CANCER

A murine preclinical syngeneic transplantation model for breast cancer precision medicine

Lorenzo Federico,[1]*[†] Zechen Chong,[2] Dong Zhang,[1] Daniel J. McGrail,[1] Wei Zhao,[1] Kang Jin Jeong,[1] Christopher P. Vellano,[1] Zhenlin Ju,[1] Mihai Gagea,[3] Shuying Liu,[4] Shreya Mitra,[1] Jennifer B. Dennison,[1] Philip L. Lorenzi,[2] Robert Cardnell,[5] Lixia Diao,[2] Jing Wang,[2] Yiling Lu,[1] Lauren A. Byers,[5] Charles M. Perou,[6] Shiaw-Yih Lin,[1] Gordon B. Mills[1]

We previously demonstrated that altered activity of lysophosphatidic acid in murine mammary glands promotes tumorigenesis. We have now established and characterized a heterogeneous collection of mouse-derived syngeneic transplants (MDSTs) as preclinical platforms for the assessment of personalized pharmacological therapies. Detailed molecular and phenotypic analyses revealed that MDSTs are the most heterogeneous group of genetically engineered mouse models (GEMMs) of breast cancer yet observed. Response of MDSTs to trametinib, a mitogen-activated protein kinase (MAPK) kinase inhibitor, correlated with RAS/MAPK signaling activity, as expected from studies in xenografts and clinical trials providing validation of the utility of the model. Sensitivity of MDSTs to talazoparib, a poly(adenosine 5′-diphosphate–ribose) polymerase (PARP) inhibitor, was predicted by PARP1 protein levels and by a new PARP sensitivity predictor (PSP) score developed from integrated analysis of drug sensitivity data of human cell lines. PSP score–based classification of The Cancer Genome Atlas breast cancer suggested that a subset of patients with limited therapeutic options would be expected to benefit from PARP-targeted drugs. These results indicate that MDSTs are useful models for studies of targeted therapies, and propose novel potential biomarkers for identification of breast cancer patients likely to benefit from personalized pharmacological treatments.

INTRODUCTION

Breast cancer is a heterogeneous disease characterized by remarkable intertumor dissimilarity in terms of presentation, progression, and prognosis (1–3). Integrative molecular analyses suggest that human breast cancers can be subclassified into at least four main groups and six intrinsic subtypes (4, 5). Although these classification systems contribute to clinical management of breast cancer, they fail to capture the broad molecular diversity of this disease (6). Under these circumstances, complementary approaches need to be deployed to reduce experimental noise and validate relevant therapeutic targets.

Murine models of diseases have helped unveil highly conserved molecular pathways and master regulators involved in human oncogenesis (7, 8). Comparative studies of transcriptomic data from genetically engineered mouse models (GEMMs) of breast cancer revealed that several models likely represent preclinical counterparts of both basal-like and human epidermal growth factor receptor 2 (HER2)–enriched human tumors (9). Although GEMMs have been used for preclinical evaluation of targeted therapy (10), in most cases, these models may have limited applicability to human disease because they are driven by a single molecular aberration and do not reflect underlying tumor heterogeneity.

Here, we developed a series of mouse-derived syngeneic transplant (MDST) models by sequential orthotopic transplantation of tumors originating from autotaxin (ATX)–lysophosphatidic acid receptor (LPAR) transgenic mice (11), a murine model of breast cancer in which receptors

for LPA [a well-known pleiotropic bioactive lipid with modulatory effects on various cellular and bodily functions (12)] and ATX [the key enzyme responsible for extracellular LPA synthesis (13)] were up-regulated in the epithelial cells of mammary glands. Studies with this model showed that tumor susceptibility was associated with a proinflammatory status (14), a feature that parallels the dynamics of human breast carcinogenesis (15).

Longitudinal and multiregional resampling of MDST tumor transplants, coupled to phenotypic and molecular characterization, showed that the core molecular identity of each MDST was independent of the number of transplantation passages and thus an intrinsic feature of each model. A group of MDSTs categorized as carcinomas displayed significant intertumor heterogeneity and harbored several distinct molecular aberrations commonly found in human ductal breast cancers.

To demonstrate the utility of the MDST as an experimental platform for pharmacodynamic studies of personalized therapy, we first showed that therapeutic response to mitogen-activated protein kinase (MAPK) kinase (MEK) inhibitors in MDSTs correlates with RAS/MAPK pathway activation, as it does in xenografts and clinical trials. We then showed that expression levels of poly(adenosine 5′-diphosphate–ribose) polymerase 1 (PARP1) protein, a key component of DNA repair pathways currently being targeted in triple-negative and homologous recombination repair–defective breast cancers (16, 17), predict response to talazoparib, a selective and potent inhibitor of PARP1 activity that traps PARP on DNA (18). MDSTs were used to test a new PARP sensitivity predictor (PSP) score, which was developed from integrated analysis of drug sensitivity data and predicted sensitivity to PARP inhibitors. Together, these results demonstrate that MDSTs are useful models for the implementation of targeted therapeutics for breast cancer patients.

RESULTS

Growth rate and across-passage stability of the MDST models

We established 12 different transplantable tumor graft lines from primary mammary gland tumors of transgenic mice expressing LPAR1,

[1]Department of Systems Biology, University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA. [2]Department of Bioinformatics and Computational Biology, University of Texas MD Anderson Cancer Center, Houston, TX 77230, USA. [3]Department of Veterinary Medicine and Surgery, University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA. [4]Department of Breast Medical Oncology, University of Texas MD Anderson Cancer Center, Houston, TX 77230, USA. [5]Department of Thoracic Head and Neck Medical Oncology, University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA. [6]Department of Genetics, University of North Carolina at Chapel Hill, Chapel Hill, NC 27514, USA.
*Corresponding author. Email: lfederico@mdanderson.org
†Present address: Department of Melanoma Medical Oncology, University of Texas MD Anderson Cancer Center, Houston, TX 77054, USA.

LPAR2, and LPAR3, as well as bigenic MDST models coexpressing the LPA-producing enzyme ATX and one of the LPARs (Fig. 1A and Table 1). These immunocompetent transplantable tumor models were termed MDSTs and designated with a random numerical identifier preceded by the name of the transgenic line from which the model was originated (LPA1, LPA2, LPA3, ATXLPA2, and ATXLPA1) (Table 1). Some MDST models grew rapidly and reached critical size in less than 20 days, whereas other models displayed a more indolent nature requiring more than 50 days to reach comparable dimensions (Fig. 1B). MDSTs originating from different mice with the same transgenic alteration showed marked differences in growth rate, and the simultaneous presence of transgenes encoding LPARs and the LPA-producing enzyme ATX did not result in any apparent synergistic effect on the growth of bigenic models (Fig. 1B; compare growth rates of LPA1-T22 and bigenic ATXLPA1-T22 MDST). Notably, growth rates of MDSTs remain stable, independent of the number of passages (Fig. 1C), although two rapidly growing MDST models, LPA2-T43 and LPA2-T13, displayed greater across-passage variation. These results suggest that the growth rate of each MDST is an intrinsic characteristic maintained over passage in mice.

## Expression of LPA and ATX transgenes in the MDST models

Although the LPA and ATX transgenes remained detectable in genomic DNA of MDSTs, irrespective of passage number and region of collection (Fig. 2A), the LPA and ATX transgenes were not detectable at the RNA or protein level (Fig. 2, B and C), despite being expressed in the original murine tumors (Fig. 2C, passage P0). Further selective inhibition of ATX activity in vivo (Fig. 2, D and E) did not alter MDST growth rates (Fig. 2F). Therefore, LPA signaling does not drive the growth of MDSTs

following passage but may contribute to the formation of the primary tumor through a proinflammatory process established by the LPA microenvironment (11).

Intertumor phenotypic heterogeneity of the MDST models

Average growth rates of MDSTs were heterogeneous across the panel and ranged from 22 to 168 mm^3/day (Table 1 and Fig. 3A). On the basis of histological features, E-cadherin (CDH1) expression, and growth rate, MDSTs are separated into two groups: SG-CDH1$^+$ carcinomas (average growth rate, <55 mm^3/day) and high-grade undifferentiated FGM sarcomas (average growth rate, >55 mm^3/day) (Fig. 3B, Table 1, and fig. S1). One MDST, LPA2-T16, was classified as an FG-CDH1$^+$ carcinoma (average growth rate, >55 mm^3/day), whereas a second one, ATXLPA2-T16 MDST, presented atypical disperse clusters of CDH1 positivity and was thus classified as an FG-mixed squamous carcinoma (Fig. 3B and Table 1). Histological examination revealed a moderate degree of mixed lymphocytic and myeloid infiltration in all carcinomas and two sarcomas and a higher degree of neutrophilic infiltration in the ATXLPA2-T16 MDST (see section S1).

Short-term in vitro cultures of cells isolated from SG-CDH1$^+$ MDSTs were more adhesive and showed an epithelial/cobblestone morphology (Fig. 3C). In contrast, cells isolated from FGM MDSTs displayed mesenchymal-like characteristics typical of loss of cell-cell contacts and spindle morphology (Fig. 3C). Cells isolated from FGM MDST models continued to grow at serum concentrations as low as 1.5%, which was not permissive for the growth of SG-CDH1$^+$ MDSTs. Additionally, in vitro drug sensitivity studies on six MDST lines using the genotoxic agent cisplatin that is used in therapy for triple-negative breast cancers (TNBCs)
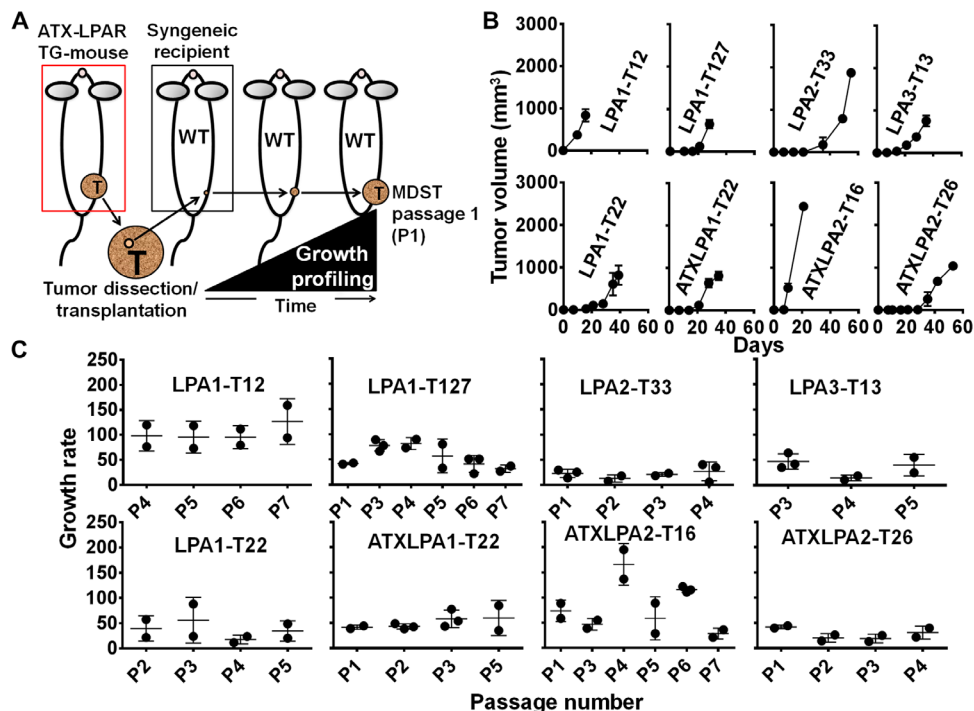
Fig. 1. Growth rate and across-passage stability of the MDST models. (**A**) Schematic of the procedure for the establishment of MDSTs. Fragments of primary tumors (T) were collected from mice with transgenic (TG) alterations of ATX-LPA signaling pathway members and orthotopically transplanted into immunocompetent wild-type (WT) mice. A growth profile of each MDST was recorded and expressed as changes in tumor volume (mm^3) per day. (**B**) Growth curves of MDST models. Average tumor volume (mm^3) ± SD (n = 2 to 3). (**C**) Average growth rates expressed as tumor volume change (ΔV) per day ± SD of MDST models at different passages (P). n = 2 to 3 grafts per passage.

Table 1. Phenotypic and molecular characteristics of the MDST models. Growth rate is expressed as the average change in volume (mm^3) per day ± SD (n = 8 to 14). Epithelial-mesenchymal transition (EMT) score is based on the levels of *Ctnnb1, Cdh1, Cldn7, mir200a, Snai1, Twist1, Tgfb1,* and *Zeb* (table S1). A detailed histopathological examination of each MDST is presented in section S1. Statistical analysis of growth rate differences between MDSTs is reported in table S5. SG-CDH1$^+$, slow-growing, CDH1-positive (<55 mm^3/day); FG-CDH1$^+$, fast-growing, CDH1-positive (>55 mm^3/day); FG-mixed, fast-growing, mixed; FGM, fast-growing/mesenchymal; N/A, not applicable.

MDST	Growth rate (mm^3/day)	Histology	Subclass	CDH1 status	Inflammatory infiltration	EMT score
LPA2-T33	22.45 (±10.45), n = 12	Adenocarcinoma	SG-CDH1$^+$	Positive	Mixed/moderated	−10
ATXLPA2-T26	22.84 (±14.82), n = 11	High-grade adenocarcinoma	SG-CDH1$^+$	Positive	Mixed/moderated	−3.3
LPA3-T13	33.07 (±19.5), n = 8	Adenocarcinoma	SG-CDH1$^+$	Positive	Mixed/moderated	−13.76
LPA1-T22	35.46 (±24.49), n = 9	Adenocarcinoma	SG-CDH1$^+$	Positive	Mixed/moderated	−6.6
ATXLPA1-T22	49.09 (±16.85), n = 11	High-grade adenocarcinoma	SG-CDH1$^+$	Positive	Mixed/moderated	−7.5
LPA2-T16	67.19 (±18.82), n = 10	Adenocarcinoma	FG-CDH1$^+$	Positive	Mixed/moderated	N/A
ATXLPA2-T16	84.46 (±50.91), n = 13	High-grade carcinoma	FG-mixed	Mixed	Mixed/moderated	−0.5
LPA1-T127	55.93 (±23.35), n = 14	High-grade sarcoma	FGM	Negative	Mixed/moderated	N/A
LPA1-T17	89.35 (±28.82), n = 13	High-grade sarcoma	FGM	Negative	N/A	N/A
LPA1-T12	104.02 (±29.13), n = 8	High-grade sarcoma	FGM	Negative	Mixed/moderated	+13
LPA2-T43	144.12 (±78.17), n = 14	High-grade sarcoma	FGM	Negative	N/A	+16.9
LPA2-T13	168.35 (±98.56), n = 10	High-grade sarcoma	FGM	Negative	N/A	+13.5

(*19*) revealed that only fast-growing lines were sensitive to treatment, as would be expected from a genotoxic agent (fig. S2).

These data demonstrate that MDSTs can be divided in two major subgroups characterized by different growth rates, E-cadherin expression levels, histological appearance, and in vitro cellular characteristics. A detailed histopathological examination of each MDST is presented in section S1.

Transcriptomic and proteomic characterization of the MDST models

To characterize the transcriptional landscapes of the MDSTs, we performed RNA-seq analysis on a representative group of five SG-CDH1$^+$ carcinomas, one FG-mixed squamous carcinoma, and three FGM sarcomas. Unsupervised clustering analysis revealed that two samples taken at different passage generations from each MDST were always nearest neighbors in the dendrogram except for two MDST sarcomas (LPA2-T13 and LPA2-T43; the two pairs were, however, in the same subcluster), indicating transcriptional stability with passage (Fig. 4A, left). The SG-CDH1$^+$ MDSTs were transcriptionally heterogeneous as they clustered in five separate branches of the dendrogram (Fig. 4A, clusters 1 to 5). Conversely, all FGM MDSTs grouped in a single cluster, indicating greater homogeneity (Fig. 4A, cluster 6). Multidimensional scaling (MDS) analysis of the same data set produced an analogous pattern, as point coordinates of samples derived from the same tumor were nearest neighbors on the MDS map (Fig. 4A, right). The FGM sarcomas were tightly clustered in the right corner of the chart, whereas SG-CDH1$^+$ carcinomas localized within a larger space (Fig. 4A). The FG-mixed MDST pair clustered in a separate region (Fig. 4A, purple circles 17 and 18), as they did in the dendrogram of Fig. 4A (left, cluster 7). Samples collected from normal mammary glands were positioned at the center of the carcinoma subgroup cluster (Fig. 4A,

right, white circles 1 and 2), suggesting similarity to the transcriptional profile of the SG-CDH1$^+$ carcinomas. A positive EMT score based on the levels of *Ctnnb1, Cdh1, Cldn7, mir200a, Snai1, Twist1, Tgfb1,* and *Zeb1* that are related to EMT was a key feature of the FGM sarcoma subgroup (Table 1 and table S1).

Molecular heterogeneity of the MDST models was not limited to transcriptional profiles. Unsupervised hierarchical clustering of averaged expression level of 170 proteins and phosphoproteins measured by a reversed-phase protein array (RPPA) assigned seven SG-CDH1$^+$ MDSTs to four distinct proteomic clusters (Fig. 4B, clusters 2 to 5), two MMTV-cMYC transgenic MDSTs to a closely related but distinct cluster (Fig. 4B, cluster 6), and all five FGM MDSTs to a single cluster (cluster 1) primarily driven by a proteomic signature characterized by low expression levels of β-catenin, claudin-7, and E-cadherin proteins (Fig. 4B, signature 1), consistent with the biology of these cancers being linked to EMT-related processes (table S1).

As compared to previously described transgenic murine models of breast cancer, where each model is relatively homogeneous (*9*), MDSTs were highly heterogeneous (Fig. 4C and fig. S3), clustering with multiple breast cancer models (PyMT, Myc, normal-like, and p53null-luminal; Fig. 4C). Both samples dissected from normal mammary tissue clustered at the center of the normal-like group (Fig. 4C), substantiating the observation that the transcriptional landscapes of SG-CDH1$^+$ carcinomas and normal tissue are related (Fig. 4A, right, white circles 1 and 2). The FG-mixed MDST clustered with a single model (squamous-like), whereas FGM/sarcoma MDSTs grouped tightly within the claudin-low expression class (Fig. 4C and fig. S3). Thus, the ATX-LPAR transgenic mice (*11*) generate the most heterogeneous group of murine tumors yet observed, encompassing multiple tumor subtypes, including luminal-like, metaplastic, and claudin-low–like tumors.

Identification of molecular targets with direct relevance to human breast cancer therapy

The PAM50 intrinsic classifier is a set of 50 genes broadly used for breast cancer prognostics and treatment selection purposes (20). Although relative expression data of 47 PAM50 classifier genes in both MDST models and patient samples did not fully merge, we identified a number of similarities between murine models and human intrinsic subtypes of breast cancer. *Foxa1*, a gene characteristically up-regulated in human HER2-positive cancers, was overexpressed in ATXLPA1-T22 MDST (Fig. 5A), a model with a *Kras* mutation (table S3). Additionally, *Actr3b* (a PAM50 gene typically overexpressed in basal breast carcinomas) and *Slc39a6* (a PAM50 gene distinctively up-regulated in both luminal A and luminal B subtypes) were highly expressed in multiple MDSTs (Fig. 5A).

Expression of the estrogen receptor was low in all MDSTs, as observed in most murine models of breast cancer. However, two MDST models, SG-CDH1+ LPA2-T33 and FGM sarcoma LPA2-T43, were sensitive to tamoxifen treatment (fig. S4).

Several potentially actionable genes, including *Notch2*, *Akt3*, *Cdkn2a*, *Kras*, *Bcl2*, *Aurka*, *Erbb3*, *Igf2*, *Fgfr3*, and *Fgf14*, were either up-regulated or mutated in one or more MDSTs (Fig. 5B and tables S2 and S3), indicating high intertumor molecular heterogeneity and enrichment in molecular targets of interest. To search for genes commonly amplified, mutated, or deleted in human breast invasive carcinomas, we interrogated The Cancer Genome Atlas (TCGA) database (21) and found that multiple proteins frequently amplified in large cohorts of patients, including PARP1, histone H3, cMYC, NOTCH1, PKCα, and P16INK4A

(*Cdkn2A* gene), were also differentially expressed or activated in MDST carcinomas (frequency alteration range, 22.2 to 5.1%; *n* = 959 patients) (Fig. 5C and table S2). *PARP1*, currently under investigation as a target in BRCA1/2 mutant tumors, was amplified in 13.5% of patients (table S2) and up-regulated in MDST carcinomas (Fig. 5C), with the exception of ATXLPA2-T16, the MDST model with mixed mesenchymal-squamous epithelial differentiation (section S1). cMYC and histone H3 proteins, whose respective genes (*MYC* and *H3F3A*) are amplified in 21.9 and 13.8% of breast cancer cases, respectively (table S2), were selectively up-regulated in distinct MDSTs (Fig. 5C). Our results indicate that each MDST is characterized by specific and stable molecular alterations, several of which with direct relevance to human breast cancer pathophysiology and therapy.

Personalized approaches for breast cancer therapy in MDST models

In terms of validating the MDSTs as models to explore targeted therapy in breast cancer, we found that RAS/MAPK pathway activity status, as assessed by RPPA, was highly predictive of response to trametinib, a specific and well-characterized MEK1/2 inhibitor (22), in four of five MDST models with different MEK phosphorylation levels (fig. S5A). PARP1 levels have been suggested to predict response to "trapping" PARP inhibitors (23); however, validation in heterogeneous tumor models is lacking. We thus tested the sensitivity of the MDST models to talazoparib, a potent and selective trapping PARP1 inhibitor currently being evaluated in TNBCs with defects on homologous recombination repair, and found that the treatment blocked tumor growth, with efficacy being predicated by
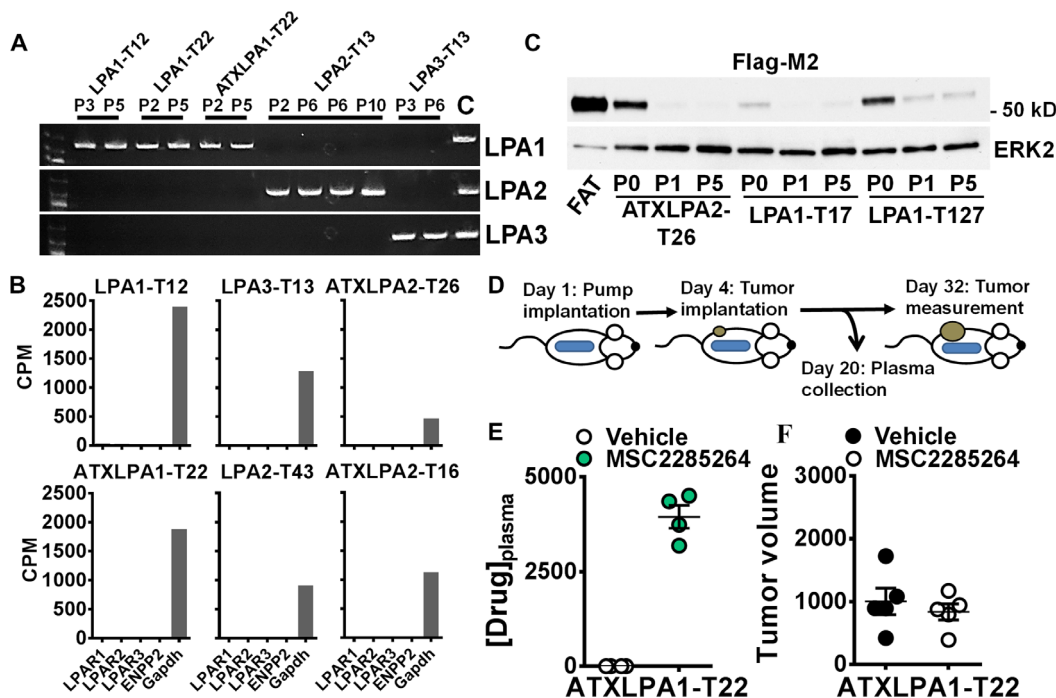
Fig. 2. Expression of LPA and ATX transgenes in the MDST models. (**A**) Polymerase chain reaction (PCR) amplification of human LPAR1, LPAR2, and LPAR3 transgenic regions in DNA isolated at different passages from distinct MDSTs. MDST IDs and passages (P) are indicated at the top. Amplification bands from known transgene-positive tissue samples are shown in column C (control). (**B**) Expression level of human LPAR1, LPAR2, LPAR3, and ENPP2 (ATX) transgenes in six different MDST models by RNA sequencing (RNA-seq) (*n* = 2). The expression level of murine *Gapdh* is used as an endogenous standard. CPM, counts per million. (**C**) Western blot analysis of transgenic Flag-M2 expression in samples from three MDST models at different passages (P). P0 represents the primary tumor. Extracellular signal–regulated kinase 2 (ERK2) protein level was used as a loading control. (**D**) Schematic of MDST transplantation and the ATX inhibitor MSC2285264 delivery via osmotic pump. (**E**) Plasma concentration level (ng/ml) of the ATX inhibitor MSC2285264 after 20 days of treatment. (**F**) Size of ATXLPA1-T22 MDST transplants (mm³) after 1 month of treatment with vehicle or the ATX inhibitor MSC2285264.
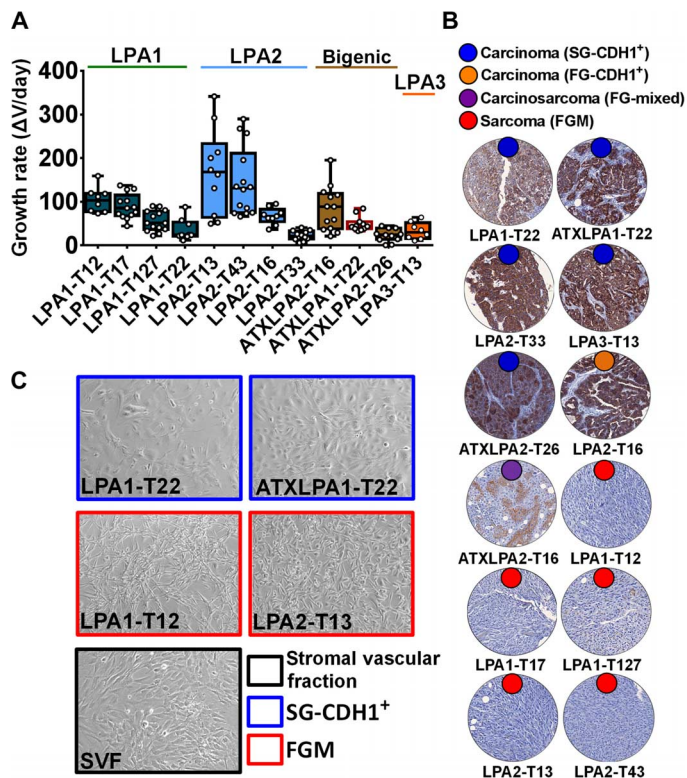
Fig. 3. Intertumor phenotypic heterogeneity of the MDST models. (**A**) Box plots of MDST growth rates. The average tumor volume changes (mm^3) per day ± SD of $n = 8$ to 14 distinct grafts are shown for each MDST. MDST models are grouped according to their transgenic background. (**B**) *CDH1* immunohistochemical staining of MDST. MDST IDs are indicated below each image. Representative hematoxylin and eosin–stained sections of these tumors are shown in fig. S1. (**C**) Phase-contrast images of cells isolated from stromal vascular fraction of healthy mammary glands, two representative carcinomas (LPA1-T22 and ATXLPA1-T22), and two representative sarcomas (LPA1-T12 and LPA2-T13).

PARP1 protein levels in four of five MDSTs (fig. S5B). The overall prognostic value of target status in MDST models was estimated via binomial testing of observed response variables (predicted versus unexpected) in all trials (talazoparib + trametinib) and found to be highly significant (80% correct prediction in $n = 10$ trials, $P < 0.0001$; fig. S5C). This suggests that the molecular features of these models are predictors of response to targeted therapy and that the models could be used to identify biomarkers of response to additional targeted therapies.

In vitro quantification of the DNA double-strand break marker γH2AX in four different MDSTs showed that basal levels of DNA damage are generally low but can be induced in vitro by talazoparib as well as by irinotecan and cisplatin (fig. S6). To understand whether PARP1 expression is more broadly associated with sensitivity to PARP inhibitors, we also interrogated responses to olaparib (AZD-2281), a second trapping PARP inhibitor, in 76 women's cancer cell lines, including breast, ovary, and endometrium. After stratifying cell lines into sensitive, intermediate, and resistant groups based on the tertile of area under the curve (AUC) values (fig. S7A), we found that PARP1 protein levels of a subset of these cell lines measured by RPPA were associated with olaparib sensitivity (fig. S7B). In addition, the response of a panel of small cell lung cancer (SCLC) cell lines to olaparib revealed a strong correlation between PARP1 expression and olaparib sensitivity ($r = -0.499$, $P = 0.049$; fig. S7C). To further validate our results, we have in-

vestigated the correlation between PARP1 protein levels and sensitivity to olaparib and talazoparib in TNBC cell lines and confirmed that the PARP protein expression level is positively correlated to sensitivity to PARP inhibitors (fig. S7, D and E). These results demonstrate that MDSTs are effective models to study targeted therapy pharmacodynamics and support the notion that PARP1 expression level is an actionable predictor of response to selective PARP1 inhibition.

PSP score

In one MDST model, PARP1 expression was not associated to sensitivity to PARP-targeted therapy. We thus developed a PSP score from integrated sensitivity data for multiple PARP inhibitors from the COSMIC (Catalogue of Somatic Mutations in Cancer) database (24) using gene expression levels from the Cancer Cell Line Encyclopedia (fig. S8A) (25) to determine whether it added to the ability of PARP levels to predict response to therapy. PSP signature genes differentially expressed ($P < 0.005$; fold change > 1.2) between sensitive cell lines (average z score < -1.0) and nonsensitive cell lines (average z score > 0) are listed in table S4. The receiver operating characteristic (ROC) for the training set yielded AUCs of 1.0 in the talazoparib training data and 0.89 in the training data with rucaparib (AG-014699), an additional PARP inhibitor, whereas validation of PSP score in a separate test set of breast cancer cell lines with olaparib was 88% accurate (fig. S8A).

Comparison between PSP scores and sensitivity to targeted treatment for each MDST revealed that PSP magnitude predicts sensitivity to targeted treatment, as positive scores were associated to sensitivity and negative scores to resistance in five of five independent MDSTs (fig. S8B). Comparison between MDST sensitivity data and homologous recombination defect (HRD) signature scores previously defined by HRD in various cancers (26) did not reveal association between therapeutic response and defects on DNA repair (fig. S8C), suggesting that, in this model, tumor inhibition by PARP1 blockade may not be dependent on synthetic lethality with HRD. Analysis of patient samples from the TCGA Data Portal (21) revealed that 21.3% of breast cancer patients have PSP scores above 0.1 indicative of potential sensitivity to PARP inhibitors (42, 7, 10, and 32% of basal, HER2, luminal A, and luminal B breast cancers, respectively) (fig. S8D). Furthermore, this classification approach indicates that a large percentage of basal and luminal B cancers, two aggressive intrinsic subtypes with limited therapeutic options (27, 28), could potentially benefit from PARP inhibitors either alone or in combination with other pharmacological agents. These data suggest that MDSTs are useful models for pharmacodynamic studies of targeted therapy and that the PSP score is a potential predictor of response to PARP inhibition that warrants assessment in human clinical studies.

DISCUSSION

We have characterized MDSTs established from a series of murine tumors that spontaneously developed from mammary glands of ATX-LPAR transgenic mice (11). The ATX and LPAR transgenes remained detectable in the MDST even after multiple regrafting but were not involved in cancer growth, supporting the idea that LPA production by ATX or action through any of the LPARs is not responsible for growth of the established MDST. However, the presence of detectable LPAR and ATX levels in the original tumor from the transgenic mice suggests that the transgenes may promote the establishment of a proinflammatory milieu that contributes to the late onset of tumor development in this model (14, 15, 29).

Molecular and phenotypic features of each MDST were surprisingly stable. Although intratumoral heterogeneity may arise during tumor
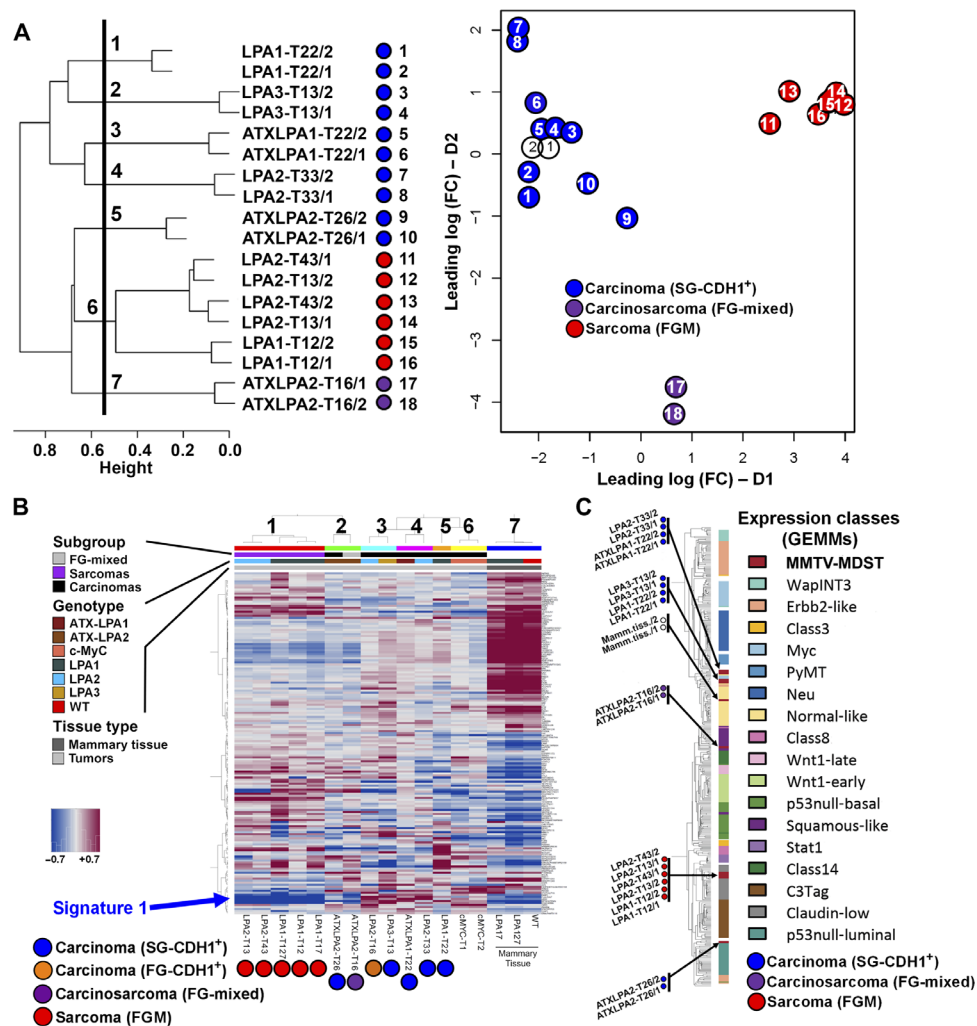
Fig. 4. Molecular profiling of MDSTs and comparison with GEMMs. (**A**) Hierarchical clustering of transcriptomic data from six carcinomas and three sarcomas ($n = 2$ samples collected from each MDST at different transplant generations). The list of MDST samples (1 to 18) (left) is mapped to a corresponding location on the MDS diagram on the right (numbered circles). Leading logFC (fold change) value coordinates D1-D2 convey transcriptional profile relatedness. Circles 1 and 2 on the MDS graph represent two samples dissected from healthy mammary glands of WT mice. (**B**) Two-way unsupervised hierarchical clustering heat map of the average $\log_2$ protein expression levels by RPPA ($n = 6$ to 20 samples collected from each MDST at different transplant generations). Horizontal color bars at the top indicate proteomic clusters (1 to 7), subgroup classification, genotype (LPA1, LPA2, LPA3, bigenic, cMYC, and WT), and tissue type. MDST IDs are listed at the bottom. Signature 1 is composed of expression levels of PARP1, claudin-7, HER3, β-catenin, and E-cadherin proteins. (**C**) Intrinsic supervised clustering analysis applied to gene expression levels in 405 mouse samples taken from 17 GEMMs. MDSTs are in dark red (arrows).

progression through an evolutionary process driven by genetic variability across subclones (30), neither evolutionary pressure potentially associated to sequential regrafting nor intrinsic regional heterogeneity of the tissue appeared sufficient to drive detectable phenotypic changes in the MDST. The stability observed in the MDSTs suggests that these syngeneic models of orthotopic tumor transplantation in immunocompetent hosts have the potential to provide effective experimental approaches for investigating relationships between macroscopic behavior and molecular features of cancers with sensitivity to targeted therapies including immuno-oncology agents.

A subgroup of MDSTs was composed of fast-growing undifferentiated sarcomas characterized by gene expression patterns suggestive of mesenchymal differentiation. Cells isolated from these tumors readily proliferated in low-serum medium, grew in a disorganized fashion, and had spindle-shaped morphology typical of mesenchymal differentiation. One MDST

model was classified as an epithelial-mesenchymal mixed carcinosarcoma, a type of metaplastic cancer rather uncommon in patients (31) likely originating from transdifferentiation processes taking place in mammary glands. Although uncommon, the very poor outcome associated with sarcomas and metaplastic tumors renders these potentially important models.

In contrast to MDST FGM sarcomas, which showed similar phenotypic and molecular features, a group of MDST SG-CDH1+ carcinomas displayed remarkable intertumor molecular diversity, closely reflecting the heterogeneous nature of human breast cancers (1, 3, 4). Transcriptional profiles of some carcinomas and normal mammary tissue were highly related, suggesting either molecular dynamic similarity or enrichment in stromal elements. Additionally, cells isolated from MDST carcinomas were morphologically similar to mammary epithelial cultures, as they organized in cobblestone-like monolayers and generally required high serum concentrations for optimal growth.
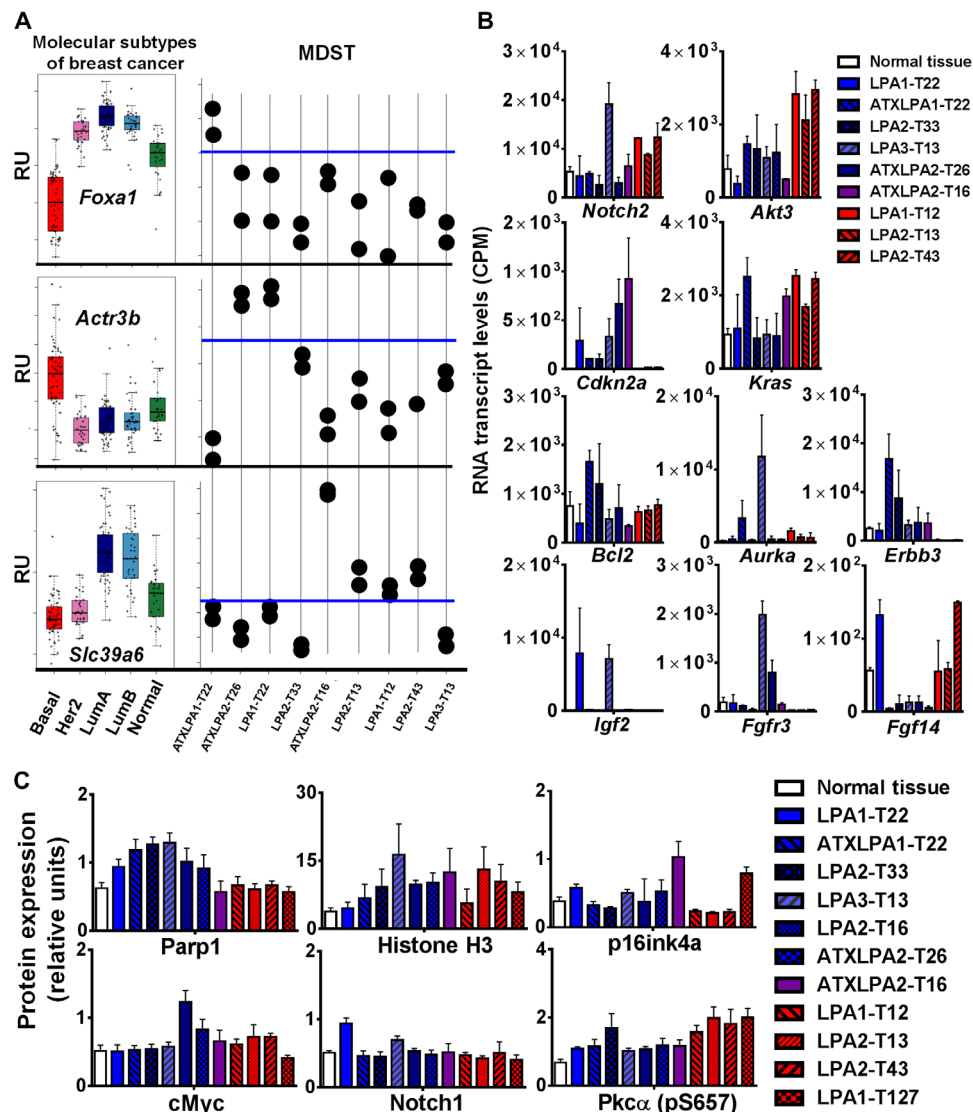
Fig. 5. Identification of molecular targets with direct relevance to human breast cancer therapy. (**A**) Similarities between murine models and human intrinsic subtypes of breast cancer. Expression pattern comparison of three murine homologs of PAM50 classifier genes in five molecular subtypes of breast cancer [basal, Her2[+], luminal A (LumA), luminal B (LumB), and normal-like] (*n* = 232) and nine distinct MDSTs (*n* = 2 samples collected from each MDST at different transplant generations). The empirical threshold for marker up-regulation is shown (blue line). RU, relative units. (**B**) The average expression level of select genes of interest in MDSTs is presented as the average CPM ± SD of *n* = 2 samples collected from each MDST at different transplant generations. (**C**) Expression levels for proteins of interest that are differentially expressed or activated in MDSTs. Each bar represents the average normalized linear intensity value ± SD of 6 to 20 samples collected from each MDST at different transplant generations.

As observed in mouse models of breast cancer, estrogen receptor expression in the MDST was low, possibly because of interspecies differences in hormone pathophysiology, but two tumors, LPA2-T33 and LPA2-T43, showed sensitivity to tamoxifen treatment, indicating that these MDSTs can be used to identify alternative estrogen-related pathways regulated by tamoxifen. In addition to sharing a number of transcriptional pattern similarities with human breast cancer subtypes from the PAM50 classifier (*20, 32, 33*), we discovered that MDSTs carry a number of tumor-specific molecular alterations in genes that are frequently amplified in human breast cancers, such as *PARP1*, *MYC*, and *H3F3A*. The variety and stability of molecular alterations found in MDST models made it possible to test targeted inhibition in models with distinct and clinically relevant molecular features.

Preclinical trials with either trametinib (a well-characterized MEK inhibitor used for the treatment of metastatic BRAF mutated melanomas) or talazoparib (a selective and potent inhibitor of PARP1 currently in phase 3 trials for BRCA-mutated metastatic breast cancer) demonstrated that MDST sensitivity to treatment depends on the level of expression or activation of the targeted molecule. These results not only confirmed drug selectivity but also demonstrated the utility of MDST models in preclinical testing of targeted therapeutics. This suggests that clinical trials based on "pathwayness" defined by the analysis of signaling pathways may prove to be powerful predictors of patient response.

PARP inhibitors are effective in cells with HRDs through synthetic lethality (*16*). Tumor inhibition by PARP1 blockade did not appear to

occur by way of synthetic lethality in the MDST models, because talazoparib treatment was relatively ineffective in a number of MDST models with defects in homologous recombination defined by an HRD signature. However, PARP1 levels in the MDST were strong predictors of response to talazoparib. Similarly, PARP1 protein expression also predicted sensitivity to PARP inhibitors in TNBC and SCLC human cell lines, a result consistent with the notion that low PARP1 expression is associated to PARP inhibitor resistance (34) and that talazoparib treatment decreases PARP1 levels in SCLC cells (35), a cell type previously shown to be sensitive to PARP inhibition (36).

PARP1 expression showed predictive power in multiple MDST models but did not predict the response to talazoparib in MDST LPA2-T33, a tumor that was resistant to treatment despite having elevated PARP1 protein levels and defective homologous recombination. The only parameter that accounted for this unexpected lack of sensitivity was the PSP score, which was low in this MDST regardless of the level of PARP1 protein and HRD score. The prevalence of the PSP signature, particularly in basal and luminal B breast cancers, suggests that subgroups of breast cancer patients may benefit from PARP inhibitor–based treatments, especially the ones without viable therapeutic options.

A subset of PARP inhibitors is believed to work by blocking enzymatic activity of PARP and by trapping PARP at DNA double-strand breaks (18). Our data suggest that talazoparib may be most effective when PARP1 is functionally connected to permissive signaling networks associated to the PSP signature. The PSP score was obtained from sensitivity data for olaparib and rucaparib, two inhibitors that show lower PARP trapping activity as compared to talazoparib. However, the ability to predict the activity of talazoparib suggests that the PSP score might be generalized to all PARP inhibitors. Analysis of key genes differentially expressed between sensitive and nonsensitive cell lines indicated that the largest cellular compartment represented in the PSP signature was nuclear protein genes, including *YEATS4* (a transcription factor involved in the regulation of drug resistance) and the DNA mismatch repair gene *MSH3*.

In conclusion, we have shown that the availability of MDSTs with a heterogeneous spectrum of molecular alterations that propagate in syngeneic mice represents a reliable preclinical tool for studies of targeted therapies. By integrating data from these models, in vitro studies, and gene expression–based modeling, we propose the PSP score as a potential selection criterion for sensitivity to PARP1-targeting therapy.

MATERIALS AND METHODS
Animals
Syngeneic grafts used in these studies were established from transgenic mice on an FVB/N inbred background with homozygous insertions of *LPAR1*, *LPAR2*, *LPAR3*, or *ENPP2* (ATX) human transgenes under the control of an MMTV-KCR promoter, as previously described (11). Animals had free access to standard rodent diet (PicoLab Rodent Diet 20, catalog no. 5053) and water and were group-housed in rooms with temperature and light/dark cycle length parameters controlled according to guidelines established by the National Institutes of Health/Department of Health and Human Services (*Guide for the Care and Use of Laboratory Animals*). All experiments involving animals were approved by the Institutional Animal Care and Use Committee.

MDST transplantation procedure
Four- to 8-week-old wild-type recipient mice syngeneic to the FVB/N background were anesthetized by isoflurane inhalation (2 to 4%). The abdominal area was shaved and cleaned with Betadine disinfectant, and a 0.5-cm incision was made on the distal side of the fourth mammary gland just above the subcutaneous adipose tissue. The subcutaneous mammary fat pad was then gently exposed using sterile fine-point forceps to create a pocket where tumor fragments can be nested. Primary mammary gland tumors were dissected from 12- to 24-month-old multiparous ATX-LPAR transgenic mice (11), carefully cleared of all stromal/fat and necrotic tissues, and cut into several small fragments of approximately 2 mm in length to be immediately seeded. Incision in the transplantation site was closed with wound clips, and tumor growth was monitored weekly. Twelve different transplantable tumor graft lines were established from this procedure and termed MDSTs (Table 1). Screening for the presence of the transgene using PCR and measurement of transgene expression using Western blot were performed as previously described (11). At the end of each experiment, portions of MDST were frozen in freezing medium composed of 90% fetal bovine serum (FBS) and 10% dimethyl sulfoxide (DMSO) and stored in liquid nitrogen for future use. Each MDST was sequentially passaged and resampled from multiple regions at different passage generations to build a repository of transgenic grafts for molecular analysis.

Histological analysis
Samples of MDST tumors were processed histologically and arrayed on a tissue microarray. Sections (4 μm thick) from each tumor were stained with hematoxylin and eosin or the primary antibody for E-cadherin (1:200; Cell Signaling Technology, #4065) and blindly examined by a trained pathologist. Localization and expression of proteins were evaluated by immunohistochemistry, as described previously (37).

Tumor cell isolation and in vitro experiments
Primary tumor and stromal vascular cells were isolated under sterile conditions from the MDST and normal mammary gland tissue, respectively. Generally, one tumor (1.5 cm in diameter) or 10 subcutaneous mammary gland pads from 5-week-old female mice were enough to establish cultures of viable cells. Particular care was taken to ensure that only compact areas of homogeneous neoplastic mass were collected, leaving out surrounding normal tissues, blood, and necrotic tissue. Tissues were finely minced in 150 μl of medium composed of RPMI, 0.2% (w/v) collagenase type II (Worthington, catalog no. 4176), and 1.5% crude serum albumin (Bovine Fraction V) and then digested in a large volume of the same medium (12 to 20 ml) at 37°C for 30 min under gentile rocking. Digested tissues were filtered twice through a 250-μm nylon filter (USFilters, BN-410-250) and spun for 10 min at 700*g*. Pelleted cells were washed once in warm growth medium and plated. Growth medium was replaced with fresh medium daily until dead cells and debris were not visible and cells began to grow. The percentage of FBS in the RPMI medium was 1% for sarcomas and 10% for carcinomas and stromal vascular fraction cell cultures. In vitro sensitivity to genotoxic treatment was assessed using the cell viability indicator PrestoBlue reagent (Molecular Probes). γH2AX foci formation assay was performed as follows: MDSTs were cultured and treated with 2.5 μM irinotecan, cisplatin, or talazoparib (BMN-673) for 24 hours. Before fixation, soluble proteins were extracted using cytoskeleton buffer [Pipes (pH 6.8), 100 mM NaCl, 300 mM sucrose, 3 mM $MgCl_2$, 1 mM EGTA, and 0.5% Triton X-100] for 5 min on ice. Cells were then washed in phosphate-buffered saline (PBS), permeabilized, and incubated with mouse anti-γH2AX (1:2000; Millipore, clone JBW301) followed by anti-mouse Alexa Fluor 488 (Invitrogen). Coverslips were sealed using VECTASHIELD with DAPI (4′,6-diamidino-2-phenylindole) to counterstain the nuclei. Images were

collected using a Nikon Eclipse TI inverted microscope at ×40 magnification. All images were quantified using custom-written MATLAB algorithms. Briefly, nuclei were segmented by Otsu's method, and any overlapping nuclei (defined by high total area and low solidity) were segmented by a watershed transform. Individual γH2AX foci were segmented using a band-pass filter. The intensity of foci was integrated for each nucleus and normalized to the total nuclear area.

RPPA and TCGA analysis

Protein expression data were collected by RPPA technology (38), and spot signal intensity was processed by the R package SuperCurve (39), which generates the relative $\log_2$ expression value for each protein. The $\log_2$ expression data were then corrected by median-centering protein expression across the samples and then across antibodies. After data correction, a two-way unsupervised hierarchical clustering heat map was generated using the Next-Generation Clustered Heat Map (NG-CHM) tool (http://bioinformatics.mdanderson.org/main/NG-CHM: Overview).

Targeted treatment trials

Targeted agents trametinib (GSK1120212) and talazoparib were obtained through the Stand Up to Cancer (SU2C) pharmacy free request agreement (Stand Up to Cancer Dream Team Translational Research Grant no. SU2C-AACR-DT0209). Trametinib and talazoparib were dissolved in a solution composed of 0.5% hydroxypropylmethylcellulose (Sigma), 0.2% Tween 80, and 2% DMSO in distilled water and administered daily via oral gavage at concentrations of 3 and 0.33 mg/kg per day, respectively. Treatment was initiated after transplants reached palpable size. In a second set of trials, the growth of MDSTs was monitored during a 4-week treatment with MSC2285264, a potent and selective ATX inhibitor provided by Merck KGaA. The compound was dissolved at a concentration of 300 mg/ml in a solution composed of 15% ethanol, 30% polyethylene glycol 400, 50% DMSO, and 5% PBS (pH 7.4) and delivered via ALZET osmotic pumps (model 2004; filling volume, 200 μl; pump rate, 0.25 μl/hour). Tumor fragments were transplanted into recipient mice 4 days after osmotic pump implantation (Fig. 2D).

RNA-seq and DNA sequencing whole-exome data analysis

RNA-seq paired-end reads of 18 samples collected from nine MDST models (2 samples per model) and 2 normal murine mammary gland samples were mapped to the mouse reference genome (NCBIM37) using TopHat2 (v2.0.10) (40). The mouse annotation file (Ensembl NCBIM37 release 66) was used to guide spliced alignment of known transcripts. Only uniquely mapped reads were kept for further analysis. We used the count-based methods edgeR (41) and DESeq (42) to analyze differentially expressed genes. The protocol by Anders et al. (43) was used to run the analysis, whereas HTSeq count (44) was used to count each transcript. A standard CPM, representing mapped read counts normalized by the number of sequenced fragments multiplied by 1 million, was calculated for each gene. Genes with CPM ≤ 1 were filtered out, and hierarchical clustering of mouse tumor samples was computed on the basis of expressed genes (CPM > 3). Euclidean distance calculation and a tree cut at 0.55 were applied to the dendrogram to reflect the relationship among the samples. To generate a thorough depiction of sample relationships, each sample was mapped into a two-dimensional MDS plot on the basis of $\log_2$ fold change (logFC) differences between sample pairs independently computed from the same expression data set. Human transgene analysis (Fig. 2B) was performed by generating a reference with the transcripts of the transgenes LPAR1, LPAR2, LPAR3, and ENPP2 (ATX), to which all sequenced RNA-seq data were mapped. Whole-exome DNA sequencing data from one mammary gland and samples from six MDSTs were mapped to the mouse reference genome (NCBIM37) using BWA-MEM (45). PCR replicates were removed using the SAMTools rmdup command, whereas all raw variants were called using SAMTools mpileup and bcftools (46). First, to identify somatic events, raw variants from each sample were annotated using an in-house tool. Next, variants from each tumor sample were filtered against the variant list of the mammary gland sample, and the remaining variants from each sample were further filtered against the mouse dbSNP database (mm9 version SNP132). Last, only variants with a minimum depth of 20, a maximum depth of 500, and a mapping quality over 30 were considered as somatic events. Only nonsynonymous substitutions and indels were retained.

EMT scoring

EMT scores shown in table S1 were calculated for each MDST by adding relative expression values of Ctnbb1 (β-catenin), Cdh1 (E-cadherin), Cldn7 (claudin-7), Mir200a, Snai1, Twist1, Tgfb1, and Zeb1 genes. Negative signs were assigned to Ctnbb1, Cdh1, Cldn7, and Mir200a, whereas positive signs were assigned to Snai1, Twist1, Tgfb1, and Zeb1 expression values according to the functional or correlative relationship that the gene has with EMT-related processes.

Clustering analysis of GEMMs

RNA-seq data of 18 samples collected from nine MDST models ($n = 2$ per MDST) and 2 normal murine mammary gland tissue samples were aligned to the National Center for Biotechnology Information (NCBI) build37 (hg19) human reference genome using TopHat2 v2.0.10 (47). The abundance of genes was quantified by Cufflinks v2.1.1 (48). A murine tumor data set of 385 DNA microarrays was downloaded from GSE42640 (9), which represents 27 GEMMs. Supervised clustering was applied to 405 mouse samples on the basis of the expression of murine intrinsic genes. The data were imputed and normalized, as described by Pfefferle et al. (9). Of these, 20 microarrays from the original ATX-LPAR murine model (11) previously incorporated in the Pfefferle et al. mega set and the 18 RNA-seq lanes corresponding to MDST samples were used as anchors for interbatch adjustment. Each tumor was standardized to have the normal distribution [N (0, 1)] across 8731 genes found in both data sets. For each gene, an adjustment factor was computed to account for the different median expression level in the anchor samples of the two data sets and was then applied to the complete set. Intrinsic gene analysis was conducted in the merged data set using a list of 1841 intrinsic genes from Pfefferle et al. (9), of which 1655 genes were found. The hierarchical clustering was performed using centroid linkage in Gene Cluster 3.0 (49) and was viewed with Java TreeView 1.1.6r4 (50).

PAM50 subtyping analysis

Clustering analysis of PAM50 intrinsic genes was performed in the merged set of 405 mouse samples using the 50-gene human breast subtype classifier PAM50 (32). Forty-seven of 50 homologs of human genes were detected in the mouse samples using the Mouse Genome Database (51). The hierarchical clustering and visualization were performed with the same approach used for the mouse intrinsic gene analysis. For across-species comparison, we analyzed the expression patterns of the 232 human breast samples that were used as PAM50 prototype samples from the training set (32).

Cell line sensitivity study

Olaparib (AZD2281) used in the vitro experiments in fig. S7D was obtained from AstraZeneca. Drug sensitivity data of olaparib and veliparib (ABT-888) for 76 women's cancer cell lines from breast, ovary, and endometrium were downloaded from the Broad Institute's Cancer Therapeutics Response Portal v2 (www.broadinstitute.org/ctrp.v2/). Cells were stratified into sensitive, intermediate, and resistant groups based on the tertile of AUC values. PARP1 protein levels of a subset of these cell lines were measured by RPPA, expressed as log-transformed linear normalized values (see "RPPA and TCGA analysis"), and associated with drug sensitivity using Student's t test. For olaparib sensitivity studies in SCLC, cells were treated with olaparib (0 to 10 µM), and proliferation was measured by CellTiter-Glo assay (Promega Corporation) after 10 days. PARP expression was quantified in lung cancer cell lines by RPPA, as described previously (36). PARP expression and olaparib median inhibitory concentration (IC_{50}) data were compared as continuous variables by Pearson correlation.

For analysis of talazoparib (Selleck) and olaparib (AstraZeneca) sensitivity of TNBC cell lines, drugs were serially diluted threefold (seven dilutions in DMSO) and stored in deep-well "master plates" at −20°C. Cancer cell lines involved in this assay were verified by short tandem repeats (Cancer Center Support Grant Characterized Cell Line Core in MD Anderson Cancer Center). The mutation status of each cell line was characterized by the T200 platform deep targeted sequencing performed in the Institute for Personalized Cancer Therapy (MD Anderson Cancer Center). Cell lines were seeded in 96-well plate at 2500 cells/100 µl per well and maintained in their optimal growth medium (5% FBS) for 24 hours. Following a serum deprivation step (overnight incubation in 2% FBS medium), serially diluted drug stocks were added and cells were incubated for 7 days. Triplicates were performed for each concentration, and 0.1% DMSO was used as control. Cell viability was determined at the end point by CellTiter-Blue Cell Viability Assay (Promega Corporation; 5 µl of the reagent per well). Cellular sensitivity was defined by 50% growth inhibition (GI_{50}) concentration calculated according to the cell viability curve.

Expression of PARP1 protein was determined in TNBC cell lines by RPPA analysis [retrieved from the MD Anderson Cell Lines Project (http://ibl.mdanderson.org/mclp/#/)]. To determine the correlation of PARP1 protein levels with sensitivity to PARP inhibitors, the GI_{50} values for both talazoparib and olaparib were log_{10}-transformed and then converted to z scores. PARP inhibitor sensitivity was determined as the average z score for both compounds and was used to determine the Pearson correlation coefficient with the PARP1 protein expression levels.

Homologous recombination defect signature score

To predict if tumors had defects in homologous recombination repair pathways, we used a previously published gene expression signature that has been shown to predict HRDs in breast, renal, lung, ovarian, and prostate cancers (26). Of the 230 genes in this original signature, 180 had murine orthologs in the MDST's RNA-seq data set. Gene expression levels for these 180 genes were then used to determine a correlation coefficient with HRD score (52).

PSP score

To generate a predictor of PARP sensitivity on the basis of gene expression levels, olaparib and rucaparib (AG014699) IC_{50} values for a variety of cell lines were obtained from the COSMIC database (24). These values were converted to z scores and then averaged to produce an overall measure of PARP sensitivity. Next, corresponding gene expression levels were obtained from the Cancer Cell Line Encyclopedia (25) for genes with expression levels evaluated during RNA-seq experiments. After reserving a subset of 16 breast cancer cell lines for testing, a grid search algorithm was used to identify optimal P value and fold change cutoffs between sensitive cell lines (average z score < −1.0) and nonsensitive cell lines (average z score > 0) by maximizing the average area under the ROC curve for the olaparib and rucaparib sensitivity data sets. Verification of this gene expression signature in the 16 reserved breast cancer cell lines resulted in an 88% accuracy rate. This signature was then correlated with the RNA-seq results to determine PARP sensitivity. For analysis of patient breast cancer samples, level 3 gene expression data were obtained for the 511 breast cancer patients from the TCGA Data Portal (4, 21). These expression data were used to classify patients into intrinsic subtypes using the PAM50 model (32).

Statistics

Unpaired two-tailed Student's t test and Kruskal-Wallis test by ranks with Dunn's correction were used to determine the significance of group differences. Values were considered significantly different at $P < 0.05$. The goodness of fit between observed and predicted incidence of two categorical variables was tested by binomial test and accepted as significant at $P \leq 0.01$.

SUPPLEMENTARY MATERIALS

REFERENCES AND NOTES

1. R. D. Baird, C. Caldas, Genetic heterogeneity in breast cancer: The road to personalized medicine? *BMC Med.* **11**, 151 (2013).
2. I. A. Mayer, V. G. Abramson, B. D. Lehmann, J. A. Pietenpol, New strategies for triple-negative breast cancer—Deciphering the heterogeneity. *Clin. Cancer Res.* **20**, 782–790 (2014).
3. V. Almendro, G. Fuster, Heterogeneity of breast cancer: Etiology and clinical relevance. *Clin. Transl. Oncol.* **13**, 767–773 (2011).
4. Cancer Genome Atlas Network, Comprehensive molecular portraits of human breast tumours. *Nature* **490**, 61–70 (2012).
5. A. Prat, C. M. Perou, Deconstructing the molecular portraits of breast cancer. *Mol. Oncol.* **5**, 5–23 (2011).
6. J. Stingl, C. Caldas, Molecular heterogeneity of breast carcinomas and the cancer stem cell hypothesis. *Nat. Rev. Cancer* **7**, 791–799 (2007).
7. A. Aytes, A. Mitrofanova, C. Lefebvre, M. J. Alvarez, M. Castillo-Martin, T. Zheng, J. A. Eastham, A. Gopalan, K. J. Pienta, M. M. Shen, A. Califano, C. Abate-Shen, Cross-species regulatory network analysis identifies a synergistic interaction between *FOXM1* and *CENPF* that drives prostate cancer malignancy. *Cancer Cell* **25**, 638–651 (2014).
8. A. Mitrofanova, A. Aytes, M. Zou, M. M. Shen, C. Abate-Shen, A. Califano, Predicting drug response in human prostate cancer from preclinical analysis of in vivo mouse models. *Cell Rep.* **12**, 2060–2071 (2015).

9. A. D. Pfefferle, J. I. Herschkowitz, J. Usary, J. C. Harrell, B. T. Spike, J. R. Adams, M. I. Torres-Arzayus, M. Brown, S. E. Egan, G. M. Wahl, J. M. Rosen, C. M. Perou, Transcriptomic classification of genetically engineered mouse models of breast cancer identifies human subtype counterparts. *Genome Biol.* **14**, R125 (2013).

10. J. Usary, W. Zhao, D. Darr, P. J. Roberts, M. Liu, L. Balletta, O. Karginova, J. Jordan, A. Combest, A. Bridges, A. Prat, M. C. U. Cheang, J. I. Herschkowitz, J. M. Rosen, W. Zamboni, N. E. Sharpless, C. M. Perou, Predicting drug responsiveness in human cancers using genetically engineered mice. *Clin. Cancer Res.* **19**, 4889–4899 (2013).

11. S. Liu, M. Umezu-Goto, M. Murph, Y. Lu, W. Liu, F. Zhang, S. Yu, L. C. Stephens, X. Cui, G. Murrow, K. Coombes, W. Muller, M. C. Hung, C. M. Perou, A. V. Lee, X. Fang, G. B. Mills, Expression of autotaxin and lysophosphatidic acid receptors increases mammary tumorigenesis, invasion, and metastases. *Cancer Cell* **15**, 539–550 (2009).

12. Y. C. Yung, N. C. Stoddard, J. Chun, LPA receptor signaling: Pharmacology, physiology, and pathophysiology. *J. Lipid Res.* **55**, 1192–1214 (2014).

13. W. H. Moolenaar, A. Perrakis, Insights into autotaxin: How to produce and present a lipid mediator. *Nat. Rev. Mol. Cell Biol.* **12**, 674–679 (2011).

14. S. Liu, M. Murph, N. Panupinthu, G. B. Mills, ATX-LPA receptor axis in inflammation and cancer. *Cell Cycle* **8**, 3695–3701 (2009).

15. D. G. DeNardo, L. M. Coussens, Inflammation and breast cancer. Balancing immune response: Crosstalk between adaptive and innate immune cells during breast cancer progression. *Breast Cancer Res.* **9**, 212 (2007).

16. M. Shaheen, C. Allen, J. A. Nickoloff, R. Hromas, Synthetic lethality: Exploiting the addiction of cancer to DNA repair. *Blood* **117**, 6074–6082 (2011).

17. M. Rouleau, A. Patel, M. J. Hendzel, S. H. Kaufmann, G. G. Poirier, PARP inhibition: PARP1 and beyond. *Nat. Rev. Cancer* **10**, 293–301 (2010).

18. J. Murai, S.-Y. Huang, A. Renaud, Y. Zhang, J. Ji, S. Takeda, J. Morris, B. Teicher, J. H. Doroshow, Y. Pommier, Stereospecific PARP trapping by BMN 673 and comparison with olaparib and rucaparib. *Mol. Cancer Ther.* **13**, 433–443 (2014).

19. P. Domagala, J. Hybiak, J. Rys, T. Byrski, C. Cybulski, J. Lubinski, Pathological complete response after cisplatin neoadjuvant therapy is associated with the downregulation of DNA repair genes in *BRCA1*-associated triple-negative breast cancers. *Oncotarget* **7**, 68662–68673 (2016).

20. T. Sørlie, C. M. Perou, R. Tibshirani, T. Aas, S. Geisler, H. Johnsen, T. Hastie, M. B. Eisen, M. van de Rijn, S. S. Jeffrey, T. Thorsen, H. Quist, J. C. Matese, P. O. Brown, D. Botstein, P. E. Lønning, A. L. Borresen-Dale, Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc. Natl. Acad. Sci. U.S.A.* **98**, 10869–10874 (2001).

21. J. Gao, B. A. Aksoy, U. Dogrusoz, G. Dresdner, B. Gross, S. O. Sumer, Y. Sun, A. Jacobsen, R. Sinha, E. Larsson, E. Cerami, C. Sander, N. Schultz, Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci. Signal.* **6**, pl1 (2013).

22. G. V. Long, D. Stroyakovskiy, H. Gogas, E. Levchenko, F. de Braud, J. Larkin, C. Garbe, T. Jouary, A. Hauschild, J. J. Grob, V. Chiarion-Sileni, C. Lebbe, M. Mandala, M. Millward, A. Arance, I. Bondarenko, J. B. Haanen, J. Hansson, J. Utikal, V. Ferraresi, N. Kovalenko, P. Mohr, V. Probachai, D. Schadendorf, P. Nathan, C. Robert, A. Ribas, D. J. DeMarini, J. G. Irani, S. Swann, J. J. Legos, F. Jin, B. Mookerjee, K. Flaherty, Dabrafenib and trametinib versus dabrafenib and placebo for Val600 BRAF-mutant melanoma: A multicentre, double-blind, phase 3 randomised controlled trial. *Lancet* **386**, 444–451 (2015).

23. J. S. Brown, S. B. Kaye, T. A. Yap, PARP inhibitors: The race is on. *Br. J. Cancer* **114**, 713–715 (2016).

24. M. J. Garnett, E. J. Edelman, S. J. Heidorn, C. D. Greenman, A. Dastur, K. W. Lau, P. Greninger, I. R. Thompson, X. Luo, J. Soares, Q. Liu, F. Iorio, D. Surdez, L. Chen, R. J. Milano, G. R. Bignell, A. T. Tam, H. Davies, J. A. Stevenson, S. Barthorpe, S. R. Lutz, F. Kogera, K. Lawrence, A. McLaren-Douglas, X. Mitropoulos, T. Mironenko, H. Thi, L. Richardson, W. Zhou, F. Jewitt, T. Zhang, P. O'Brien, J. L. Boisvert, S. Price, W. Hur, W. Yang, X. Deng, A. Butler, H. G. Choi, J. W. Chang, J. Baselga, I. Stamenkovic, J. A. Engelman, S. V. Sharma, O. Delattre, J. Saez-Rodriguez, N. S. Gray, J. Settleman, P. A. Futreal, D. A. Haber, M. R. Stratton, S. Ramaswamy, U. McDermott, C. H. Benes, Systematic identification of genomic markers of drug sensitivity in cancer cells. *Nature* **483**, 570–575 (2012).

25. J. Barretina, G. Caponigro, N. Stransky, K. Venkatesan, A. A. Margolin, S. Kim, C. J. Wilson, J. Léhar, G. V. Kryukov, D. Sonkin, A. Reddy, M. Liu, L. Murray, M. F. Berger, J. E. Monahan, P. Morais, J. Meltzer, A. Korejwa, J. Jané-Valbuena, F. A. Mapa, J. Thibault, E. Bric-Furlong, P. Raman, A. Shipway, I. H. Engels, J. Cheng, G. K. Yu, J. Yu, P. Aspesi Jr., M. de Silva, K. Jagtap, M. D. Jones, L. Wang, C. Hatton, E. Palescandolo, S. Gupta, S. Mahan, C. Sougnez, R. C. Onofrio, T. Liefeld, L. MacConaill, W. Winckler, M. Reich, N. Li, J. P. Mesirov, S. B. Gabriel, G. Getz, K. Ardlie, V. Chan, V. E. Myer, B. L. Weber, J. Porter, M. Warmuth, P. Finan, J. L. Harris, M. Meyerson, T. R. Golub, M. P. Morrissey, W. R. Sellers, R. Schlegel, L. A. Garraway, The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature* **483**, 603–607 (2012).

26. G. Peng, C. Chun-Jen Lin, W. Mo, H. Dai, Y.-Y. Park, S. M. Kim, Y. Peng, Q. Mo, S. Siwko, R. Hu, J.-S. Lee, B. Hennessy, S. Hanash, G. B. Mills, S.-Y. Lin, Genome-wide

27. F. Ades, D. Zardavas, I. Bozovic-Spasojevic, L. Pugliano, D. Fumagalli, E. de Azambuja, G. Viale, C. Sotiriou, M. Piccart, Luminal B breast cancer: Molecular characterization, clinical management, and future perspectives. *J. Clin. Oncol.* **32**, 2794–2803 (2014).

28. F. Bertucci, P. Finetti, D. Birnbaum, Basal breast cancer: A complex and deadly molecular subtype. *Curr. Mol. Med.* **12**, 96–110 (2012).

29. L. M. Coussens, Z. Werb, Inflammation and cancer. *Nature* **420**, 860–867 (2002).

30. A. Kreso, C. A. O'Brien, P. van Galen, O. I. Gan, F. Notta, A. M. Brown, K. Ng, J. Ma, E. Wienholds, C. Dunant, A. Pollett, S. Gallinger, J. McPherson, C. G. Mullighan, D. Shibata, J. E. Dick, Variable clonal repopulation dynamics influence chemotherapy response in colorectal cancer. *Science* **339**, 543–548 (2013).

31. E. McKinnon, P. Xiao, Metaplastic carcinoma of the breast. *Arch. Pathol. Lab. Med.* **139**, 819–822 (2015).

32. J. S. Parker, M. Mullins, M. C. U. Cheang, S. Leung, D. Voduc, T. Vickery, S. Davies, C. Fauron, X. He, Z. Hu, J. F. Quackenbush, I. J. Stijleman, J. Palazzo, J. S. Marron, A. B. Nobel, E. Mardis, T. O. Nielsen, M. J. Ellis, C. M. Perou, P. S. Bernard, Supervised risk predictor of breast cancer based on intrinsic subtypes. *J. Clin. Oncol.* **27**, 1160–1167 (2009).

33. T. Sørlie, R. Tibshirani, J. Parker, T. Hastie, J. S. Marron, A. Nobel, S. Deng, H. Johnsen, R. Pesich, S. Geisler, J. Demeter, C. M. Perou, P. E. Lønning, P. O. Brown, A.-L. Børresen-Dale, D. Botstein, Repeated observation of breast tumor subtypes in independent gene expression data sets. *Proc. Natl. Acad. Sci. U.S.A.* **100**, 8418–8423 (2003).

34. A. Montoni, M. Robu, É. Pouliot, G. M. Shah, Resistance to PARP-inhibitors in cancer therapy. *Front. Pharmacol.* **4**, 18 (2013).

35. R. J. Cardnell, Y. Feng, L. Diao, Y.-H. Fan, F. Masrorpour, J. Wang, Y. Shen, G. B. Mills, J. D. Minna, J. V. Heymach, L. A. Byers, Proteomic markers of DNA repair and PI3K pathway activation predict response to the PARP inhibitor BMN 673 in small cell lung cancer. *Clin. Cancer Res.* **19**, 6322–6328 (2013).

36. L. A. Byers, J. Wang, M. B. Nilsson, J. Fujimoto, P. Saintigny, J. Yordy, U. Giri, M. Peyton, Y. H. Fan, L. Diao, F. Masrorpour, L. Shen, W. Liu, B. Duchemann, P. Tumula, V. Bhardwaj, J. Welsh, S. Weber, B. S. Glisson, N. Kalhor, I. I. Wistuba, L. Girard, S. M. Lippman, G. B. Mills, K. R. Coombes, J. N. Weinstein, J. D. Minna, J. V. Heymach, Proteomic profiling identifies dysregulated pathways in small cell lung cancer and novel therapeutic targets including PARP1. *Cancer Discov.* **2**, 798–811 (2012).

37. J. B. Dennison, J. R. Molina, S. Mitra, A. M. Gonzalez-Angulo, J. M. Balko, M. G. Kuba, M. E. Sanders, J. A. Pinto, H. L. Gómez, C. L. Arteaga, R. E. Brown, G. B. Mills, Lactate dehydrogenase B: A metabolic marker of response to neoadjuvant chemotherapy in breast cancer. *Clin. Cancer Res.* **19**, 3703–3713 (2013).

38. M. S. Carey, R. Agarwal, B. Gilks, K. Swenerton, S. Kalloger, J. Santos, Z. Ju, Y. Lu, F. Zhang, K. Coombes, D. Miller, D. Huntsman, G. B. Mills, B. T. Hennessy, Functional proteomic analysis of advanced serous ovarian cancer using reverse phase protein array: TGF-β pathway signaling indicates response to primary chemotherapy. *Clin. Cancer Res.* **16**, 2852–2860 (2010).

39. J. Hu, X. He, K. A. Baggerly, K. R. Coombes, B. T. Hennessy, G. B. Mills, Non-parametric quantification of protein lysate arrays. *Bioinformatics* **23**, 1986–1994 (2007).

40. D. Kim, G. Pertea, C. Trapnell, H. Pimentel, R. Kelley, S. L. Salzberg, TopHat2: Accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.* **14**, R36 (2013).

41. M. D. Robinson, D. J. McCarthy, G. K. Smyth, edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).

42. S. Anders, W. Huber, Differential expression analysis for sequence count data. *Genome Biol.* **11**, R106 (2010).

43. S. Anders, D. J. McCarthy, Y. Chen, M. Okoniewski, G. K. Smyth, W. Huber, M. D. Robinson, Count-based differential expression analysis of RNA sequencing data using R and Bioconductor. *Nat. Protoc.* **8**, 1765–1786 (2013).

44. S. Anders, P. T. Pyl, W. Huber, HTSeq—A Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166–169 (2015).

45. H. Li, Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv: 13033997 (2013).

46. H. Li, B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin; 1000 Genome Project Data Processing Subgroup, The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).

47. C. Trapnell, L. Pachter, S. L. Salzberg, TopHat: Discovering splice junctions with RNA-Seq. *Bioinformatics* **25**, 1105–1111 (2009).

48. C. Trapnell, B. A. Williams, G. Pertea, A. Mortazavi, G. Kwan, M. J. van Baren, S. L. Salzberg, B. J. Wold, L. Pachter, Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* **28**, 511–515 (2010).

49. M. J. L. de Hoon, S. Imoto, J. Nolan, S. Miyano, Open source clustering software. *Bioinformatics* **20**, 1453–1454 (2004).

transcriptome profiling of homologous recombination DNA repair. *Nat. Commun.* **5**, 3361 (2014).

50. A. J. Saldanha, Java TreeView—Extensible visualization of microarray data. *Bioinformatics* **20**, 3246–3248 (2004).

51. J. T. Eppig, J. A. Blake, C. J. Bult, J. A. Kadin, J. E. Richardson; Mouse Genome Database Group, The Mouse Genome Database (MGD): Facilitating mouse as a model for human biology and disease. *Nucleic Acids Res.* **43**, D726–D736 (2015).

52. M. J. van de Vijver, Y. D. He, L. J. van't Veer, H. Dai, A. A. Hart, D. W. Voskuil, G. J. Schreiber, J. L. Peterse, C. Roberts, M. J. Marton, M. Parrish, D. Atsma, A. Witteveen, A. Glas, L. Delahaye, T. van der Velde, H. Bartelink, S. Rodenhuis, E. T. Rutgers, S. H. Friend, R. Bernards, A gene-expression signature as a predictor of survival in breast cancer. *N. Engl. J. Med.* **347**, 1999–2009 (2002).