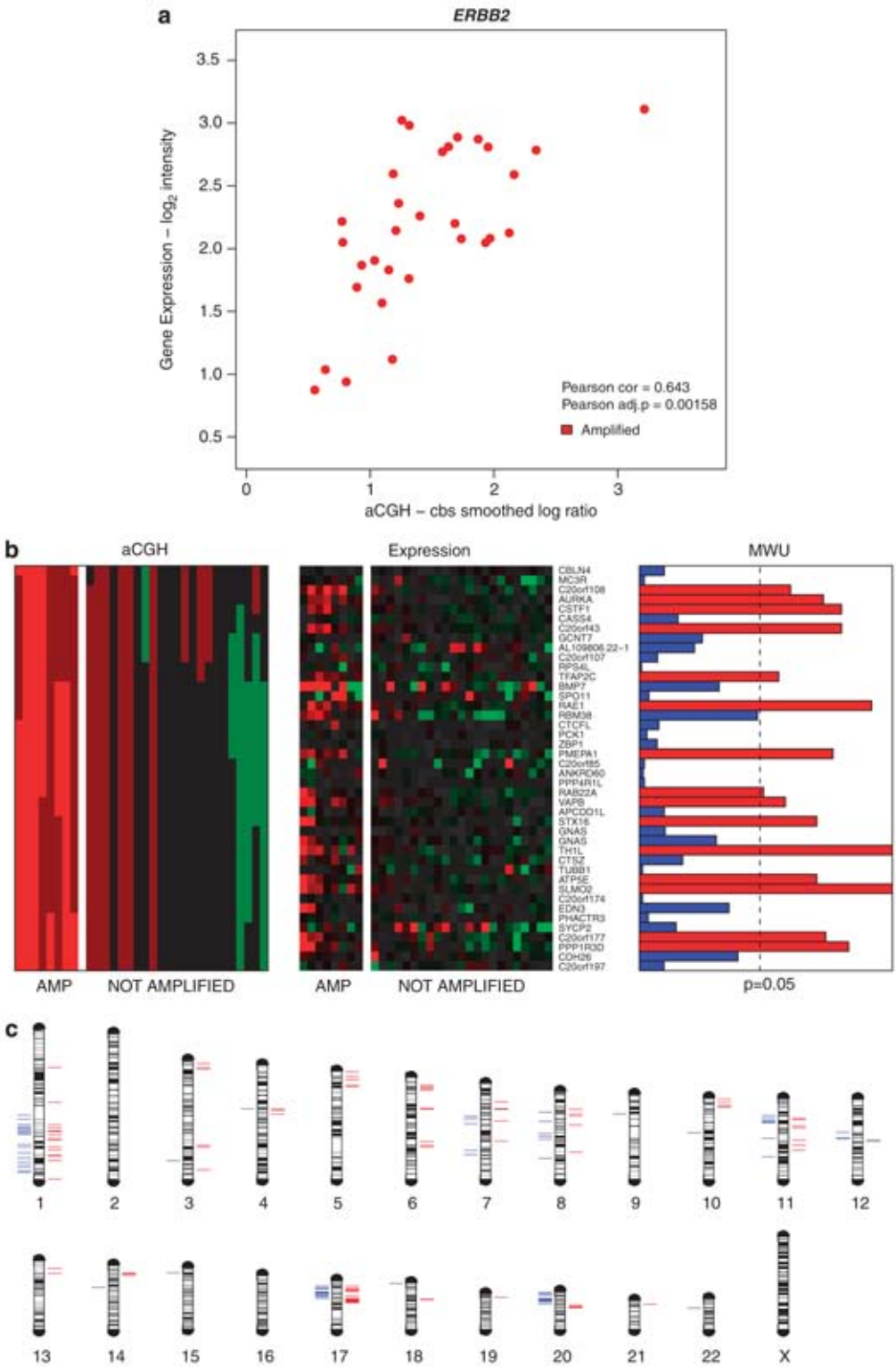

Figure 2. Integration of copy number and expression arrays and mapping to recurrent cell line amplicons. (**a**) Example of a significant Pearson correlation of increasing ERBB2 copy number with increasing gene expression (normalized to median array expression Illumina Ref 6 version 2 chip). (**b**) Example of a matched heatmap of expression and aCGH within the recurrent amplicon of 20q13.2-q13.33. aCGH and gene expression heatmaps for each amplicon are shown with genes mapping to the region. The bar on the right represents the result of the Mann–Whitney *U* (MWU) test for amplified vs un-amplified cases for each gene. The vertical interrupted line represents a MWU *P* value < 0.05. Bars are coloured in red if significantly correlated, and in blue if not. aCGH: Green: copy number loss, Black: no copy number change, Dark red: copy number gain, bright red: gene amplification; gene expression: green: downregulation, red: upregulation. (**c**) Karyogram showing location of 93 recurrent amplified loci in the 45 *ERBB2*-amplified breast cancers (red bars) and 67 recurrently amplified loci in the 16 breast cancer cell lines used in the primary siRNA screen.

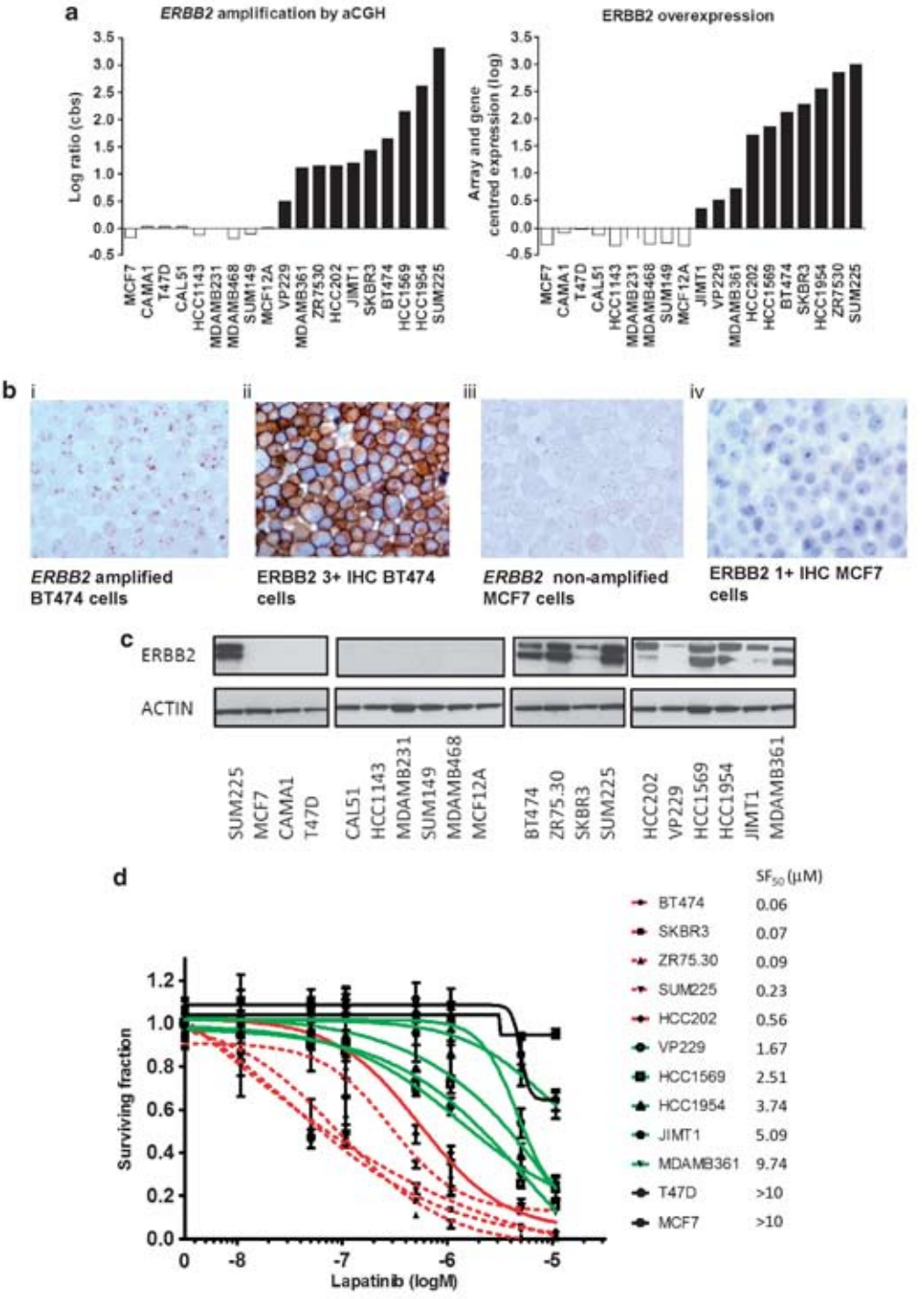

Figure 3. Characterization of the breast cancer cell lines. (**a**) Matched aCGH and mRNA expression profiles of *ERBB2* in 10 *ERBB2* amplified cell lines (black) and 9 non-ERBB2 amplified cell lines (white). *ERBB2* gene copy number was categorized as amplified if the circular binary segmentation (cbs)-smoothed $\log_2$ ratios were >0.45. The transcript was defined as overexpressed if the mRNA expression $\log_2$ intensity value was greater than median-scaled expression following variance-stabilizing transformation and robust spine normalization of the raw expression data using the *lumi* package in the Bioconductor software.[52] (**b**) Cell line tissue microarray showing: (i) *ERBB2* CISH amplified BT474 cells; (ii) ERBB2 3+ immunohistochemistry (IHC) in BT474 cells; (iii) *ERBB2* CISH non-amplified MCF7 cells; and (iv) ERBB2 − ve IHC in MCF7 cells. (**c**) Western blot showing ERBB2 expression in 20 breast cancer cell lines. (**d**) Representative drug curves of cell lines treated with lapatinib. Dotted red = *ERBB2*-amplified lapatinib sensitive cell lines, Solid red = *ERBB2*-amplified Lapatinib resistant and Black = Non-*ERBB2*-amplified cell lines.

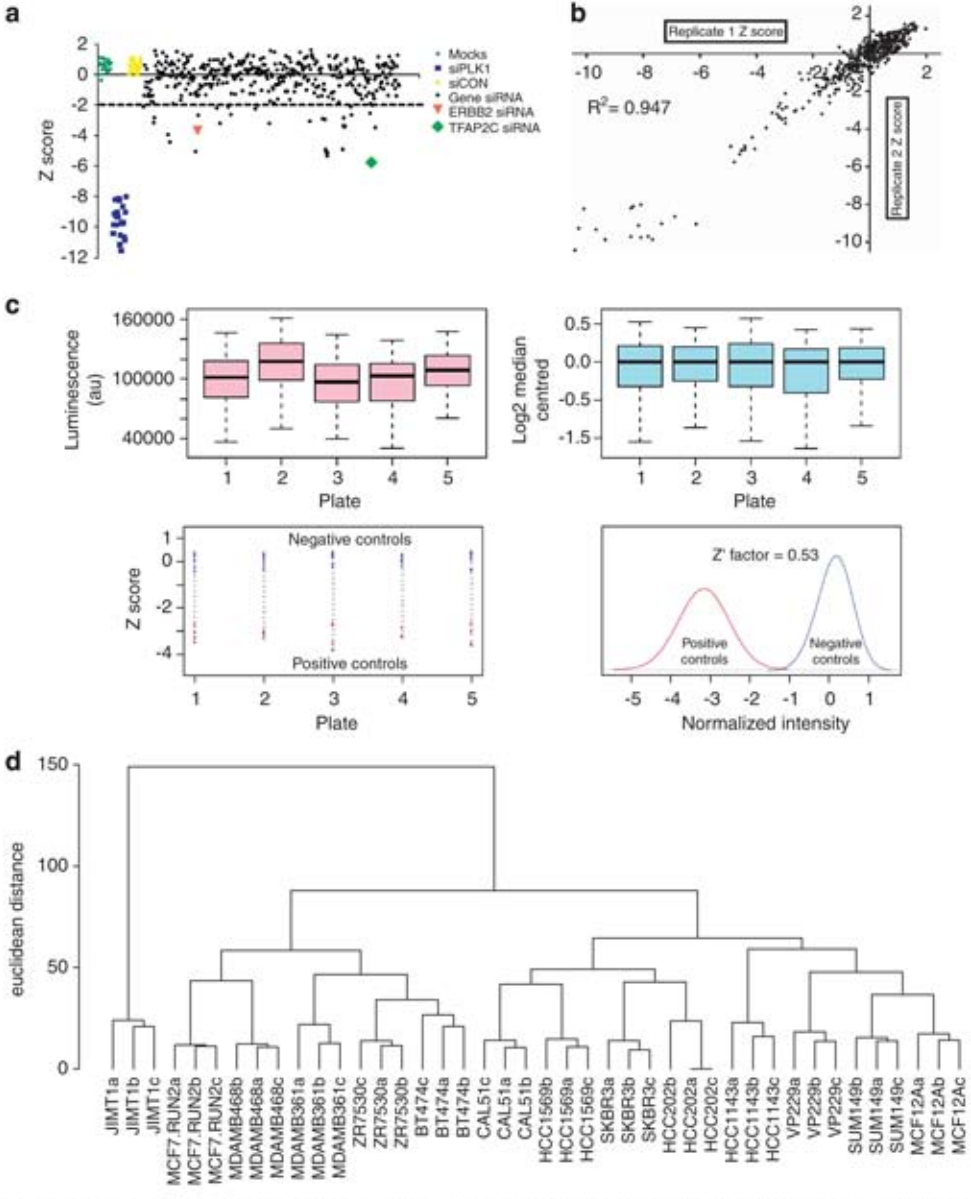

Figure 4. Analysis and quality control of parallel siRNA screens in multiple cell lines. (**a**) Representative scatter plot of *Z*-scores from cell viability screen of SKBR3. Dots represent individual siRNA targeting the cell line. (**b**) Each screen was done in triplicate and *Z*-scores were plotted against each replicate. Replicates where $r^2 < 0.7$ were discarded and/or screen repeated. (**c**) Raw values from Cell titre glo assays were also normalized using cellHTS in the ROCK database. Significant separation of the negative (siCON) and positive controls (siPLK1) *Z*-score values was assessed by *z'*. (**d**) Unsupervised clustering of the *Z*-score in each cell line screen replicate.

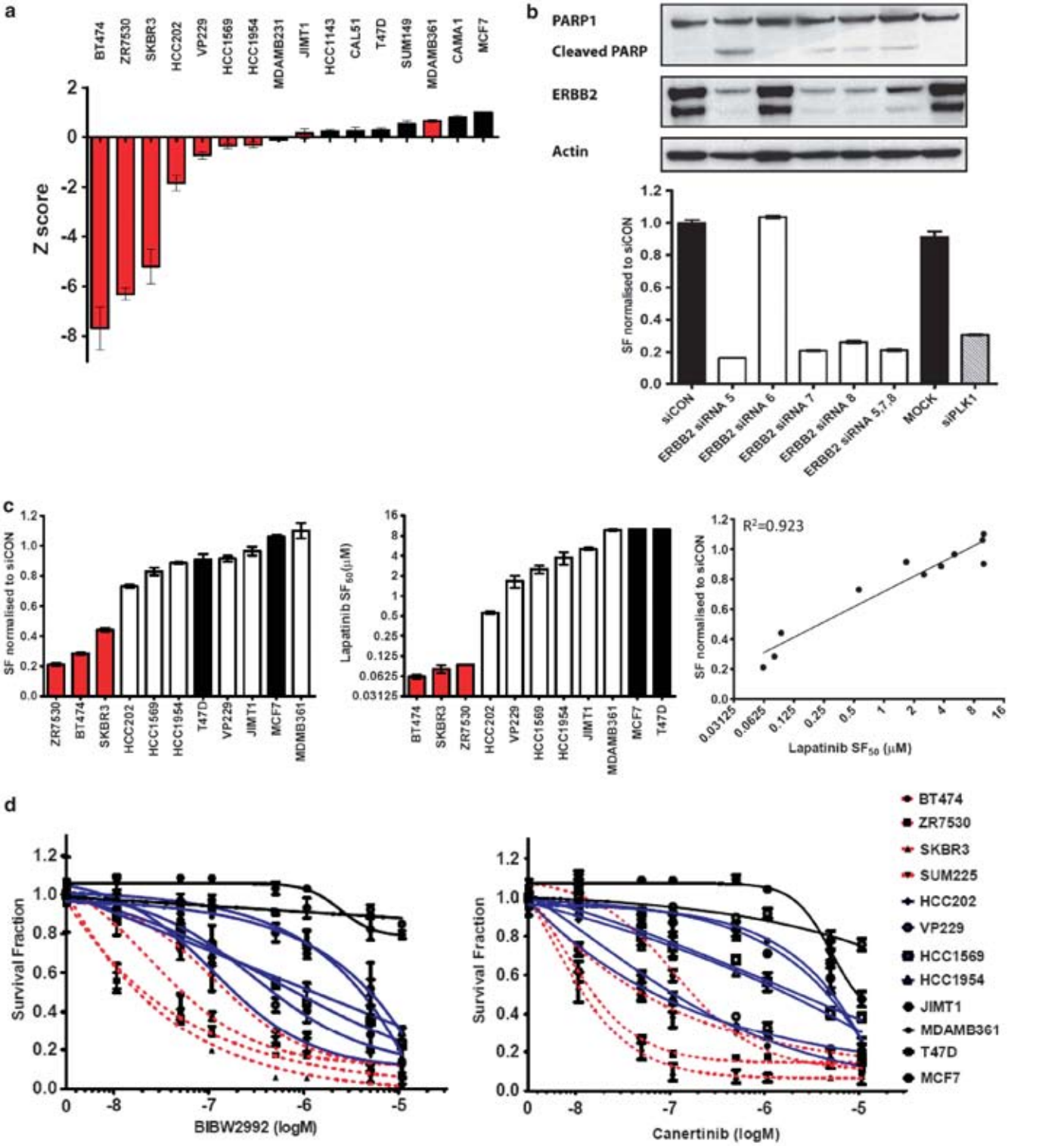

Figure 5. Validation of screen with ERBB2 siRNA. (**a**) *Z*-scores are plotted against each cell line with mean and s.e. values in triplicate experiments. *Z*-score of < -2 represents significant loss of viability effect. Lapatinib-sensitive *ERBB2*-amplified cell lines are blocked in red, lapatinib-resistant *ERBB2*-amplified cell lines are blocked in green, and non-*ERBB2*-amplified cell lines are blocked in purple. (**b**) Expression of ERBB2 and PARP1 cleavage protein following transfection of siRNA to ERBB2 in ZR7530 cells, as well as (siCON) and Mock. (**c**) Survival fraction normalized to siCON of each cell line following silencing with ERBB2 smart pool siRNA is shown, mean and s.e. in triplicate experiments. SF_{50}(μM) to lapatinib for each cell line is also shown. Survival fraction for each cell lines is plotted against lapatinib IC_{50} (x axis $\log_2$) showing mean and s.e. in sextuplet experiments. (**d**) A panel of cell lines were treated with a range of drug concentrations to ascertain their SF_{50}s of BIBW2992 and canertinib. Loss of viability was measured using CellTiter-Glo cell viability assay. Values represent mean and s.e. of sextuplet replicates.

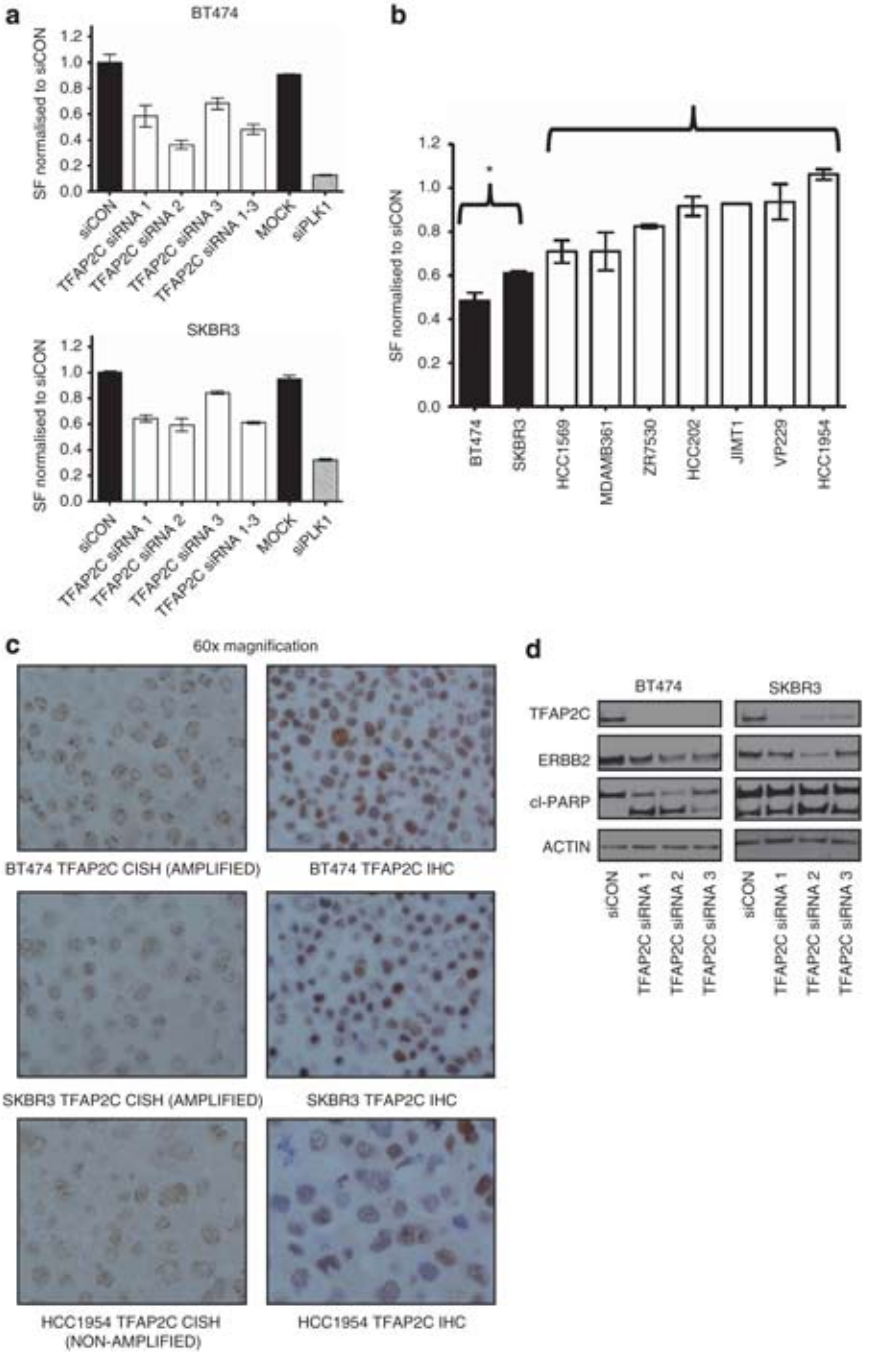

Figure 6. *TFAP2C* is novel target for a subgroup of *ERBB2*-amplified cancers and a potential predictive marker for lapatinib sensitivity. (**a**) Survival fractions following transfection of siRNA pool of TFAP2C normalized to siCON in a panel of *ERBB2* amplified cell lines. (**b**) Cell viability effect of transfecting siRNA pools of TFAP2C in three subgroups of cell lines. Mean and s.e. bars are shown of triplicate experiments are shown. Survival fractions are significantly different between *TFAP2C* amplified cells (black bars) vs non amplified cells for *TFAP2C*. *unpaired *t*-test $P = 0.014$ (two sided). (**c**) TFAP2C copy number and expression (i) BT474, (ii) SKBR3, (iii) HCC1954 on cell line microarray as assessed by CISH and immunohistochemistry (IHC). (**d**) Expression of TFAP2C, ERBB2, total and cleaved PARP1 expression 72 h following transfection of siRNA to TFAP2C vs siCON in BT474 and SKBR3 cells.

Table 1. Correlation of significant loss of viability after gene silencing by siRNA in recurrently amplified genes

	MCF12A	*MCF7*	*HCC1143*	*CAL51*	*MDAMB468*	*SUM149*	*BT474*	*SKBR3*	*ZR7530*	*HCC202*	*VP229*	*HCC1569*	*JIMT1*	*MDAMB361*
ERBB2	0.50	−0.69	−0.17	−0.74	0.11	0.36	**−5.86**	**−3.30**	**−5.90**	−1.88	−1.33	−0.67	0.21	0.58
STARD3	−1.00	−1.03	−2.87	0.67	−1.01	−1.60	**−2.47**	0.03	−0.37	−0.96	−0.88	0.48	**−2.11**	**−2.25**
PSMD3	−9.35	−2.06	−10.15	−8.32	−4.19	−5.45	−1.08	**−5.15**	−3.64	−7.84	**−10.22**	−7.23	**−7.86**	−3.98
TATDN1	**−0.38**	−1.35	**−2.25**	−2.19	−3.16	−0.31	−0.63	**−2.43**	**−2.27**	−3.78	−0.02	−2.48	−0.36	0.98
TFAP2C	−0.32	**−2.21**	−0.82	−2.63	−1.38	−1.58	**−3.73**	**−5.17**	−1.62	−1.51	0.08	−3.43	−1.54	−1.95
ANO1	−0.51	−1.89	**−5.86**	−4.97	−3.44	−2.87	−2.32	−3.30	−0.59	−0.36	−3.23	−1.42	−4.01	**−2.16**
AURKA	−2.88	**−1.19**	−1.18	−4.95	−4.31	−4.07	**−3.62**	**−3.46**	−5.01	−5.18	−5.96	−1.92	**−12.73**	−2.26
NUF2	−2.34	−1.37	−1.67	−2.60	−3.37	−3.31	**−3.37**	−3.72	**−2.18**	−3.27	−2.57	−3.05	−16.08	−2.11
SUPT6H	−7.33	−2.46	−4.86	−4.97	−2.40	−4.90	−1.92	−0.94	**−3.01**	**−0.77**	**−2.21**	−2.97	−6.56	−0.88

Abbreviation: SiRNA, short interfering RNA. Amplified genes are highlighted in red. *Z*-scores <2 in amplified genes are in bold.