Functional Viability Profiles of Breast Cancer

Rachel Brough*  
Jessica R. Frankum*  
David Sims  
Alan Mackay  
Ana M. Mendes-Pereira  
Ilirjana Bajrami  
Sara Costa-Cabral  
Rumana Rafiq  
Amar S. Ahmad  
Maria Antonietta Cerone  
Rachael Natrajan  
Rachel Sharpe  
Kai-Keen Shiu  
Daniel Wetterskog  
Konstantine J. Dedes  
Maryou B. Lambros  
Teeara Rawjee  
Spiros Linardopoulos  
Jorge S. Reis-Filho  
Nicholas C. Turner  
Christopher J. Lord  
Alan Ashworth
**ABSTRACT**

The design of targeted therapeutic strategies for cancer has largely been driven by the identification of tumor-specific genetic changes. However, the large number of genetic alterations present in tumor cells means that it is difficult to discriminate between genes that are critical for maintaining the disease state and those that are merely coincidental. Even when critical genes can be identified, directly targeting these is often challenging, meaning that alternative strategies such as exploiting synthetic lethality may be beneficial. To address these issues, we have carried out a functional genetic screen in >30 commonly used models of breast cancer to identify genes critical to the growth of specific breast cancer subtypes. In particular, we describe potential new therapeutic targets for PTEN-mutated cancers and for estrogen receptor-positive breast cancers. We also show that large-scale functional profiling allows the classification of breast cancers into subgroups distinct from established subtypes.

**SIGNIFICANCE:** Despite the wealth of molecular profiling data that describe breast tumors and breast tumor cell models, our understanding of the fundamental genetic dependencies in this disease is relatively poor. Using high-throughput RNA interference screening of a series of pharmacologically tractable genes, we have generated comprehensive functional viability profiles for a wide panel of commonly used breast tumor cell models. Analysis of these profiles identifies a series of novel genetic dependencies, including that of PTEN-null breast tumor cells upon mitotic checkpoint kinases, and provides a framework upon which additional dependencies and candidate therapeutic targets may be identified. *Cancer Discovery; 1(3); 260–73. ©2011 AACR.*

**INTRODUCTION**

Central to the design of novel therapeutic strategies for cancer is the identification of genes that are critical to the survival of tumor cells but largely redundant in normal cells. Correlating molecular changes with tumorigenesis has offered one route to the identification of potential drug targets and provides the rationale behind efforts to characterize genetic variation and gene expression in tumors. However, the correlative nature of these data means that it is frequently not possible to determine whether the observations are causative or merely an effect of the disease state. For example, tumor cells commonly exhibit anywhere between 10^3 and 10^5 genetic changes, compared with germline DNA, but theoretical estimates suggest that only a few of these mutations (probably <10) are critical drivers of tumor formation and survival (reviewed in ref. 1). In breast cancer, defining the critical genes involved in tumor cell survival can, in some cases, lead to the development of novel therapeutic approaches to the disease, the most notable recent example being the development of agents such as trastuzumab and lapatinib, which target the reliance of some breast tumors upon the oncogene *ERBB2* (2). Despite the wealth of molecular, genetic, and histologic information on breast tumors and cell-line models, our understanding of the genetic dependencies in this disease is relatively poor.

RNA interference (RNAi) screening has already enabled the identification of genetic dependencies in cancer cells (1, 3) but has not as yet been applied to a comprehensive study of breast cancer models. In this article, we describe the first attempt to define in detail genetic dependencies for a set of pharmacologically tractable genes in a wide range of breast tumor cell-line models, using a library of siRNA targeting the kinome. In doing so, we not only reaffirm the impact of PI3 kinase and ERBB2 signaling in the disease but, importantly, also show that combining functional RNAi analysis with gene expression, gene mutation, and genomic analysis provides a new strategy for identifying essential determinants of specific breast cancer subtypes, which are potential novel drug targets.

**RESULTS**

**Generation of Functional Viability Profiles of Breast Cancer**

To generate functional profiles of breast cancer, we used an approach that involves high-throughput RNA interference/siRNA viability screening of multiple cell lines and the integration of viability profiles with gene expression, genetic, genomic, and histologic analysis (3). We reasoned that siRNAs causing significant loss of cell viability in all cell lines assayed probably targeted genes that had an essential ubiquitous function in both normal and tumor cells. Similarly, siRNAs with no major effect on viability in any of the cell
lines either were not functional or targeted nonessential genes. Finally, siRNAs causing significant lethality in some, but not all, cell lines likely identified genes that represented tumor-specific dependencies and candidate therapeutic targets (Supplementary Fig. S1A).

To generate functional viability profiles for breast cancer, we selected a panel of 34 breast cancer–derived cell lines and optimized these for high-throughput siRNA screening. Subsequently, each cell line was transfected with a 96-well plate-arrayed siRNA library targeting 714 kinases and kinase-related genes (see Methods; summarized in Fig. 1 and Supplementary Fig. S1B). After 5 population doublings, cell viability in each well was estimated with a highly sensitive luminescent assay measuring cellular ATP levels. To identify loss of viability, inhibition, or failure to proliferate effects in each cell line, luminescence readings from each well were log transformed and then centered by the plate median, to account for plate-to-plate variation (Fig. 1; Supplementary Fig. S1B). To allow data to be compared between different cell lines, plate-centered data from each screen were standardized by the use of a Z-score statistic, where $Z = 0$ represents no effect on viability and negative Z-scores represent loss of viability. RNA interference screens were carried out in triplicate and comparison of Z-score data from replica screens of each cell line showed the screening process to be highly robust (Supplementary Fig. S1C–E; Supplementary Table S1).

Having screened and processed data from 34 breast cancer lines, we selected 20 for further analysis, based on stringent QC thresholds (such as Z prime factors; Supplementary Table S2). This panel encapsulated each of the major breast cancer subtypes (4) and included hormone receptor [estrogen receptor (ER), progesterone receptor (PR)] and ERBB2-positive as well as ER-, PR-, and ERBB2-negative models (triple-negative), as characterized by parallel immunohistochemical, fluorescence in situ hybridization (FISH), and Western blot analysis (Supplementary Table S3; Supplementary Fig. S1F). To complement this functional viability profiling, we also characterized the breast tumor panel using transcript microarrays, array-based comparative genomic hybridization (aCGH), drug sensitivity for a small number of targeted agents, and

**Figure 1.** Functional profiling of a breast cancer tumor cell-line panel. Thirty-four breast cancer cell lines were selected for study and characterized at the molecular level, using techniques such as aCGH, transcript microarray, and immunohistochemistry. Functional profiling of the panel was performed using an siRNA viability screen targeting 714 kinases and kinase-related genes. Each cell line was screened in triplicate. After processing data (see Methods; Supplementary Fig. S1A), the quality of each screen was assessed, and data from 20 breast cancer lines were selected for subsequent interrogation.
gene mutation data (summarized in Supplementary Tables S3–S8). In the siRNA screens we used a stringent threshold of $Z < -2$ for defining loss of viability effects ($Q = 0.02275$ in 1 screen replica, and $Q = 1.21 \times 10^{-3}$ in triplicate screens; ref. 5) and identified 330 genes whose depletion by siRNA caused loss of viability in $\geq 1$ cell line, and 180 genes that caused loss of viability in $\geq 2$ cell lines. This analysis enabled us to identify the predominant dependencies in each cell line (Supplementary Tables S3 and S4). All 20 cell lines were reliant upon Polo-Like Kinase 1 (PLK1), thus providing an ideal internal control to monitor successful transfection. Furthermore, significant fractions of the breast tumor cell panel were sensitive to silencing of the mitotic kinases AURKA (19 of 20; 95%) and WEE1 (15 of 20; 75%).

**Identification of De Facto Genetic Dependencies Using Functional Viability Profiling**

The delineation of functional viability profiles for the breast cell-line panel provided a framework for identifying dependencies in particular genetically defined subgroups of the disease. As proof of this principle, we used supervised hierarchical clustering of the viability data and integration with DNA sequence, aCGH, FISH, and Western blot data to identify additional gene dependencies in PIK3CA mutant or ERBB2-amplified cell lines (summarized in Supplementary Tables S5 and S6). Somatic mutations of PIK3CA, which encodes the p110α catalytic subunit of PI3 kinase, have been shown to induce oncogenic transformation in vitro and in vivo (6). PIK3CA mutations are found in 8% to 35% of human breast cancers, making them one of the most common genetic aberrations in this disease (7). Supervised hierarchical clustering of functional viability data according to PIK3CA mutation status identified 30 siRNAs that caused loss of viability preferentially in the PIK3CA mutant subgroup (Fig. 2A; Supplementary Fig. S2A). Of 714 genes profiled in aCGH, aCGH, FISH, and Western blot data to identify additional gene dependencies in PIK3CA mutant or ERBB2-amplified cell lines (summarized in Supplementary Tables S3 and S6) identified a number of genes that, when targeted, were selectively lethal in ERBB2-amplified breast cancer lines (Fig. 3a; Supplementary Fig. S2E). One of the most dominant effects was targeting of ERBB2 itself ($Q = 0.003$), supporting the hypothesis that ERBB2 amplification can lead to oncogene addiction (Fig. 3a and b; Supplementary Fig. S2E–G). This finding reaffirms that our approach may have merit in identifying genes functionally relevant to tumor cell viability in clinically defined subgroups. Moreover, many ERBB2-amplified cell lines were highly sensitive to PIK3CA siRNA (Fig. 3a and c), supporting the clinical development of PI3K inhibitors in ERBB2-dependent breast cancer regardless of PIK3CA mutation status.

In associated work examining gene addiction effects in melanoma, we performed a similar analysis on COLO-829 cells, a melanoma cell line whose entire genome sequence has recently been determined (11). Deep sequencing of COLO-829 and a normal B-cell line from the same patient identified $>33,000$ tumor-associated mutations, including a $BRAF\ V600E$ mutation recurrently found in melanoma (11). Using siRNA viability profiling of COLO-829, we showed that of 779 genes targeted, siRNA targeting of $BRAF\ had$ the most significant inhibitory effect (Supplementary Fig. S3A), supporting the hypothesis that functional viability profiling of tumor cell lines has the ability to identify real addiction effects. Similarly, functional viability profiling of the $BRAF\ V600E$ mutant breast cancer cell line MDAMB231 also identified the addiction to $BRAF$ (Supplementary Table S3; Supplementary Fig. S3B).

**Candidate Genetic Dependencies for Breast Cancer Subgroups**

On the basis of these proof-of-concept examples, we assessed whether functional viability profiles could identify novel addictions or candidate therapeutic targets for specific genetic or phenotypic backgrounds. To do this, we annotated the viability dataset with gene mutation, transcript expression, and phenotypic data and then used hierarchical clustering of viability data to identify candidate gene dependencies in different genetic contexts. For example, integrating functional viability data with genomic profiles generated
Figure 2. PIK3CA oncogene addiction effects in breast cancer confirmed by functional viability profiling. A, heat map showing the results of a supervised clustering of siRNA Z-scores. Breast tumor cell lines were clustered according to PIK3CA gene mutation status and differential effects between PIK3CA mutant (MUT) and wild-type (WT) groups identified using the median permutation test. Statistically significant effects (P < 0.05) are shown. siRNA targeting PIK3CA is marked by the arrow. B, waterfall and box/whiskers plots of PIK3CA siRNA Z-scores across the breast tumor cell-line panel. In the box/whiskers plot, *P = 0.014 between PIK3CA mutant and wild-type groups, using the median permutation test. C, box/whiskers plots of AKT2 and AKT3 siRNA Z-scores in PIK3CA mutant and wild-type models (*, P = 0.047 and ***, P = 0.003, respectively, using the median permutation test). D, box/whiskers plots of individual and pooled PIK3CA siRNA in breast tumor cell lines with kinase catalytic domain or noncatalytic domain mutations (where *P < 0.05 when comparing survival of *helical mutant and wild type, **kinase mutant and wild type, and ***helical mutant and kinase mutant, using a Student t test). NPI, normalized percent inhibition compared to nonsilencing control siRNA and siRNA targeting PLK1.
**Figure 3.** Candidate genetic dependencies for breast cancer subtypes. 

**A,** heat map showing the results of a supervised clustering of siRNA Z-scores according to ERBB2 amplification status in the breast tumor cell-line panel. Statistically significant effects (P < 0.05, median permutation test) are shown. ERBB2 and PIK3CA siRNA are marked by arrows. 

**B,** waterfall and box/whiskers plots of ERBB2 siRNA Z-scores across the breast tumor cell-line panel. In the box/whiskers plot, *p* = 0.003 between ERBB2-amplified and nonamplified groups, using the median permutation test. 

**C,** waterfall plots of kinases that selectively kill breast cancer cell lines containing an ERBB2 amplification. 

**D,** TN/ basal breast cancer cell lines. 

**E,** luminal breast cancer cell lines. 

P < 0.05 for A–C, using a Student t test.
by aCGH analysis enabled the identification of candidate genetic dependencies for breast tumor cells carrying amplification events commonly found in breast cancer (12, 13). As with the ERBB2 paradigm, we identified candidate genetic dependencies associated with chromosome 11q13.2-q13.3 amplification; the chromosome 8q23.3-q24.3 amplification encompassing the MYC oncogene; amplification of FGFR1 on chromosome 8p12; amplification of AURKA on chromosome 20q13; amplification of CDK4 on chromosome 12q14.1; the chromosome 17q21-q23 amplification encompassing the PPM1D oncogene; and amplification of MDM2 on chromosome 12q14.3-q15 (Supplementary Fig. S4A–G). A similar analysis allowed us to identify candidate dependencies in models with common tumor suppressor or oncogene mutation events, such as those for PTEN mutation, CDKN2A (p16) mutation, KRAS mutation, p53 mutation, BRCA1 mutation, or RB mutation (Supplementary Fig. S5A–F). Finally, we compared functional viability profiles with transcript microarray–defined subtypes of breast cancer (e.g., ERBB2, basal/triple-negative, and luminal status; ref. 14) to identify candidate dependencies in each group (Fig. 3C–E). Further candidate dependencies identified in the ERBB2-amplified models included the calcium/calmodulin-dependent protein kinase CAMK1 and a member of the mitogen-activated protein kinase family, MAP2K3 (Fig 3C). Key potential targets for the triple-negative (negative for ER, PR, and ERBB2 amplification) subgroup included the ribosomal S6 kinase, RPS6KA3; the protein kinase C–related protein kinase 2, PRKCL2; and the pyruvate kinase PKLR (Fig 3D). Finally, candidate luminal subtype dependencies included PIK3CA, the casein kinase CSNK1E, and the serine/threonine kinase MINK (Fig. 3E).

**PTEN-Deficient Breast Tumor Cells Are Dependent upon the Mitotic Checkpoint Kinase TTK**

To validate our approach, we selected 2 genetic dependencies for more detailed analysis. By comparing the candidate genetic dependencies of PTEN-deficient models with those having wild-type PTEN expression (Supplementary Fig. S6A and B; Supplementary Table S7), we observed the apparent dependency of PTEN-deficient breast cancer lines for a series of genes controlling the mitotic spindle assembly checkpoint (Supplementary Fig. S6C). The most significant dependency among these was that on the TTK protein kinase gene (also known as MPS1; \(P = 0.0012\); Fig. 4A and B; Supplementary Fig. S6D and E; ref. 15). To directly address the hypothesis that TTK inhibition was selective for PTEN mutant cancer cells, we used isogenic cancer cell lines in which both copies of the *PTEN* gene had been rendered dysfunctional by gene targeting (16). We showed that multiple siRNAs silencing TTK (Fig. 4C and D; Supplementary Fig. S6F) or chemical inhibition of TTK (using 2 distinct TTK small molecule inhibitors; ref. 15) was selective for PTEN deficiency (Fig. 4E and F), confirming observations made in the more genetically diverse breast cancer line panel. This finding indicated that TTK inhibition may be a novel therapeutic strategy for treating PTEN mutant tumors. Aneuploidy is frequently observed in both human breast carcinomas with low expression of PTEN and prostatic intraepithelial neoplasia from *Ptten* mutant mice (17). The latter phenomenon is perhaps explained by the centromeric dysfunction in PTEN mutant tumor cells, most likely mediated by a loss of the interaction between PTEN and CENP-C, a key kinetochore component (18). TTK is required for normal function of the mitotic spindle checkpoint, and it is established that TTK inhibition drives early exit from mitosis and chromosomal aneuploidy. In tumor cells with an aneuploid phenotype, TTK inhibition further exacerbates aneuploidy and is particularly lethal (19). By examining metaphases from PTEN-deficient tumor cells exposed to a small-molecule TTK inhibitor, we showed that the frequency of abnormal metaphases was substantially increased in PTEN-deficient cells (Supplementary Fig. S6G and H), which may explain the PTEN-deficient–selective effect of inhibiting this mitotic checkpoint kinase.

**Genetic Dependency on Uncharacterized aarF Domain–Containing Protein Kinase 2 in ER+ Breast Tumor Models**

As a second example of the potential of functional viability profiling, we interrogated our datasets to identify candidate genes whose targeting was selective for ER+ versus ER- subtypes. The most significant effect we identified was the sensitivity of ER+ breast tumor cell-line models to ADCK2 siRNA (\(P = 0.004\); Fig. 5A and B; Supplementary Fig. S7A and B), an effect validated in an additional 4 ER+ breast tumor cell-line models not examined in the original screen (Supplementary Fig. S7C and D). This observation suggested that the genetic dependency of ER+ breast tumor cells on uncharacterized aarF domain–containing protein kinase 2 (ADCK2) is a relatively common phenomenon. Integration of viability data with transcript and protein profiling also identified a correlation between sensitivity to ADCK2 silencing and ADCK2 mRNA/protein levels (Supplementary Table S4; Supplementary Fig. S7E–G), perhaps indicating that elevated expression of this protein could be essential for the survival of particular tumor cells. ADCK2 is a member of a family of aarF domain–containing proteins. Other members of this protein family are localized to the mitochondria and mitochondrial membranes and have been implicated in ubiquinone biosynthesis (20). Given that the most dominant driver in ER+ breast cancers is estrogen signaling itself, we hypothesized that ADCK2 could target ER+ breast tumor cells by modulating estrogen signaling. In support of this hypothesis, we noted that sensitivity to siRNA targeting of *ESR1*, the gene encoding ERα, closely correlated with the inhibitory effect of ADCK2 siRNA (Supplementary Fig. S7C and H; Supplementary Table S8), indicating that ADCK2 is an important dependency in breast tumor cells with a strong addiction to estrogen signaling. More specifically we noted that ADCK2 silencing reduced estrogen signaling, as measured by the expression of well-established ER target genes, including *ESR1* itself (Fig. 5C); estrogen stimulation stimulated ADCK2 expression; suppression of estrogen signaling (with faslodex/ICI182780) inhibited ADCK2 expression (Fig. 5D); and, finally, ADCK2 coimmunoprecipitated with ERα (Fig. 5E), indicating that the ER and ADCK2 interact, either directly or indirectly. Taken together, these observations suggest that the genetic dependency of ER+ breast tumor cells for ADCK2 could be explained by the role of ADCK2 in estrogen signaling itself. To extend this observation, we examined ADCK2 expression in publicly available breast cancer gene expression datasets. ER+ breast tumors divide into 2 broad...
and overlapping subgroups termed luminal A and B, based predominantly upon proliferative rate and strength of estrogen signaling (luminal A being low proliferative with strong estrogen signaling; ref. 21). We noted in 2 independent studies that ADCK2 expression was elevated in the luminal A subgroup when compared with luminal B (Fig. 5F; P = 0.0135 and 0.0054 in the 2 studies, respectively; refs. 12, 22), in concordance with our in vitro observations.

Figure 4. Inhibition of TTK is PTEN selective. 

A, heat map showing the results of a supervised clustering of siRNA Z-scores. Breast tumor cell lines were clustered according to PTEN gene mutation status and differential effects between PTEN mutant and wild-type groups identified using the median permutation test. Statistically significant effects (P < 0.05) are shown. siRNA targeting TTK is marked by the arrow. B, waterfall and box/whiskers plots of TTK siRNA Z-scores across the breast tumor cell-line panel. In the box/whiskers plot, *P = 0.0012 between PTEN mutant and wild-type groups, using the median permutation test. C and D, surviving fraction data from 2 independently derived isogenic PTEN mutant/wild-type models transfected with siRNA targeting TTK. Surviving fractions were determined 7 days after transfection. *, P < 0.05 (Student t test). E and F, surviving fraction data from 2 independently derived isogenic PTEN mutant/wild-type models exposed to 2 chemically distinct TTK inhibitors (AZ3146 or CCT132774). The surviving fraction was determined after 7 days of exposure to the inhibitor. In each case, the 2-way ANOVA statistic of differences in survival curves is P < 0.0001.)
Figure 5. Genetic dependency of ER+ breast tumor cells upon ADCK2. A, heat map showing the results of a supervised clustering of siRNA Z-scores. Breast tumor cell lines were clustered according to ER status and differential effects between ER+ and ER− groups identified using the median permutation test. Statistically significant effects (P < 0.05) are shown. siRNA targeting ADCK2 is marked by the arrow. B, waterfall and box/whiskers plots of ADCK2 siRNA Z-scores across the breast tumor cell-line panel. In the box/whiskers plot, *P = 0.004 between ER+ and ER− groups, using the median permutation test. C, ADCK2 silencing by multiple ADCK2 siRNA species reduces the expression of ER target genes, as shown by Western blotting of MCF7 total cell lysates generated after siRNA transfection. D, Western blot of total cell lysates from MCF7 cells treated with either estradiol (E2) or faslodex (FAS, ICI182780). An E2-dependent increase in ADCK2 expression is shown. E, Western blot of ERα immunoprecipitated material from MCF7 cells, indicating an ERα–ADCK2 interaction. F, tumor ADCK2 mRNA expression in ER+ breast tumors classified according to luminal A or B status. Transcript expression data from 2 independent studies were used (12, 22) and luminal A or B status determined according to Sorlie et al. (14). P values were calculated using a Student t test.
Functional Viability Profiling Identifies Candidate Functional Taxonomies

Gene expression and genomic profiles have been used to subclassify breast cancers. We investigated whether the patterns of response to gene silencing by siRNA could also be used to define a “functional taxonomy.” Hierarchical cluster analysis of transcript profiles classifies breast tumors and tumor cell lines into luminal and basal-like molecular subtypes (Supplementary Fig. S8A and B). Hierarchical clustering of genes that, when silenced, caused loss of viability ($Z < -2$) in $\geq 2$ cell lines revealed 2 separate groups, distinct from those formed by the clustering of expression data (Fig. 6; Supplementary Fig. S8C and D). Group 1 was enriched for breast cancer lines with PTEN mutations (5 of 5 PTEN mutant lines in group 1; $P = 0.0016$), whereas group 2 was enriched for PIK3CA mutant lines (5 of 7 PIK3CA mutant lines; $P = 0.038$). CAL51 and MDAMB453, which carry both PTEN and PIK3CA mutations, were classified into group 1. Of interest, although the ERBB2-amplified cell lines were distributed evenly between the 2 groups, the cell lines resistant to ERBB2 silencing and also lapatinib treatment (Supplementary Table S6) were all contained within group 1 (JIMT1, MDAMB453, and VP229), and those sensitive to ERBB2 silencing and lapatinib (HCC202, BT474, and SKBR3) fell into group 2 (Fig. 6; Supplementary Fig. S8E). In general, the distinction between groups 1 and 2 implies that our panel of breast cancer cell lines functionally divides into 2 groups according to their dependency on well-established essential cancer networks, and is independent from the currently used clinical and transcriptionally defined breast cancer subgroups.

DISCUSSION

A major limiting factor in cancer drug discovery is identifying targets for specific subtypes of the disease. One important approach to addressing this issue has been to extensively profile the molecular changes that specifically occur in tumor cells. Although this has proven productive and has led, in part, to the development of drugs such as trastuzumab and imatinib, it is often difficult with molecular profiling alone to identify critical proteins that, when targeted, cause inhibition of tumor cell growth or induce tumor cell death. In this article, we show that multidimensional datasets that include molecular profiling data as well as functional viability profiles can be used to identify genetic dependencies and candidate targets in tumor cells. Our reidentification of well-validated targets such as ERBB2 and PIK3CA, using this approach, demonstrates the potential of this method and suggests that examining the clinical potential of other genetic dependencies noted in this paper may be worthwhile.

In this study, we have focused on integrating functional viability profiles for breast tumor cell-line models with accompanying molecular profiling datasets, such as gene mutation status, aCGH, transcriptomic data, and hormone receptor status. An obvious extension of this approach is to apply a similar method to other cancer types, as well as to add new molecular and phenotypic annotation to the breast cancer dataset defined in this article. Low-depth, whole-genome sequences have already been generated for a number of breast tumor cell-line models (23), and hundreds of complete cancer genomes and their normal counterparts are expected to be available within a year or 2 (24). Somatic mutations in cancer genomes include both “driver” mutations, which confer clonal growth advantage on the cancer cell and have therefore been selected during development of the cancer, and “passenger” mutations, which are not implicated in cancer development. As an example, the recent comprehensive identification of somatic DNA mutations in a melanoma cell line, COLO829, identified 33,345 substitutions, 66 small insertions or deletions, and 37 rearrangements (11). On the basis of sequencing data alone, it may be difficult to interpret the significance of most of these mutations and therefore to identify what drives survival in this particular tumor cell. Although computational approaches may indeed direct the discrimination of driver from passenger mutations, many thousands of samples may need to be profiled. Therefore, functional assessments of cancer cell dependencies, as described in this article, will no doubt complement these large-scale sequencing efforts and aid the identification of therapeutic targets (25).

The selectivity of TTK inhibition for tumor cell models with PTEN mutations we describe is a clear example of synthetic lethality/sickness, in which a combination or synthesis of effects (i.e., PTEN mutation and TTK inhibition) has a greater influence on cellular fitness than either defect alone. One standard approach to identifying such synthetic lethalities and candidate therapeutic targets in cancer is to use isogenic pairs of cell lines in which the only engineered genetic difference between isogenic comparitors is the tumorogenic mutation in question (in this case, PTEN; ref. 26). Although these genetically simple isogenic systems can be very powerful in identifying the prime determinants of response to a particular therapeutic approach (27), when used in isolation they do not address the impact that changes in additional genes may have on any therapeutic response (28). Although the true extent of intra- or intertumoral genetic heterogeneity is not yet fully known, it seems likely to be a key determinant of the response to treatment. On this basis alone, identifying synthetic lethal effects that are relatively unaffected by changes in additional genes (i.e., hard synthetic lethali- ties; ref. 28) is critical. With these issues in mind, one rational approach to this problem is to use genetically heterogeneous, nonisogenic panels of tumor cell lines to first identify candidate genetic dependencies (as we have done in our study) and then, where available, use isogenic systems to validate effects. Such an approach goes some way toward assessing the effects of genetic heterogeneity upon candidate therapeutic approaches and could even be complemented with additional techniques, such as the use of synthetic rescue screens (29). Given the cost and time involved in the processes of preclinical and clinical drug development, it seems reasonable to propose that similar or complementary approaches be routinely used to address the impact of genetic heterogeneity prior to and during the drug development process.

Our identification of TTK as a candidate therapeutic target for PTEN-deficient tumors is particularly interesting, given the current focus on development of clinical inhibitors of this mitotic kinase. At present, a number of TTK inhibitors have been created, including relatively potent and orally bioavailable small molecules (19, 30). Primarily, the utility of TTK inhibitors is thought to be in the targeting of tumors...
defined by high levels of chromosome instability (31) and/or those reliant upon the function of the spindle assembly checkpoint (32). In this paper, we define PTEN deficiency as an additional biomarker that could be used to direct the use of these agents. As well as defining PTEN gene defects by conventional Sanger sequencing or FISH, tumor-specific PTEN deficiency can also be effectively delineated by immunohistochemistry (33). Given this, real potential exists that, once developed into clinically usable agents, TTK inhibitors could be assessed in patient cohorts defined by tumor PTEN status. Moreover, the PTEN selectivity of agents such as PARP inhibitors (26) also suggests that combinatorial approaches using these agents together with TTK inhibitors in PTEN-null tumors should also be investigated. In terms of understanding the mechanistic basis underlying the PTEN/TTK synthetic lethality, we show that PTEN-null cells possess an underlying aneuploidy or genomic instability that is exacerbated by the effect of TTK inhibition. This paradigm has some precedence. For example, in the study by Gray and colleagues (19), a small-molecule TTK inhibitor caused cells to exhibit abnormal numbers of chromosomes and was particularly lethal in the U2OS cell line that has an underlying aneuploidy phenotype. Aneuploidy is frequently observed in both human breast carcinomas that have low expression of PTEN and prostatic intraepithelial neoplasias from Pten mutant mice (17, 34). It is reasonable to conclude that inhibition of mitotic checkpoint kinases such as TTK exploit these underlying defects and thus elicit tumor-selective cell death.

We have also shown that ADCK2 silencing is selectively lethal in ER$^+$ breast tumor cell lines. In addition, we noted that ADCK2 inhibition abrogates estrogen signaling. The ADCK2 protein encompasses a predicted protein kinase catalytic domain, an aarF domain, and is highly conserved (99% amino acid identity between Homo sapiens and Pan troglodytes and 54% between H. sapiens and Danio rerio). Although the function of ADCK2 is little understood, other aarF-domain proteins, such as ADCK3, have been implicated in the metabolism of coenzyme Q (ubiquinone), a lipid electron and proton carrier in the electron transport chain (35). In addition, ADCK2 has been identified as a genetic dependency in a glioblastoma tumor cell line, a phenotype that also correlates with an increase in ADCK2 copy number (36). At present, it is unclear whether ADCK2 itself has a coenzyme Q metabolism role or whether this plays a part in impairing estrogen signaling. No ADCK2 inhibitors...
are currently available; however, our results implicate ADCK2 as a possible candidate target for future drug development, as a means to target ER− breast cancer.

In summary, we describe how functional viability profiles of breast tumor cell lines may be used to identify novel candidate therapeutic targets, as well as increasing our understanding of the fundamental dependencies that breast tumor cells carry. As far as we are aware, this is the first attempt to comprehensively identify genetic dependencies for a set of potentially druggable genes in breast cancer cell lines. In a system akin to the free availability of transcriptomic datasets, we believe that these functional data, and subsequent expansions of this approach, will be of great use to the community. The addition of further annotation to our dataset (e.g., data describing the phenotypic properties of each model and the whole-genome sequence of each model) will, it is hoped, enable wider utility of these data. Therefore, we have deposited all raw and processed data on the ROCK database (37) as well as the Stand Up To Cancer Breast Cancer Browser (due for release in 2012), along with tools that facilitate data analysis. We envisage that expanding this dataset to other cancer types should enable identification of novel candidate targets in specific genetic and phenotypic subsets of cancer, and ultimately speed the development of novel therapeutic approaches.

METHODS

Cell Lines, Compounds, and siRNA

All cell lines were obtained from American Type Culture Collection (ATCC) and maintained according to the supplier’s instructions. Cell lines were grown and transfected with SMARTpool siRNAs, using Dharmafect 3, Dharmafect 4 (Dharmacon), Oligofectamine, Lipofectamine 2000, or RNAiMAX (Invitrogen) transfection reagents as indicated in Supplementary Table S9. The siRNA library (siARRAY—targeting 714 known and putative human protein kinase genes) was obtained in 9 96-well plates from Dharmacon. Each well in this library contained a SMARTpool of 4 distinct siRNA species targeting different sequences of the target transcript. Each plate was supplemented with siCONTROL (10 wells; Dharmacon).

Antibodies

Antibodies targeting the following were used per the manufacturers’ instructions: ACTIN and C-MYC (Santa Cruz Biotechnology), TTK, ER, PR, ERBB2, CYCLIN D1, TFF1, FOXO1, C-JUN, PTEN C-terminal (Cell Signaling Technology, Danvers, Massachusetts), and ADCK2 (Abcam). All secondary antibodies used for Western blotting and by viability assays of silencing effects with individual oligos. Cells were transfected with individual ERBB2, ADCK2, ESR1, PIK3CA or TTK siGenome oligos (Dharmacon). Protein lysates were collected 48 hours following transfection for Western blot analysis. Cell viability was measured using CellTiter-Glo (Promega) and MR or XAR film (Kodak).

RNA Interference Screening

We transfected cells with the SMARTpool library, in which each well of the 96-well plate contained a pool of 4 different siRNAs (a SMARTpool) targeting 1 gene. After 5 population doublings (Supplementary Table S9), cell viability in each well was estimated via a highly sensitive luminescent assay measuring cellular ATP levels (CellTiter-Glo; Promega). To identify loss of viability or failure to proliferate effects in each cell line, luminescence readings from each well were log transformed and then centered by the plate median, to account for the plate-to-plate variation common in high-throughput screens. Well-position effects were identified and eliminated, and the quality of data from each plate was estimated by calculating Z prime factors (38) based on positive (siRNA targeting PLK1) and negative (nontargeting siRNA) controls in each plate. The dynamic range of each screen was determined by calculating Z prime values (38); we use a threshold of Z prime > 0.3 to define acceptable screens, based upon previous ones in which Z prime > 0.3 is predictive of reproducible data (39, 40). To allow data to be compared between different cell lines and to minimize the impact of moderate variations in transfection efficiency between cell lines, plate-centered data from each screen were standardized by a Z-score statistic, where Z = 0 represents no effect on viability and negative Z-scores represent loss of viability. Z-scores were calculated using the median absolute deviation of all effects in each cell line (5). In each screen the Z-score data approximated a normal distribution, allowing comparison of the individual siRNA effects across cell lines (Supplementary Fig. S1D–F). siRNA screens were carried out in triplicate, and comparison of Z-score data from replica screens of each cell line showed the screening process to be highly robust, as demonstrated by Spearman r values approaching 1 for all comparisons (Supplementary Table S1). Quality metrics for each screen are shown in Supplementary Table S2. Raw luminescence values and processed data are now deposited on the ROCK Breast Cancer Functional Genomics Database (ref. 37; http://www.rock.icr.ac.uk) as a community resource. This rigorous quality control method reduced our cell-line panel to 20 cell lines with robust viability profiles (see Supplementary Table S3). Heat maps of Z-scores that separated different phenotypes, as defined by a median difference permutation test, were created to show the significant genes ranked in rows according to their median difference and ordered in columns by phenotype.

Western Blots

Protein lysates were prepared using radioimmunoprecipitation assay lysis buffer (50 nM Tris, pH 8.0; 150 mM NaCl; 0.1% SDS; 0.1% sodium deoxycholate; 1% Triton X-100; 50 mM NaF; 1 mM Na3VO4; and protease inhibitors). A total of 100 mg of total cell lysate was loaded onto prefabricated 4% to 12% Bis-Tris gels (Invitrogen), with full-range rainbow molecular weight marker (GE Healthcare) as a size reference, and resolved by SDS-PAGE electrophoresis. Proteins were transferred to nitrocellulose membrane (Bio-Rad), blocked, and probed with primary antibody diluted 1 in 1,000 in 5% milk overnight at 4°C. Secondary antibodies were diluted 1 in 5,000 in 5% milk and incubated for 1 hour at room temperature. Protein bands were visualized using ECL (GE Healthcare) 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 oligos. Cells were transfected with individual ERBB2, ADCK2, ESR1, PIK3CA or TTK siGenome oligos (Dharmacon). Protein lysates were collected 48 hours following transfection for Western blot analysis. Cell viability was measured using CellTiter-Glo (Promega) after 5 population doublings.

Survival Assays

For measurement of sensitivity to TTK inhibitor treatment, cells were plated in 96-well plates and exposed to the drug at the indicated concentrations. Cells were dosed at 24 and 96 hours. After 7 days, cell viability was measured using CellTiter-Glo (Promega) and MR or XAR film (Kodak).

Metaphase Spreads

After 24 hours of treatment with 2 μM AZ3146, HCT116 PTEN wild type or PTEN-null cells were treated with colcemide (10 ng/mL; Sigma-Aldrich) and MG132 (20 μM; Sigma-Aldrich) for 1 hour. Cells were next lysed in hypotonic solution (0.03 M sodium citrate) for 20 minutes at 37°C and fixed in methanol/acetic acid (3:1). Then, 2 or 3 drops of suspended cells were applied to glass slides, and chromosomes were stained with 4',6-diamidino-2-phenylindole.
PTEN FISH

Cell pellets of CAL51 and Hec-1b were washed twice with PBS and fixed with methanol/acetic acid solution (3:1) onto slides. LSI PTEN (10q23; red/orange)/chromosome 10 centromere (CEP 10; green) Dual Colour Probe (Abbott Molecular, Abbott Park, Illinois) was hybridized to representative slides of the cell lines according to the manufacturer’s instructions. Signals were counted in 100 non-overlapping nuclei, using the Leica TCS SP2 confocal microscope (Leica, Milton Keynes, UK).

Transcript Profiling

RNA was extracted from cell lines with TRIZol and phenol/chloroform extraction followed by isopropanol precipitation. For each cell line, triplicate extractions and profiles were performed. Biotin-labeled cRNA was produced by means of a linear amplification kit (IL1791; Ambion, Austin, Texas) using 250 ng of quality-checked total RNA as input. Chip hybridizations, washing, Cy3-streptavidin (Amersham Biosciences) staining, and scanning were performed on an Illumina BeadStation 500 (Illumina, San Diego, California) platform, using reagents and following protocols supplied by the manufacturer, as previously described (12). cRNA samples were hybridized on Illumina human-6 v2 BeadChips, covering approximately 47,000 RefSeq transcripts. The random distribution of large populations of oligonucleotide-coated beads across the available positions within the human-6 v2 chip enables, on average, 30 intensity measurements per RefSeq, yielding quantitative assessments of gene expression. All basic expression data analysis was carried out using the manufacturer's software, BeadStudio 3.1. Illumina expression profiles were performed in triplicate; the raw data were then variance-stabilizing transformed, and the robust spine normalized using the lumi package in the Bioconductor software. Expression values for each sample were median scaled, and the mean expression value was established over the 3 replicates. Genes with significant difference in expression between cell lines were identified by 1-way ANOVA.

aCGH Analysis Method

Genomic DNA was extracted from cell lines using the QiAamp DNA Blood Mini Kit (51104; QiaGen), according to the manufacturer's instructions. Microarray-based CGH analysis was performed on an in-house 32K tiling path bacterial artificial chromosome (BAC) array platform, as previously described (41). For copy-number correlations, the circular binary segmentation ratios of BACs containing the gene of interest were used for copy-number correlations, and copy number was assigned as previously described (41). Briefly, adaptive weight smoothed (AWS) log2 ratio values < −0.12 were categorized as losses, those in between as gains, and those in between as unchanged. Amplifications were defined as smoothed log2 ratio values > 0.4 (41). Data processing and analysis were carried out in R 2.9.0 (42) and BioConductor 1.5 (43), making extensive use of modified versions of the aCGH packages, marray and AWS.

Correlation of Gene Expression and aCGH Data

aCGH and gene expression were compared by direct Pearson correlation of gene expression log intensity values with smoothed log ratio values for every probe in the gene expression data. Correlation P values were adjusted for multiple comparison testing, using the local false discovery rate (FDR) method of Benjamini and Hochberg (44), as previously described (12). Mann–Whitney U tests were performed for each gene to compare gene expression values in groups defined as amplified or not, gained or not, lost or not, and deleted or not, using the thresholded aCGH calls. Wilcox test P values were corrected for multiple comparison testing within contiguously altered regions. For each gene, cases in which genes were amplified, gained, lost, or deleted were recorded, along with the fold changes between samples carrying a copy-number change and those that did not. Total counts of gains, losses, amplifications, and deletions were also recorded.

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 in which siRNA caused significant loss of viability (Z < −2). The Z-score was compared with normalized gene expression, using a Pearson correlation coefficient. A gene was taken as being significantly correlated if the Pearson correlation coefficient was significantly different from the null hypothesis, the correlation was inverse, and the variation in gene expression between cell lines was significantly different as assessed by 1-way ANOVA.

For aCGH overlay, Z-scores for the common 714 genes in the gene expression and kinase library were correlated directly, with Pearson correlation and P values adjusted for multiple comparison testing using the FDR (12). Gene expression values were subsequently compared in those cell lines that displayed Z-score hits for each gene and those that did not. Fold changes between these groups for each gene were also calculated. aCGH copy-number changes for each gene were used to divide cell lines carrying amplifications, gains, losses, or deletions from those that did not, and Mann–Whitney U tests were performed to compare gene expression between these groups for each gene.

Clustering Methods

Supervised analysis was performed by calculating the absolute difference in the medians of the 2 groups, followed by estimating a P value by permuting the labels to create a distribution of median differences against which to compare the actual effects. We used a significance cut-off of P = 0.05 with no corrections for multiple comparisons. The combined gene expression, Z-scores, and aCGH matrices were subjected to hierarchical clustering using a Pearson correlation distance measure and the Ward clustering algorithm. Equivalent heat maps were drawn that clustered the samples and genes in the combined overlay, using the corresponding Z-scores. Heat maps of gene expression and copy-number changes were then displayed under the same dendograms derived by clustering according to Z-score. Stability of the clustering was established by pvclust, giving an approximately unbiased P value > 0.9. All analysis was performed in R 2.9.0, using in-house scripts.

Cell-Line Verification

We used short tandem repeat (STR) profiling to verify authenticity of the panel of cell lines. We simultaneously amplified 8 STR loci in a multiplex PCR reaction (Promega PowerPlex 1.2 System) and used the ATCC database for comparison, where possible. In addition, gain and loss of expression of particular subtype markers were confirmed by Western blot as well as microarray analysis.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Acknowledgments

We thank members of the Gene Function and Drug Target Discovery Laboratories for helpful comments and suggestions and Chris Torrance at Horizon Discovery for providing isogenic cell lines.

Grant Support

This work was supported by grants from Breakthrough Breast Cancer and Cancer Research UK. Research supported by Stand Up To Cancer-American Association for Cancer Research Dream Team Translational Cancer Research Grant, Grant Number SU2C-AACR-DT0409. We also acknowledge National Health Service funding to the National Institute for Health Research Royal Marsden Hospital Biomedical Research Centre.

Received May 11, 2011; revised June 24, 2011; accepted June 24, 2011; published OnlineFirst August 2, 2011.
REFERENCES


