

ARTICLE OPEN



Identification of biomarkers of response to preoperative talazoparib monotherapy in treatment naïve gBRCA+ breast cancers

Xuan Liu^{1,11}, Zhongqi Ge^{2,11}, Fei Yang³, Alejandro Contreras⁴, Sanghoon Lee⁵, Jason B. White⁶, Yiling Lu⁷, Marilyn Labrie⁸, Banu K. Arun⁶, Stacy L. Moulder⁶, Gordon B. Mills⁸, Helen Piwnica-Worms⁹, Jennifer K. Litton⁶ and Jeffrey T. Chang^{1,10}✉

Germline mutations in *BRCA1* or *BRCA2* exist in ~2–7% of breast cancer patients, which has led to the approval of PARP inhibitors in the advanced setting. We have previously reported a phase II neoadjuvant trial of single agent talazoparib for patients with germline *BRCA* pathogenic variants with a pathologic complete response (pCR) rate of 53%. As nearly half of the patients treated did not have pCR, better strategies are needed to overcome treatment resistance. To this end, we conducted multi-omic analysis of 13 treatment naïve breast cancer tumors from patients that went on to receive single-agent neoadjuvant talazoparib. We looked for biomarkers that were predictive of response (assessed by residual cancer burden) after 6 months of therapy. We found that all resistant tumors exhibited either the loss of SHLD2, expression of a hypoxia signature, or expression of a stem cell signature. These results indicate that the deep analysis of pre-treatment tumors can identify biomarkers that are predictive of response to talazoparib and potentially other PARP inhibitors, and provides a framework that will allow for better selection of patients for treatment, as well as a roadmap for the development of novel combination therapies to prevent emergence of resistance.

npj Breast Cancer (2022)8:64; <https://doi.org/10.1038/s41523-022-00427-9>

INTRODUCTION

Pathogenic germline variants in *BRCA1* and *BRCA2* (gBRCA+) confer an elevated risk (57% and 49%, respectively) of developing breast cancer¹ and are found across breast cancer subtypes². Together, approximately 1.6–7% of all breast cancers have hereditary aberrations in *BRCA*^{3,4}, although up to 39% are predicted to be deficient in the homologous recombination (HR) DNA damage repair pathway where *BRCA* are key mediators⁵. PolyADP ribose polymerase (PARP) inhibitors (PARPi) were found to be synthetic lethal in cancers with aberrations in HR mediated by *BRCA*⁶ and multiple PARPi have now been approved for the treatment of several cancer subtypes, including gBRCA+ breast cancers⁷.

In gBRCA+ breast cancers, the randomized OlympiAD⁸ and EMBRACA⁹ trials both demonstrated that single agent PARPi (olaparib and talazoparib, respectively) increased response rates and progression-free survival compared to physician's choice of chemotherapy, with substantial improvements in quality of life¹⁰. While objective response rates were approximately 60% in both trials, resistance eventually developed in most patients and neither trial has shown an improvement in overall survival^{11,12}. Recently, PARPi have shown promise as single agent therapy in the neoadjuvant setting achieving a pathologic complete response (pCR) rate of 53%¹³, which is similar to that historically seen with anthracycline/taxane-based chemotherapy. Notably, nearly half of the gBRCA+ breast cancers treated with

neoadjuvant PARPi did not result in pCR; thus, primary resistance to PARP inhibition remains a substantial problem.

Several PARPi resistance mechanisms have been identified in vitro and in various mouse models. Both olaparib¹⁴ and talazoparib¹⁵ are substrates of the ABCB1 multidrug efflux channel¹⁶, which functions to decrease intracellular concentrations of the drug. PARPi resistance can also be mediated by reinstatement of HR through restoration of *BRCA* function by reversion^{17,18} or hypomorphic *BRCA* mutations¹⁹, and promotion of homologous recombination through inhibition of the shieldin complex^{20,21}, as well as by stabilization²² or repriming²³ of replication forks. The shieldin complex determines the balance between low fidelity non-homologous end joining (NHEJ) that can result in accumulation of further mutations and eventual cell death and HR that is compromised but not completely abrogated by *BRCA* mutations.

Although a range of resistance mechanisms has been demonstrated in models, only a limited number have been linked to resistance to PARPi in the clinical setting. Most commonly seen are secondary *BRCA* mutations that restore the HR repair compromised by gBRCA mutations. These mutations have been reported in breast and ovarian cancer tumors^{24–26}, and also in circulating cell-free DNA in PARPi or platinum resistant prostate²⁷, ovarian^{28–31}, and breast³⁰ cancers. While reversion of *BRCA* is the most frequently documented mechanism, other mutations that could potentially restore HR have been sporadically reported in clinical

¹Department of Integrative Biology and Pharmacology, University of Texas Health Science Center at Houston, Houston, TX, USA. ²Department of Immunology, The University of Texas MD Anderson Cancer Center, Houston, TX, USA. ³Department of Translational Molecular Pathology, The University of Texas MD Anderson Cancer Center, Houston, TX, USA. ⁴Department of Anatomical Pathology, The University of Texas MD Anderson Cancer Center, Houston, TX, USA. ⁵Department of Systems Biology, The University of Texas MD Anderson Cancer Center, Houston, TX, USA. ⁶Department of Breast Medical Oncology, The University of Texas MD Anderson Cancer Center, Houston, TX, USA. ⁷Department of Genome Medicine, The University of Texas MD Anderson Cancer Center, Houston, TX, USA. ⁸Department of Cell, Developmental, and Cancer Biology, Knight Cancer Institute, Oregon Health and Science University, Portland, OR, USA. ⁹Department of Experimental Radiation Oncology, The University of Texas MD Anderson Cancer Center, Houston, TX, USA. ¹⁰Department of Bioinformatics and Computational Biology, The University of Texas MD Anderson Cancer Center, Houston, TX, USA. ¹¹These authors contributed equally: Xuan Liu, Zhongqi Ge. ✉email: Jeffrey.T.Chang@uth.tmc.edu

samples such as *RAD51C* and *RAD51D* mutations in ovarian cancer³², or loss of *TP53BP1* and amplification of *MRE11A* in breast cancer²⁶. Reversion of *PARP* trapping may also occur; a *PARP1* mutation that can prevent trapping was seen in a resistant ovarian cancer patient³³.

Notably, most clinical mechanisms of disease resistance have been studied in more heavily pre-treated patients with disease recurrence or metastasis at the time of PARPi therapy. To identify biomarkers of response and to elucidate mechanisms underlying response and resistance for PARPi in treatment naïve *gBRCA*+ breast cancers, we performed multi-omic (RNA-Seq and whole-exome sequencing (WES)) analysis on 13 treatment-naïve *gBRCA*+ tumors from a neoadjuvant clinical trial of the single-agent PARPi talazoparib¹³. Response, determined by residual cancer burden (RCB)³⁴, was assessed after six months of treatment. Our results identified possible resistance mechanisms and suggest that biomarkers present in pre-treatment biopsies have the potential to enrich for disease response to single agent talazoparib therapy. Our data may also inform targeted therapy strategies for combination regimens.

RESULTS

Genomic profiling of treatment naïve *gBRCA*+ tumors from patients treated with single agent PARPi

As part of an IRB approved protocol, pretreatment core needle biopsies were collected from 20 treatment-naïve patients with *BRCA*-associated breast cancers who consented to receive neoadjuvant talazoparib monotherapy on clinical trial (Table 1). We excluded one patient who opted to also receive chemotherapy prior to surgery and one patient due to poor RNA data quality, leaving 18 samples in all.

We then profiled the tumors using whole-exome sequencing (WES) and RNA-sequencing. Copy number analysis with FACETS³⁵ revealed a large variation of tumor cellularity. The copy number profiles of five of the 18 tumors were not called, indicating little evidence of copy number changes (Supplementary Fig. S1). An ABSOLUTE³⁶ analysis substantiated these results, calling four of these five tumors *non-aneuploid*. The fifth produced only an implausible prediction of a tumor with 100% purity and ploidy = 1, with little evidence of copy number change (Supplementary Fig. S2). Further, these five tumors also had a significantly lower mutation burden ($p = 0.01$) and no detectable somatic mutations in *TP53*, which is almost universally mutated in *gBRCA*+ breast cancers³⁷ (Supplementary Fig. S3). The H&E (hematoxylin and eosin) stained tissue slides from these five tumors confirmed a deficiency of cancer cells in two of them (P32, P35). The remaining three had adequate cancer cell purity (50–70%). The discrepancy between the histology and genomics may be due to spatial heterogeneity within the tumor, and we opted to discard these samples due to uncertainty in the quality of the genomic profiles, leaving 13 tumors for analysis.

This cohort ($n = 13$) included 10 patients with deleterious *gBRCA1* mutations and three with deleterious *gBRCA2* mutations. Nine of the 13 (69%) *gBRCA* mutant tumors were triple negative breast cancers (TNBC). 67% (2 out of 3) of the *gBRCA2* mutant tumors were ER+, while only 20% (2 out of 10) *gBRCA1* were ER+, a frequency consistent with previous reports³⁸. RCB was assessed after patients completed six months of neoadjuvant talazoparib, and we defined response as pCR (38% of tumors) and RCB-I (15%), and resistance as RCB-II (31%) and RCB-III (15%) (Table 1).

We called single nucleotide mutations from the WES data and identified 859 non-synonymous somatic mutations with a median of 2.44 mutations per Mb. *TP53* mutations were the most common mutations identified and were present in 10 out of 13 tumors

Table 1. Description of samples.

Sample	Germline	ER Status	TP53	Response	Subtype	Mutations	Purity	Ploidy
P15	BRCA1	TNBC	G266E	RCB-III	IM	0.67	23%	1.5
P16	BRCA1	ER 5% weak, PR–	E258G	RCB-I		1.17	50%	1.8
P17	BRCA1	TNBC	C141Y	pCR	IM	0.73	48%	1.8
P18	BRCA1	TNBC	R175H	RCB-III	BL2	2.44	48%	3.1
P21	BRCA2	TNBC	17:7,579,310	pCR	M	5.22	81%	3.4
P23	BRCA1	TNBC	R175H	RCB-II	BL2	3.71	34%	1.8
P24	BRCA1	TNBC	L265P	pCR	BL1	3.74	50%	3.1
P25	BRCA1	ER+, PR+, HER2–	wt	RCB-II		2.54	63%	2.0
P26	BRCA1	TNBC	Y220C	pCR	MSL	4.11	43%	1.8
P30	BRCA2	ER+, PR+, HER2–	wt	pCR		2.14	23%	2.0
P31	BRCA1	TNBC	R175H	RCB-II	BL1	3.66	40%	2.2
P34	BRCA1	TNBC	V216M	RCB-II	BL1	1.71	21%	1.7
P36	BRCA2	ER+, PR+, HER2–	wt	RCB-I		2.39	27%	2.3
P19	BRCA1	TNBC		pCR		1.09	0%	2.0
P27	BRCA1	TNBC		RCB-II		0.78	0%	2.0
P29	BRCA1	TNBC		pCR		0.97	0%	2.0
P32	BRCA2	TNBC		pCR		0.98	0%	2.0
P35	BRCA1	ER+, PR–, HER2–		pCR		0.68	0%	2.0

Each sample contains a germline *BRCA1* or *BRCA2* mutation (*Germline* column). The *ER Status* column includes the expression of ER, PR, or HER2 based on immunohistochemistry. TNBC (for triple-negative breast cancer) indicates that the patient was negative in all three receptors. Somatic *TP53* mutations were identified by whole-exome sequencing. *wt* indicates no mutation identified. The *Response* column delineates whether the patient was responsive to treatment at surgery, assessed by RCB status. For the TNBC tumors, the TNBC *Subtype* is shown. BL1 is basal-like 1, BL2 is basal-like 2, M is mesenchymal, MSL is mesenchymal stem-like, and IM is immunomodulatory. *Mutations* quantifies the number of non-synonymous coding mutations per megabase. *Purity* contains the percent of tumor cells found in the sample, as estimated from the copy number profiles of the exome sequencing. The *Ploidy* of the tumor is also estimated from the exome sequencing.

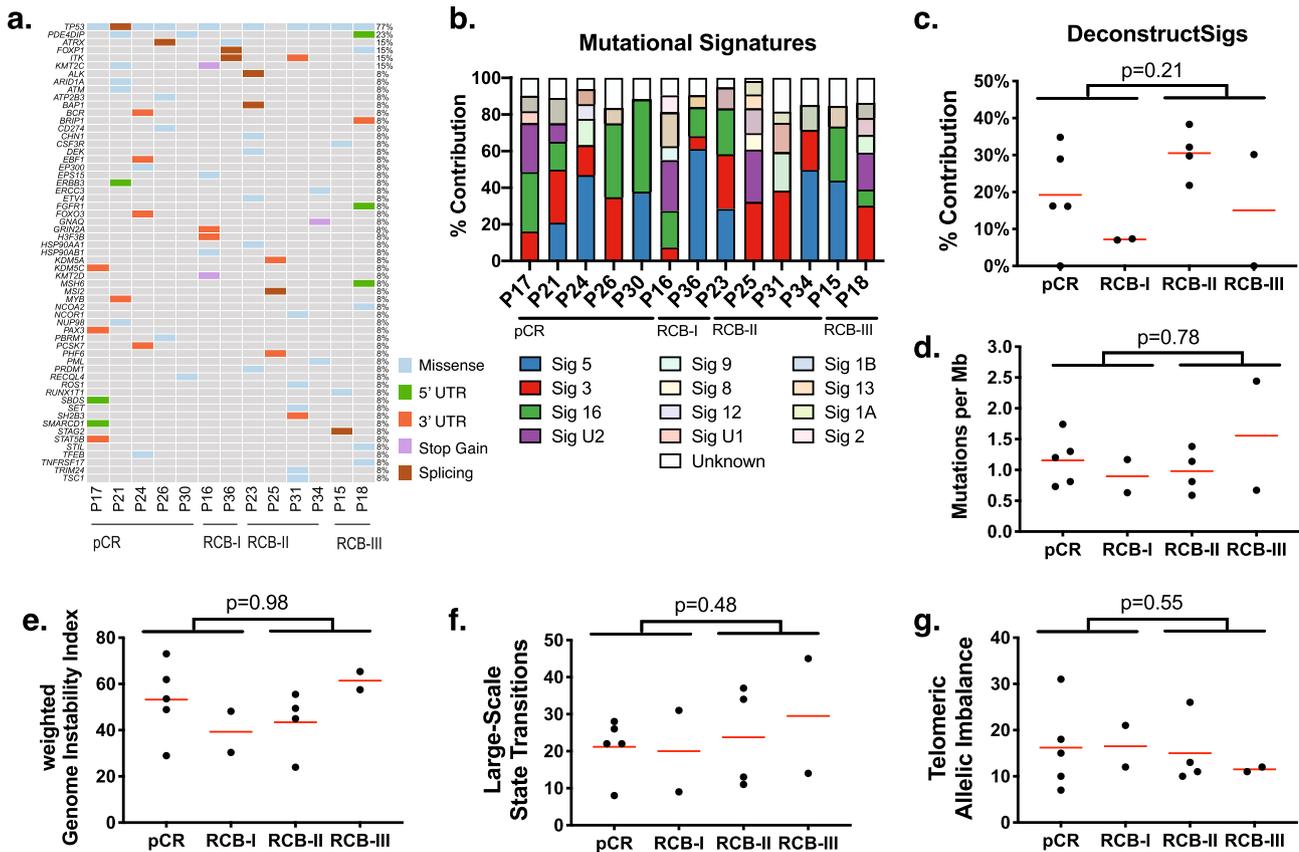


Fig. 1 Genomic profile of data set. **a** This oncoplot shows the somatic mutations in cancer-associated genes from the COSMIC database found in this cohort. Each row represents a mutated gene and each column represents a patient. The patients are grouped by RCB. The rectangles are colored according to mutation status. **b** This plot shows the predicted contribution of each mutation signature (y-axis), as predicted by DeconstructSigs, to the overall mutational processes in the tumors (columns). The signatures are distinguished by color. **c** The percent contribution of signature 3 (indicative of homologous recombination deficiency) as predicted by DeconstructSigs, are shown on the y-axis for each of the tumors. The % contribution from signature 3 is separated according to the response of each tumor. The red line indicates mean contribution, and the p -values are calculated using an unpaired two-sided Student's t test comparing pCR and RCB-I against RCB-II and RCB-III. **d** The tumor mutation burden (y-axis) is shown for each of the tumors, separated by response. The red lines indicate average mutations. The p -values are calculated as in **c**. **e** The weighted genome instability index (y-axis) is shown for each tumor, separated by response. The mean indexes are shown by red lines. The p -values are calculated as in **c**. **f** The number of large-scale state transitions (y-axis) is shown for each tumor, separated by response. The red lines indicate means. The p -values are calculated as in **c**. **g** The number of telomeres with imbalanced alleles (y-axis) is shown for each tumor, separated by response. The red lines indicate means. The p -values are calculated as in **c**.

(77%), which is comparable to prior studies showing frequent *TP53* mutations in *BRCA* mutant breast cancers (Fig. 1a)³⁹. Nine of the mutations targeted residues in the core DNA-binding domain⁴⁰, while one was predicted to affect splicing. Three tumors harbored R175H hotspot mutations⁴¹, and three others (G266E, L265P, variant 17:7,579,310) were previously documented in the COSMIC database⁴². All coding region mutations were predicted to be deleterious by both Polyphen2⁴³ and SIFT⁴⁴. No other known cancer driver genes were found to be recurrently mutated. However, there were individual cases with mutations in genes or pathways previously associated with PARPi response such as *ARID1a*, *ATM*, *BAP1*, and *TSC1*.

In addition to driver genes, mutations can also reveal the mutagenic processes that promote the acquisition of mutations throughout the progression of the tumor⁴⁵. *BRCA* mutant breast cancers have defects in double strand break repair that result in mutational patterns characterized as *signature 3*, which has previously been found to be a marker for *BRCAness* and HR deficiency^{46,47}. We confirmed that this signature is seen in this cohort of *gBRCA+* tumors using SigMA, a tool that was optimized to detect Signature 3, which predicted that this signature contributed 74–85% of the mutational patterns across these

tumors⁴⁸ (Fig. S4). However, we found no correlation between this signature and sensitivity to talazoparib ($p = 0.61$), when sensitivity was defined as pCR or RCB-I, and resistance as RCB-II or RCB-III, consistent with the endpoint of the clinical trial that generated these samples¹³, or when comparing pCR and non-PCR tumors ($p = 0.58$). In subsequent analyses, we will primarily defined sensitivity according to the definition from the trial, but also report results from the pCR and non-PCR comparison in the Supplementary Information.

To determine if other processes contribute to the mutations, we repeated the analysis with DeconstructSigs⁴⁹, which can score the contribution of a range of signatures. From this analysis, we found that the mutations across all tumors predominantly reflect signature 3, confirming the SigMA prediction, as well as signatures 5, 16, and U2 (Fig. 1b). The etiology of signatures 5 and 16 is not known, and signature U2 is suspected to be a sequencing artifact⁴⁵. The predictions of signature 3 from the two algorithms were highly correlated ($p = 0.0004$, Supplementary Fig. S4), although the absolute percentages from DeconstructSigs were lower, likely due to contributions from other signatures and artifacts. Accordingly, the predicted contributions from DeconstructSigs also failed to correlate with response (Fig. 1c,

Supplementary Fig. S5). It is possible that this signature is best suited for differentiating tumors with and without HR deficiency, and is not sensitive enough to distinguish more subtle degrees of HR competency seen in a cohort consisting entirely of *gBRCA*+ breast cancers.

We next associated tumor mutation burden with response, due to the potential for neoantigens in *BRCA*+ breast cancers to contribute to an immune response⁵⁰. However, we found no association between response to PARPi with either tumor mutation burden (Fig. 1d, Supplementary Fig. S6) or with the weighted genome instability index⁵¹, a measure of genome instability based on copy number aberrations (Fig. 1e). Similarly, we saw no associations between response and other surrogate measures of HRD including large-scale structural chromosomal aberrations, quantified as a large-scale state transition score^{52,53} (Fig. 1f, Supplementary Fig. S7); or telomeric allelic imbalance, asymmetrical copy numbers of the major and minor alleles extending to the telomere, which was predicted to result from failures in double stranded break repair⁵⁴ (Fig. 1g, Supplementary Fig. S8). Taken together, we observed no correlation between measures of HR deficiency and response to single agent talazoparib in this *gBRCA*+ cohort.

Loss of shieldin complex member SHLD2 is associated with resistance

We next searched for evidence of genomic or transcriptomic changes in 51 genes previously associated with PARPi resistance in pre-clinical models (Table 2). We started by identifying non-synonymous somatic single nucleotide mutations from the WES data (Fig. 2a). We also estimated copy numbers from this data using FACETS³⁵ and plotted amplifications or deletions that corresponded with a two-fold deviation from the average expression in responsive tumors. We found no significant structural rearrangements using Manta⁵⁵, although it is highly likely that rearrangements were not well captured in the WES data. In addition, alternative splicing changes can also lead to *BRCA* reversions⁵⁶. However, the RNA-Seq data did not reveal any *BRCA* splicing events.

The oncoplot of the variants revealed mutations in potential resistance genes in nearly every tumor (Fig. 2a). There was frequent somatic copy number deletion in *BRCA1* and *BRCA2*, likely reflecting loss of heterozygosity events⁵⁷. In the resistant tumors, alterations that may associate with resistance were seen across every patient except for P31 in genes including *SHLD2*, *RAD51*, *ATMIN*, *PTEN*, *PARG*, *MLL4*, *EZH2*, *SLFN11*, *AR*, *KRAS*, and *GPBP1*. We prioritized these genes by looking for mutations or copy number alterations that led to predicted loss or gain of function in PARPi resistant but not PARPi sensitive tumors. Four genes were identified as potential candidates mediating resistance including *SHLD2* (FAM35A), *PARG*, *MLL4* (KMT2B), and *SLFN11*. Of these, *SHLD2* was the only one altered recurrently, with copy number deletion seen in three of the six resistant tumors. *SHLD2* is a member of the shieldin complex, which mediates 53BP1-dependent non homologous end joining^{20,58}. Loss of shieldin promotes HR, leading to resistance to PARPi. Here, all tumors with *SHLD2* loss had *gBRCA1* mutations. The fact that *SHLD2* loss was not observed in *gBRCA2* mutant tumors may be due to the fact that *BRCA2* acts downstream of shieldin, and thus shieldin loss would not confer resistance in *gBRCA2*+ tumors⁵⁹.

To verify the copy number aberrations, we used an additional copy number caller Sequenza⁶⁰, which confirmed copy number loss in the tumors of patients 23 and 34, but did not recapitulate the copy number loss seen in the tumor from patient 15. We then compared the gene expression of shieldin complex proteins between PARPi resistant and sensitive tumors (Fig. 2b, Supplementary Fig. S9). Expression of *SHLD2* alone was 2.2x lower in PARPi resistant tumors ($p = 0.003$) and tumors with the lowest

SHLD2 expression also had copy number loss in at least one of the callers. We repeated this analysis across the remainder of the 51 established resistance genes, regardless of mutation or copy number change. The expression of *CDK12* and *SMARCA1* were both significantly higher in the sensitive tumors (both $p = 0.03$), but the average changes in gene expression were modest (1.2x and 1.3x, respectively) and therefore unlikely to indicate a general resistance mechanism (Supplementary Fig. S10). Thus, our analysis identified loss of *SHLD2*, either by copy number loss or reduced expression, as a potential mechanism for PARPi resistance in *gBRCA*+ tumors.

Gene expression signatures are associated with both sensitivity and resistance to PARPi

To look more broadly for markers of resistance, we performed a pathway analysis. We used a pre-ranked GSEA (prGSEA) test (see Methods) on the Hallmarks database⁶¹ to identify the pathways that correlated with response. We found that sensitive tumors were associated with activation of E2F (adjusted $p = 0.003$), MYC (adjusted $p = 0.003$), and G2/M (adjusted $p = 0.02$) (Fig. 3a, b). This could represent more proliferative tumors or alternatively tumors with premature entry into S phase and replication stress that has been associated with response to PARPi⁶². In contrast, resistant tumors had high predicted activity of hypoxia (adjusted $p = 0.0004$), EMT (adjusted $p = 0.0004$), altered metabolism (glycolysis adjusted $p = 0.003$, oxidative phosphorylation adjusted $p = 0.007$, cholesterol homeostasis adjusted $p = 0.007$), along with several other pathways.

To verify the association between MYC targets and sensitivity to talazoparib, we extracted the genes from the leading edge of the signature (i.e., the genes with the largest change in expression) and visualized their expression in a heatmap (Fig. 3c). This confirmed that nearly all the tumors that achieved pCR, with the exception of P26, had high expression of MYC target genes, while the tumors with RCB-II and -III had lower expression across those genes.

In resistant tumors, the hypoxia gene expression signature was predicted to be active. As an initial validation of this result, we scored the tumors with a validated hypoxia signature⁶³ and found that its scores were highly correlated with the Hallmarks signature from our analysis (Supplementary Fig. S11). Next, we asked whether gene expression of known mediators of hypoxia, HIF1A (a downstream transcriptional regulator of a hypoxia response) and VEGF-A (a hypoxia responsive growth factor) (reviewed in⁶⁴) had concomitant changes (Fig. 3d, Supplementary Fig. S12). VEGF-A was expressed higher in resistant tumors, although it did not achieve statistical significance ($p = 0.06$). Expression of HIF1A was not associated with response ($p = 0.53$). This could be because HIF1A is regulated post-translationally by ubiquitination, and its activity may not correlate with its gene expression levels. Therefore, we examined whether expression of HIF target genes⁶⁵ correlated with the hypoxia signature and found that HIF target genes were upregulated in resistant tumors ($p < 0.0001$) (Fig. 3e). The expression levels of HIF1 targets (seen on the leading edge of a GSEA analysis), including VEGF-A, were highly upregulated in 4 out of the 6 most aggressive tumors (Fig. 3f). This provides evidence that hypoxia is a common feature in talazoparib-resistant tumors.

An EMT signature and stem cell markers are correlated with resistance

In addition to the hypoxia signature, the pathway analysis also showed an equally strong association of epithelial mesenchymal transition (EMT) with resistance to talazoparib (Fig. 3a), which was also previously seen in PDX models of small cell lung cancer⁶⁶. To corroborate the EMT signature, we evaluated tumors for the expression of CDH1, the primary cadherin responsible for

Table 2. Mechanisms of resistance.

	Gene	Mechanism	Resistance	Reference	
DNA Damage Response	ABCB1	Drug efflux	Gain	Multiple ABCB1 transcriptional fusions in drug resistant high-grade serous ovarian and breast cancer	
	BRCA1	Promotes DNA end resection.	Gain	Secondary BRCA1 mutations in BRCA1-mutated ovarian carcinomas with platinum resistance	
	BRCA2	Promotes homologous recombination.	Gain	Secondary mutations as a mechanism of cisplatin resistance in BRCA2-mutated cancers	
	BRIP1	Binds BRCA1.	Gain	Inactivation of the Tumor Suppressor BRIP1 Gene Confers Increased Susceptibility to Platinum Antineoplastic Agents and Augments the Synergistic Response to PARP Inhibition in Ovarian Epithelial Cells	
	TP53BP1	Inhibits DNA end resection.	Loss	53BP1 inhibits homologous recombination in Brca1-deficient cells by blocking resection of DNA breaks	
	RIF1	Inhibits DNA end resection.	Loss	53BP1 regulates DSB repair using Rif1 to control 5' end resection	
	SHLD1 (C20orf196)	Inhibits DNA end resection. Shieldin complex.	Loss	DNA repair network analysis reveals Shieldin as a key regulator of NHEJ and PARP inhibitor sensitivity	
	SHLD2 (FAM35A)	Inhibits DNA end resection. Shieldin complex.	Loss	DNA repair network analysis reveals Shieldin as a key regulator of NHEJ and PARP inhibitor sensitivity	
	SHLD3	Inhibits DNA end resection. Shieldin complex.	Loss	DNA repair network analysis reveals Shieldin as a key regulator of NHEJ and PARP inhibitor sensitivity	
	REV7/MAD2L2	Inhibits DNA end resection. Shieldin complex.	Loss	MAD2L2 controls DNA repair at telomeres and DNA breaks by inhibiting 5' end resection	
	TRIP13	Promotes dissociation of REV7-Shieldin complex.	Gain	TRIP13 regulates DNA repair pathway choice through REV7 conformational change	
	STN1	Inhibits DNA end resection. CST Complex.	Loss	The CST complex mediates end protection at double-strand breaks and promotes PARP inhibitor sensitivity in BRCA1-deficient cells	
	CTC1	Inhibits DNA end resection. CST Complex.	Loss	The CST complex mediates end protection at double-strand breaks and promotes PARP inhibitor sensitivity in BRCA1-deficient cells	
	TEN1	Inhibits DNA end resection. CST Complex.	Loss	The CST complex mediates end protection at double-strand breaks and promotes PARP inhibitor sensitivity in BRCA1-deficient cells	
	RAD51	Stimulates strand invasion. Stabilizes replication forks.	Gain	Secondary somatic mutations restoring RAD51C and RAD51D associated with acquired resistance to the PARP inhibitor rucaparib in high-grade ovarian carcinoma	
	PALB2	Stimulates strand invasion.	Gain	Cooperation of breast cancer proteins PALB2 and piccolo BRCA2 in stimulating homologous recombination	
	SHFM1	Stimulates strand invasion.	Gain	Cooperation of breast cancer proteins PALB2 and piccolo BRCA2 in stimulating homologous recombination	
	CDK12	Promotes homologous recombination.	Gain	Ovarian cancer-associated mutations disable catalytic activity of CDK12, a kinase that promotes homologous recombination repair and resistance to cisplatin and poly(ADP-ribose) polymerase inhibitors	
	Replication Fork	HELB	Inhibits DNA end resection.	Loss	HELB is a feedback inhibitor of DNA end resection
		DYNLL1	Inhibits DNA end resection.	Loss	DYNLL1 binds to MRE11 to limit DNA end resection in BRCA1-deficient cells
ATMIN		Transcriptionally activates DYNLL1.	Loss	DYNLL1 binds to MRE11 to limit DNA end resection in BRCA1-deficient cells	
MRE11		Promotes end resection and replication fork degradation.	Unclear ^a	Double-strand break repair-independent role for BRCA2 in blocking stalled replication fork degradation by MRE11	
PTEN		Controversial.	Loss		
PIK3CA		PI3K controls DSB repair.	Gain	PI3K inhibition impairs BRCA1/2 expression and sensitizes BRCA-proficient triple-negative breast cancer to PARP inhibition	
PIK3CB		PI3K controls DSB repair.	Gain	PI3K inhibition impairs BRCA1/2 expression and sensitizes BRCA-proficient triple-negative breast cancer to PARP inhibition	
AKT1		PI3K controls DSB repair.	Gain	PI3K inhibition impairs BRCA1/2 expression and sensitizes BRCA-proficient triple-negative breast cancer to PARP inhibition	
POLQ		Catalyzes MMEJ, promoting cell survival in HR-deficient cells.	Gain	Polymerase theta inhibition kills homologous recombination deficient tumors	
PARP1		Promotes PARP trapping.	Loss	A genetic screen using the PiggyBac transposon in haploid cells identifies Parp1 as a mediator of olaparib toxicity	
PARG		Inhibits PARP parylation.	Loss	Selective loss of PARG restores PARYlation and counteracts PARP inhibitor-mediated synthetic lethality	
PTIP (PAXIP1)		Promotes replication fork degradation.	Loss	Replication fork stability confers chemoresistance in BRCA-deficient cells	

Table 2 continued

Gene	Mechanism	Resistance	Reference
MLL3	Promotes replication fork degradation.	Loss	Replication fork stability confers chemoresistance in BRCA-deficient cells
MLL4	Promotes replication fork degradation.	Loss	Replication fork stability confers chemoresistance in BRCA-deficient cells
EZH2	Promotes replication fork degradation.	Loss	EZH2 promotes degradation of stalled replication forks by recruiting MUS81 through histone H3 trimethylation
SMARCAL1	Promotes replication fork degradation.	Loss	Restoration of replication fork stability in BRCA1- and BRCA2-deficient cells by inactivation of SNF2-family fork remodelers
ZRANB3	Promotes replication fork degradation.	Loss	Restoration of replication fork stability in BRCA1- and BRCA2-deficient cells by inactivation of SNF2-family fork remodelers
HLTF	Promotes replication fork degradation.	Loss	Restoration of replication fork stability in BRCA1- and BRCA2-deficient cells by inactivation of SNF2-family fork remodelers
SLFN11	Stalls stressed replication forks.	Loss	Resistance to PARP inhibitors by SLFN11 inactivation can be overcome by ATR inhibition
CHD4	Inhibits translesion synthesis.	Loss ¹	Resistance to therapy in BRCA2 mutant cells due to loss of the nucleosome remodeling factor CHD4
RADX	Inhibits RAD51 at replication forks.	Loss	RADX promotes genome stability and modulates chemosensitivity by regulating RAD51 at replication forks
TLK1	Stabilizes replication forks.	Gain	Tousled-like kinases stabilize replication forks and show synthetic lethality with checkpoint and PARP inhibitors
TLK2	Stabilizes replication forks.	Gain	Tousled-like kinases stabilize replication forks and show synthetic lethality with checkpoint and PARP inhibitors
PRIMPOL	Promotes repriming of stalled forks.	Gain	PRIMPOL-mediated adaptive response suppresses replication fork reversal in BRCA-deficient cells
Other			
AR	Regulates HR genes.	Gain	Androgen receptor inhibitor enhances the antitumor effect of PARP inhibitor in breast cancer cells by modulating DNA damage response
KRAS	RAS mutant cell lines are resistant to PARPi.	Gain	Rational combination therapy with PARP and MEK inhibitors capitalizes on therapeutic liabilities in RAS mutant cancers
HRAS	RAS mutant cell lines are resistant to PARPi.	Gain	Rational combination therapy with PARP and MEK inhibitors capitalizes on therapeutic liabilities in RAS mutant cancers
NRAS	RAS mutant cell lines are resistant to PARPi.	Gain	Rational combination therapy with PARP and MEK inhibitors capitalizes on therapeutic liabilities in RAS mutant cancers
BRAF	RAS mutant cell lines are resistant to PARPi.	Gain	Rational combination therapy with PARP and MEK inhibitors capitalizes on therapeutic liabilities in RAS mutant cancers
ARID1A	Unknown	Unclear	ARID1A Deficiency Impairs the DNA Damage Checkpoint and Sensitizes Cells to PARP Inhibitors, A quantitative chemotherapy genetic interaction map reveals factors associated with PARP inhibitor resistance
GPBP1	Unknown	Loss	A quantitative chemotherapy genetic interaction map reveals factors associated with PARP inhibitor resistance
TDG	Unknown	Loss	A quantitative chemotherapy genetic interaction map reveals factors associated with PARP inhibitor resistance
RNF168	Multiple	Unclear	

¹Phenotype appears to be BRCA-dependent.

Genes that are known or suspected to be linked to resistance to PARPi are shown. Resistance indicates whether a Gain or Loss of function has been either shown or inferred to lead to resistance to PARPi.

epithelial cell adhesion that is suppressed in EMT^{67,68}. Surprisingly, we found that resistant tumors had higher, rather than lower, levels of CDH1 expression, although the differences were not statistically significant ($p = 0.16$) (Fig. 4a). Next, we examined tumors for expression of both CDH1 (an epithelial state marker) and vimentin (VIM, a mesenchymal state marker) but did not see a clear distinction between canonical epithelial cells (high CDH1 and low VIM) and mesenchymal cells (low CDH1 and high VIM)⁶⁹ (Fig. 4b). Therefore, despite the correlation between PARPi resistance and an EMT pathway signature, resistant tumors did not exhibit classic EMT markers.

Because the EMT is associated with a number of phenotypes, we investigated the possibility that activation of a subset of the EMT program may be responsible for the high scores in the gene

expression signature, even if an EMT itself is not achieved. We visualized the expression of the 93 genes on the leading edge of the prGSEA analysis (i.e., the most differentially expressed genes from the EMT gene set) and saw that they are more highly expressed in the resistant tumors (Fig. 4c). To understand their function, we performed a GATHER analysis⁷⁰ with a collection of oncogenic gene sets, which is more comprehensive than Hallmarks, from MSigDB. This revealed 63 pathways that are significantly (false discovery rate (FDR) < 5%) enriched in the subset of the EMT program (Fig. 4d). The most significant one was associated with genes upregulated in embryonic stem cells (ESC_V6.5_UP_EARLY.V1_DN; FDR = 10^{-23}). The fact that the EMT signature was not correlated with EMT markers, but was associated with a stem cell signature, suggests that the Hallmark

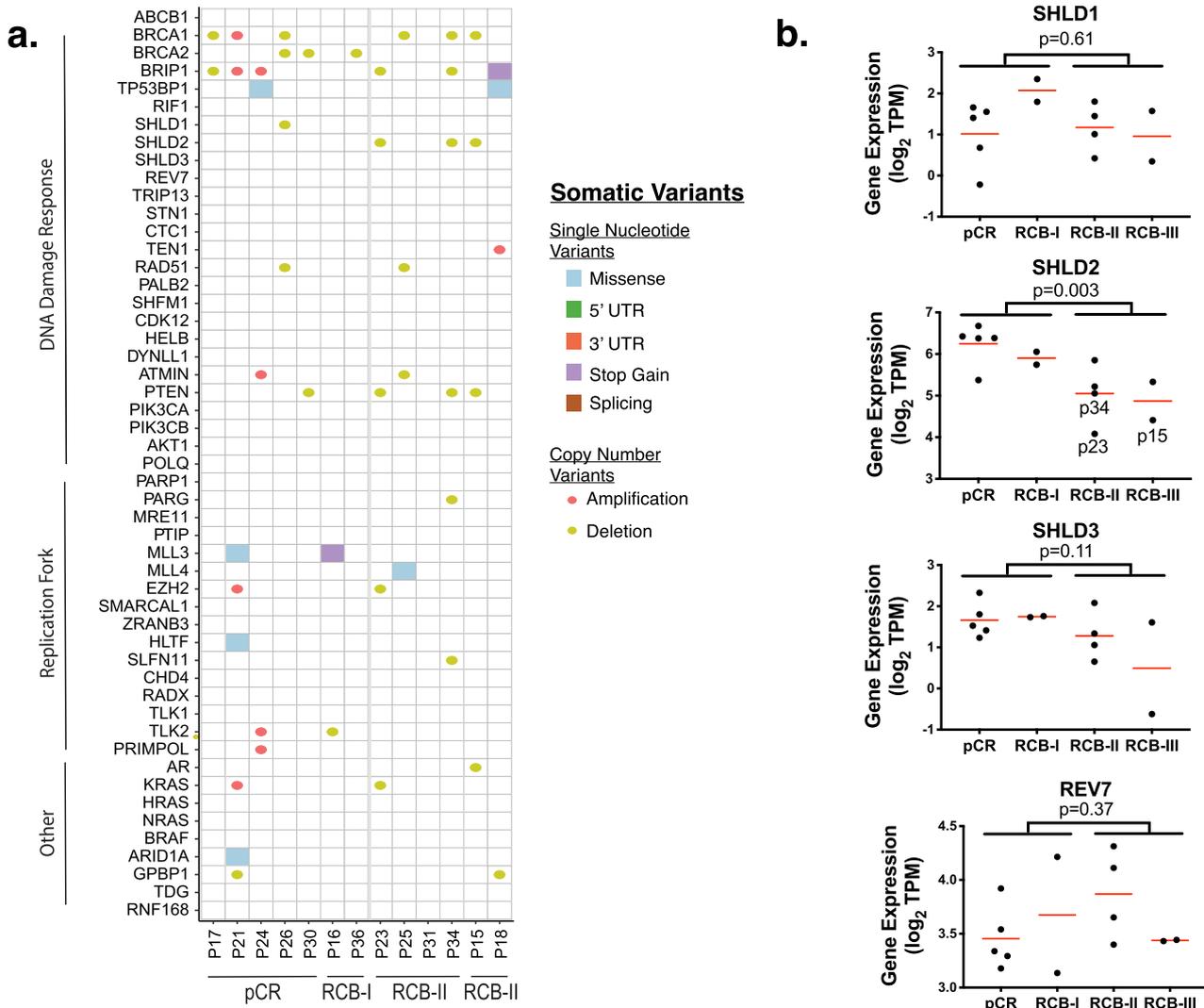


Fig. 2 Known resistance genes associated with resistance. **a** This oncoplot shows the mutations seen across a curated list of genes linked with resistance to PARPi. **b** The gene expression of members of the shieldin complex are shown on the Y-axis as \log_2 TPM. Individual patients (dots) are grouped according to RCB response. Mean expressions are shown by red lines. The p -values are calculated from unpaired two-sided Student's t tests. In the SHLD2 plot, patients predicted to have copy number loss of the gene are labeled.

EMT signature was evincing an EMT-associated stem cell program⁷¹. To determine whether the stem cell program was seen in patient tumors, we scored it on the tumors using prGSEA and found that it was significantly upregulated in resistant tumors ($p = 2 \times 10^{-5}$) (Fig. 4e).

DISCUSSION

We previously demonstrated the efficacy and reduced toxicity of talazoparib as a single-agent neoadjuvant therapy for previously untreated *gBRCA+* cancers¹³. However, not all patients demonstrated benefit as indicated by complete responses to neoadjuvant therapy. Thus, biomarkers of both sensitivity and resistance to PARPi are needed to identify patients most likely to benefit, and resistance remains a critical issue that must be addressed to achieve durable responses in a greater number of patients. Prior studies to determine mechanisms underlying sensitivity and resistance to PARPi have been performed predominantly in cell culture and genetically engineered mouse models, and little is known about the emergence of resistance in the clinical setting of breast cancer, particularly in the up-front neoadjuvant setting.

To identify correlates of resistance, we performed an extensive genomic and transcriptomic analysis of treatment-naïve *gBRCA+* breast cancer tumors treated with neoadjuvant talazoparib monotherapy. This study establishes, for the first time, correlates of sensitivity and resistance to a single agent PARPi in a treatment naïve neoadjuvant breast cancer clinical setting. It reveals potential resistance mechanisms including copy number loss coupled with low expression of the shieldin complex member SHLD2, or a gain of a hypoxia, or EMT/stem cell signature (Fig. 5). The potential role that the shieldin complex plays in promoting resistance to PARPi has been well established in pre-clinical studies. Here, all resistant tumors exhibited either loss of shieldin, gain of hypoxia signature, or gain of an EMT/stem cell signature. Potential biomarkers of response included MYC and E2F1 signatures.

The link between EMT, cancer stem cells, and resistance to chemotherapies is well established⁷² and reported for PARPi as well⁷³. However, the connection between EMT and PARP function remains an intense area of investigation, as PAR appears to be present in epithelial adherens junctions⁷⁴, and PARP1 has both

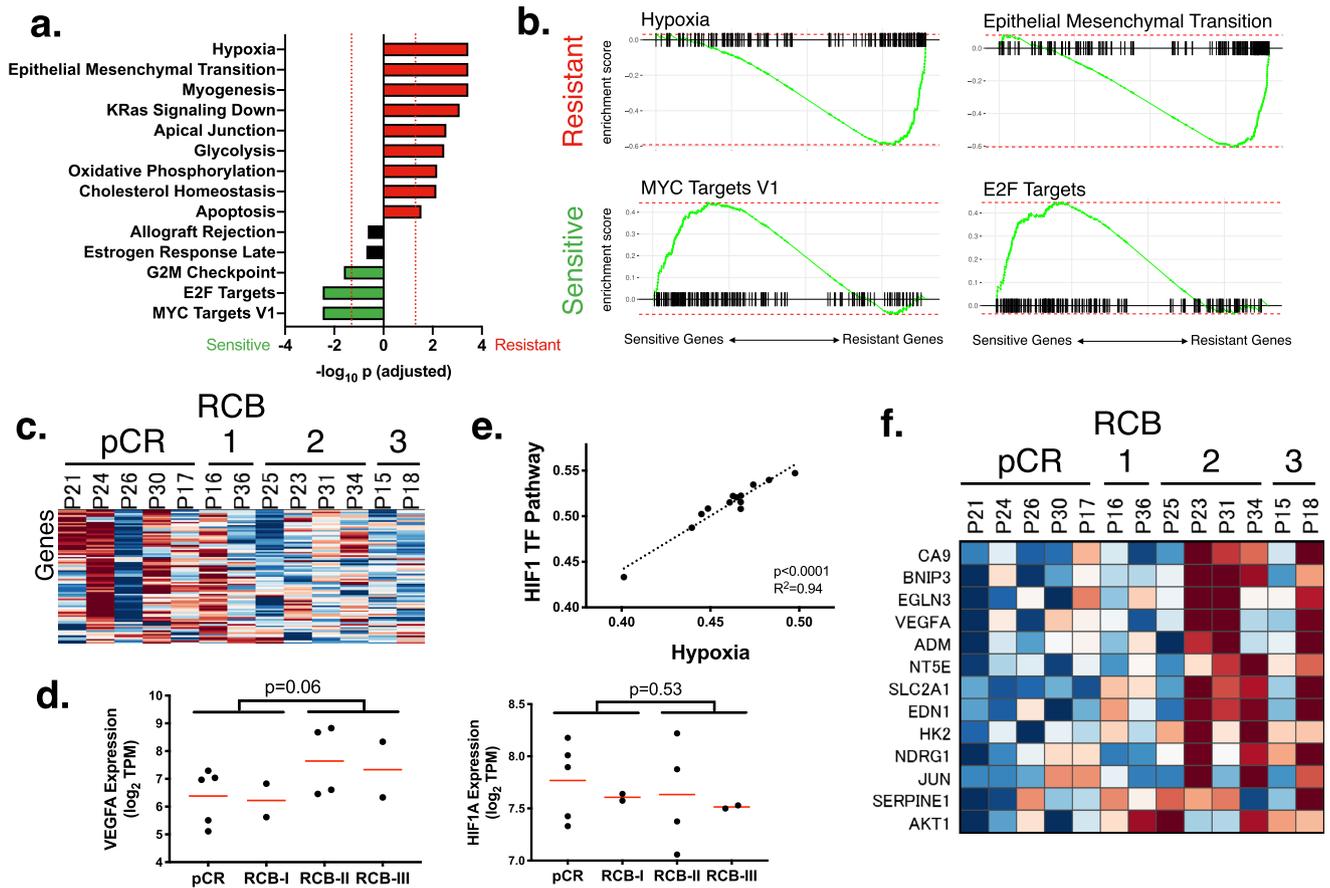


Fig. 3 Pathways correlated with resistance. **a** This shows the association between Hallmarks pathways and response to talazoparib. Each bar stands for a pathway, and the height indicates the statistical significance as the $-\log_{10}$ of the adjusted p -value. The pathways that are associated with sensitivity to treatment are shown in green and those associated with resistance are red. **b** This contains the GSEA enrichment plots for pathways most strongly associated with resistance to talazoparib. Pathways that are higher in resistant tumors are in the top row, and the bottom contains the pathways higher in sensitive tumors. The genes are ranked such that the ones associated with sensitivity are on the left, and those with resistance are on the right. **c** This heatmap shows the expression of the most differentially regulated genes (leading edge) of the MYC Targets V1 pathway (rows) for each of the tumors (columns), organized by response. Warmer colors indicate higher expression, and cooler colors are lower expression. **d** These plots show the association of the expression of VEGFA (left) and HIF1A (right) with response. The mean expression is shown as a red line. p -values (unpaired two-sided Student's t tests) are shown after comparing patients with no progression (pCR and RCB-I) against those that progressed (RCB-II and RCB-III). **e** This plot compares the scores of the hypoxia pathway (x-axis) against that with a HIF1 transcriptional network (y-axis). The p -value is calculated from a two-sided Pearson's correlation coefficient test. **f** This heatmap shows the expression of the genes in the leading edge of the hypoxia pathway. The heatmap is arranged as in panel **c**.

positive^{75,76} and negative^{77,78} associations with TGF- β signaling, a potent inducer of an EMT⁷⁹ that may be context dependent.

In contrast to EMT, the connection between PARPi and hypoxia has been explored more deeply, although with mixed success thus far in the clinic. Under a severe hypoxic condition (0.5% oxygen), cells are more sensitive to PARPi due to the synthetic lethality induced by decreased expression of HR proteins⁸⁰. However, the opposite is seen in moderate hypoxic conditions (2% oxygen) where a hypoxia signature predicted resistance to a PARPi⁸¹ due to decreased ROS-induced DNA damage⁸². In a breast cancer mouse model, hypoxic cells were shown to contribute to PARPi resistance, and targeting tumors with a combination of hypoxia and PARP inhibitors leads to synergistic toxicity. However, the clinical trial results are inconsistent, with phase 2 trials of olaparib and cederinib showing an increased progression free survival^{83,84} whereas the phase 3 trial failed to confirm benefit⁸⁵. As a possible explanation, these studies may be confounded by the pleiotropic effects of cederinib, which can inhibit both VEGFR and decrease expression of HR proteins⁸⁶. Furthermore, the trials were done in an unselected patient population and the number of patients with a pre-existing tumor hypoxia signature may have been sufficiently

low as to confound the outcomes analysis. In a number of other disease contexts, including myocardial ischemia, stroke, and traumatic brain injury, PARPi have been observed to be protective against ischemic cell death (reviewed in^{87–89}). Unfortunately, the mechanism has not yet been established, and it is not known whether a similar process might be activated in hypoxic tumors and render them refractory to PARPi treatment. Nevertheless, the observation that a hypoxia/stem cell signature exists in the majority of PARPi resistant tumors prior to therapy argues for additional studies to understand whether hypoxia can induce a PARPi resistant state, to elucidate underlying mechanisms, and to determine whether targeting hypoxia will increase the benefit of PARPi in patients with a pre-existing hypoxia signature.

Markers of high MYC activity were frequently observed in PARPi sensitive tumors. The relationship between MYC and DNA damage has been well-documented and controversial. Activated MYC can induce chromosomal damage^{90,91} including defects in double-stranded break repair^{92,93}. Another consequence of activated MYC, increased proliferation, may enhance sensitivity to PARPi via PARP trapping, whose toxicity has been shown to be specific to S-phase⁹⁴. In contrast, it has also been reported that MYC can

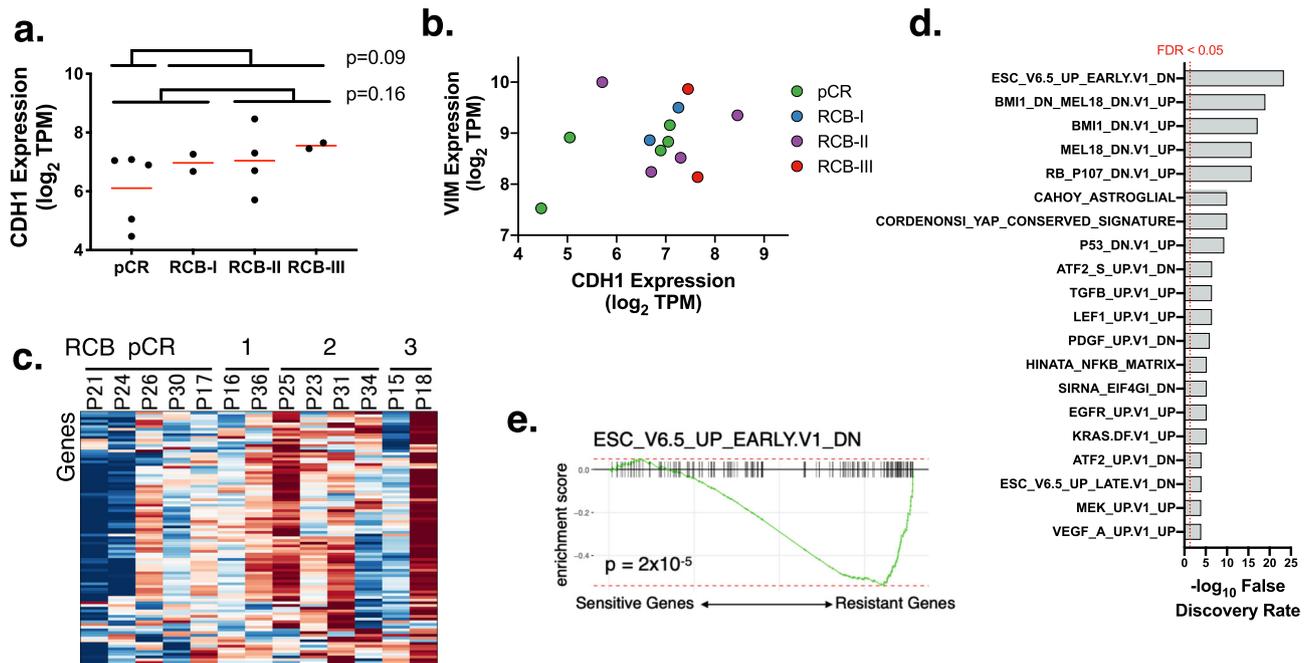


Fig. 4 A stem cell transcriptional program was associated with response. **a** The gene expression (\log_2 TPM) of CDH1 is shown on the y-axis. Each dot is a tumor, grouped according to the response to talazoparib. The red lines indicate mean expression. The p -values (by unpaired two-sided Student's t test) are shown for comparisons between pCR vs other patients, as well as a comparison between pCR and RCB-I, and RCB-I and RCB-II and RCB-III. **b** The gene expression for CDH1 (x-axis; \log_2 TPM) and VIM (y-axis; \log_2 TPM) are shown for tumors colored by response. **c** The rows in the heatmap contain the genes in the leading edge of the EMT gene expression signature, and the columns are samples organized according to response. Warm colors indicate high expression of the genes. **d** The bars in this plot show the significance (as $-\log_{10}$ of the false discovery rate) of the top 20 most significant oncogenic pathway (rows). The dotted red line shows the cutoff for a 5% false discovery rate. **e** The enrichment plots for the oncogenic pathway that is most significantly associated with response.

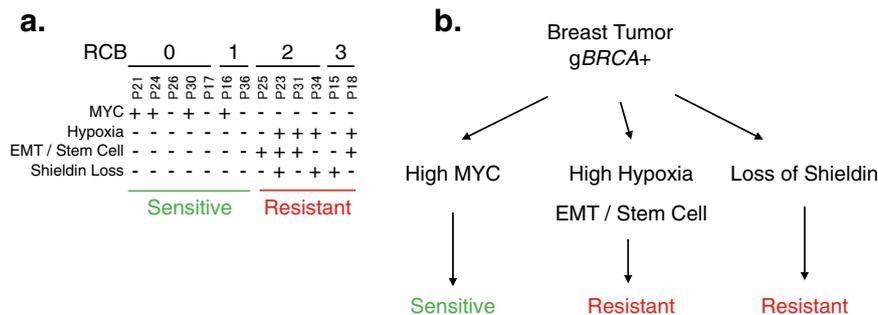


Fig. 5 Model. **a** This table shows each of the tumors in the trial, grouped by RCB status. The tumors predicted to have an activated Hypoxia pathway, based on the pathway signature and expression of HIF targets, are indicated in the first row with a "+". The EMT / Stem Cell status is determined from the expression of the leading edge EMT genes. Shieldin Loss reflects the concomitant loss of copy number and expression of SHLD2. **b** In this cohort of $gBRCA$ mutant breast cancer patients, loss of shieldin, high hypoxia signature, or high EMT/Stem Cell signature is predictive of resistance.

facilitate DNA repair by inducing the expression of genes involved in HR, including BRCA1⁹⁵, which may lead to PARPi resistance⁹⁶. The impact of MYC is likely to be context dependent and the relationship between MYC activation and DNA repair in the human tumor environment requires further investigation.

There are several major limitations in this neoadjuvant study of $gBRCA+$ tumors. 1) The sample size was limited. We profiled 18 tumors, and only 13 could be used to interrogate the biology of the cancer cells. A prior study of a larger cohort of ovarian cancer patients found ABCB1 fusions in 8% of patients and BRCA1/2 reversions in another 5%, suggesting that resistance mechanisms are heterogeneous and likely to be found at low frequencies⁹⁷. 2) We profiled pre-treatment samples and therefore can only detect intrinsic mechanisms of resistance and sensitivity. Longitudinal

profiling will be needed to identify resistance mechanisms that are acquired during the course of treatment. The majority of patients had insufficient tumor at surgery following completion of talazoparib therapy therefore limiting detailed longitudinal analysis. 3) Our bulk genomic assays may not have sufficient sensitivity to identify mechanisms that are present in pre-treatment samples at low clonality levels, such as subclones with BRCA reversions^{27,30}. Deep sequencing, single cell sequencing, or ctDNA profiling may facilitate the identification of genetic or transcriptomic events that occur subclonally. 4) Our exome sequencing data may not reveal the vast majority of possible genomic re-arrangements, leaving the possibility that gene fusions contributing to PARPi resistance and sensitivity were missed. 5) We did not interrogate non-coding RNAs, which have

previously been suggested to mediate resistance⁹⁸. 6) We considered only the cancer cells and did not profile the microenvironment, which can impact the efficacy of PARPi. As one example, it was recently shown that PARPi leads to DNA damage that triggers the cGAS/STING pathway, type I interferon signaling, and then cytotoxic T cell function; resulting in tumor clearance^{99–101}. Finally, 7) this is an observational study of a single data set, and follow-up studies will be needed to determine whether our findings causally lead to resistance, and thus, would be viable targets for therapy.

In conclusion, our findings demonstrate that a number of biomarkers for sensitivity to single-agent talazoparib can be identified in treatment-naïve gBRCA+ samples. This supports previous findings showing that intrinsic activities in cancer cells have an impact on the clinical response to PARPi. The markers we identified provides a potential explanation for the high rates of relapse seen in patients treated with PARP inhibitors, several of which exhibited copy number loss of SHLD2. These results predict that defects in the shieldin complex may be a relatively frequent mechanism of resistance to single agent PARPi in patients, and if verified in larger cohorts, may have a role as a biomarker to exclude patients from therapies that will not be effective. Other resistant patients exhibited a hypoxia signature. As this manuscript was under review, a paper was published showing the efficacy of a combinatorial treatment targeting PARP and hypoxia in a breast cancer PDX model⁸², demonstrating the potential for developing therapeutic strategies to overcome this resistance. Further mechanistic studies are needed to understand the signaling networks engaged by these biomarker pathways to select viable therapeutic targets, and translational studies are needed to develop clinical tests and treatments to reverse resistance.

METHODS

Patient cohort

The data was derived from a neoadjuvant trial of single agent talazoparib¹³. The trial was approved by the institutional review board under protocol 2014-0045, and each participant provided written informed consent. Briefly, twenty patients with clinical stage I-III breast cancer and an identified germline *BRCA1* or 2 pathogenic variant were enrolled. Patients received 6 months of single agent talazoparib starting at a dose of 1 mg orally daily. Surgery was scheduled within 5 weeks of the last dose of talazoparib. Final pathology was determined by a dedicated breast pathologist at our institution, and we used residual cancer burden as a measure of response¹⁰².

Genomic DNA, library prep, and capture

Genomic DNA was quantified by Picogreen (Invitrogen) and quality was accessed using Genomic DNA Tape for the 2200 TapeStation (Agilent). DNA from each sample (100–500 ng of genomic DNA) was sheared by sonication with the following conditions: Peak Incident Power 175, Duty Cycle 20%, Intensity 5, Cycles per Burst 200, and 120 s using Covaris E220 instrument (Covaris). To ensure the proper fragment size, samples were checked on TapeStation using the DNA High Sensitivity kit (Agilent). The sheared DNA proceeded to library prep using the KAPA library prep kit (KAPA) following the “with beads” manufacturer protocol. Briefly, this protocol consists of 3 enzymatic reactions for end repair, A-tailing and Adaptor ligation, followed by barcode insertion by PCR using KAPA HiFi polymerase (6 cycles). PCR primers were removed by using a 1.8x volume of Agencourt AMPure PCR Purification kit (Agencourt Bioscience Corporation). At the end of the library prep, samples were analyzed on TapeStation to verify correct fragment size and to ensure the absence of extra bands. Samples were quantified using the KAPA qPCR quantification kit. Equimolar amounts of DNA were pooled for capture (2–6 samples per pool). We used whole-exome biotin labeled probes from Roche Nimblegen (Exome V3) and followed the manufacturer’s protocol for the capture step. Briefly, DNA was pooled (2–6 samples), dried out and after addition of the capture reagents and probes, samples were incubated at 47 °C on a thermocycler with a heated lid (57 °C) for 64–74 h. The targeted regions were recovered using streptavidin beads and the streptavidin-biotin-

probe-target complex was washed and another round of PCR amplification was performed according to the manufacturer’s protocol. The quality of each captured sample was analyzed on TapeStation using the DNA High Sensitivity kit and the enrichment was accessed by qPCR using specific primers designed by Roche Nimblegen. The cutoff for the enrichment was 50-fold minimum.

Mutation data analysis

The captured libraries were sequenced on a HiSeq 2000 (Illumina Inc., San Diego, CA, USA) on a version 3 TruSeq paired end flowcell at a cluster density between 700–1000 K clusters/mm². Sequencing was performed for 2 × 100 paired-end reads with a 7 nt read for indexes using Cycle Sequencing v3 reagents (Illumina). The resulting BCL files containing the sequence data were converted into “.fastq.gz” files and individual libraries within the samples were demultiplexed using CASAVA 1.8.2 with no mismatches.

We called single nucleotide variants (SNVs) using a pipeline¹⁰³ implemented in the Betsy system¹⁰⁴. Starting with the FASTQ files, we trimmed adapter and low-quality sequence using Trimmomatic¹⁰⁵. Next, we aligned the trimmed reads to human reference assembly hg19 using BWA¹⁰⁶ and flagged duplicated reads using Picard¹⁰⁷. We realigned indels and recalibrated base quality scores using GATK¹⁰⁸. Then, we called somatic variants using a consensus of six callers (MuSE¹⁰⁹, MuTect¹¹⁰, MuTect2, SomaticSniper¹¹¹, Strelka2¹¹², and VarScan2¹¹³). For the Strelka analysis, we skipped the depth filter and set minPruning = 3 in MuTect2 to speed up computation¹¹⁴. To confirm the identities of the germline and tumor samples, we used GATK to call variants on the samples and confirmed that the tumor samples shared mutations with its cognate germline. We annotated variants with Annovar¹¹⁵ and SnpEff¹¹⁶, and filtered for nonsynonymous changes supported by at least 2 callers, 20 reads, and 5% variant allele frequency. We also annotated the mutations using Variant Effect Prediction¹¹⁷ and selected the most severe consequence per variant. Synonymous, intergenic, intronic and non-coding transcript mutations were removed.

We estimated copy number profiles using the FACETS algorithm³⁵. We varied the parameters used for FACETS (critical value from 100–500, nbhd from 100–500, and nhet from 15–50). Following a manual examination of the log-ratio and log-odds-ratio plots, we decided to use the models with parameters critical value 300, nbhd 250, and nhet 30 because they had the most plausible models across the samples. We also estimated copy number profiles using Sequenza⁶⁰ with default parameters and manually selected the most plausible of the alternative models. Finally, we estimated tumor purity and ploidy with ABSOLUTE³⁶ with a sigma_p 0.02, max_sigma_h 0.02, max.non.clonal 0.1, min.ploidy 0.95, and max.ploidy 10.

We performed mutational signature analysis using the DeconstructSigs⁴⁹ and SigMA⁴⁸ R packages. We ran DeconstructSigs on the mutations found by at least 2 callers and supported by at least 20 reads and 5% VAF using the *signatures.nature2013* reference with the *exome2genome* mapping. In contrast to the set of mutations above, we did not filter for nonsynonymous changes here. We ran SigMA using the same filtering criteria as DeconstructSigs on the *breast* tumor type with the *seqcap* data option.

We calculated the tumor mutation burden as the number of nonsynonymous mutations per megabase of targeted exome, as determined from the BED files downloaded from the Roche Nimblegen website. We calculated the weighted Genome Instability Index¹¹⁸ as the percent of the genome with either gain or loss of copy number relative to the median copy number of the sample. This was calculated separately for chromosomes 1–22 and then averaged so that the score was not biased by chromosome length.

We calculated the Large-Scale State Transition (LST) and Telomeric Allelic Imbalance (TAI) scores from the FACETS allele-specific copy number predictions described above. We implemented the LST scores⁵² following bio-protocol (<https://doi.org/10.21769/BioProtoc.814>). Briefly, we split the genome into segments with identical copy number profiles (i.e., the same number of major and minor alleles). Then, we counted the number of segments that were at least 10 megabases long that were within 3 megabases of another 10 megabase segment with a different copy number profile. Although the original algorithm smoothed the copy number profiles of short (<3 megabase) segments, we dropped the short segments entirely, as this had only a minor impact on the scores.

We calculated TAI scores⁵⁴. To do this, we looked for regions at the end of the chromosome where the major and minor alleles had different numbers. If this imbalance originated from the telomeres and did not extend to the centromeres (extracted from the UCSC Genome Browser), we increased the TAI score. We made one alteration to the original algorithm and ignored the

p-arms of chromosomes 13, 14, 15, and 22. Because these chromosomes were acrocentric, the p-arms lacked coverage in our whole-exome sequencing panel, and thus, we were not able to obtain estimates of copy number there.

Gene expression data analysis

We used STAR¹¹⁹ to align RNA-sequencing reads and HTSeq-count¹²⁰ to quantify gene-level counts. We performed pathway analysis using a pre-ranked GSEA implemented in fGSEA¹²¹, after ranking the genes by differential expression statistics from DESeq2¹²². We used gene sets found in MSigDB v6.2⁶¹. We annotated gene sets using GATHER⁷⁰.

Reporting summary

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

DATA AVAILABILITY

RNA-seq data described in this study are deposited in the Gene Expression Omnibus database under accession GSE160568. The raw RNA-Seq and whole-exome sequencing data are available under controlled access from the European Genome-phenome Archive under accession EGAD0000100827.

Received: 9 April 2021; Accepted: 12 April 2022;

Published online: 10 May 2022

REFERENCES

- Chen, S. & Parmigiani, G. Meta-analysis of BRCA1 and BRCA2 penetrance. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* **25**, 1329–1333 (2007).
- Sonderstrup, I. M. H. et al. Subtypes in BRCA-mutated breast cancer. *Hum. Pathol.* **84**, 192–201 (2019).
- Antoniou, A. et al. Average risks of breast and ovarian cancer associated with BRCA1 or BRCA2 mutations detected in case Series unselected for family history: a combined analysis of 22 studies. *Am. J. Hum. Genet.* **72**, 1117–1130 (2003).
- Beitsch, P. D. et al. Underdiagnosis of hereditary breast cancer: are genetic testing guidelines a tool or an obstacle? *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* **37**, 453–460 (2019).
- Akashi-Tanaka, S. et al. BRCAness predicts resistance to taxane-containing regimens in triple negative breast cancer during neoadjuvant chemotherapy. *Clin. Breast Cancer* **15**, 80–85 (2015).
- Farmer, H. et al. Targeting the DNA repair defect in BRCA mutant cells as a therapeutic strategy. *Nature* **434**, 917–921 (2005).
- Sachdev, E., Tabatabai, R., Roy, V., Rimel, B. J. & Mita, M. M. PARP inhibition in cancer: an update on clinical development. *Target Oncol.* **14**, 657–679 (2019).
- Robson, M. et al. Olaparib for metastatic breast cancer in patients with a germline BRCA mutation. *N. Engl. J. Med.* **377**, 523–533 (2017).
- Litton, J. K. et al. Talazoparib in patients with advanced breast cancer and a germline BRCA mutation. *N. Engl. J. Med.* **379**, 753–763 (2018).
- Ettl, J. et al. Quality of life with talazoparib versus physician's choice of chemotherapy in patients with advanced breast cancer and germline BRCA1/2 mutation: patient-reported outcomes from the EMBRACA phase III trial. *Ann. Oncol. Off. J. Eur. Soc. Med. Oncol. ESMO* **29**, 1939–1947 (2018).
- Robson, M. E. et al. OlympiAD final overall survival and tolerability results: Olaparib versus chemotherapy treatment of physician's choice in patients with a germline BRCA mutation and HER2-negative metastatic breast cancer. *Ann. Oncol. Off. J. Eur. Soc. Med. Oncol. ESMO* **30**, 558–566 (2019).
- Litton, J. K. et al. Talazoparib versus chemotherapy in patients with germline BRCA1/2-mutated HER2-negative advanced breast cancer: final overall survival results from the EMBRACA trial. *Ann. Oncol. Off. J. Eur. Soc. Med. Oncol. ESMO*, <https://doi.org/10.1016/j.annonc.2020.08.2098> (2020).
- Litton, J. K. et al. Neoadjuvant talazoparib for patients with operable breast cancer with a germline BRCA pathogenic variant. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* **38**, 388–394 (2020).
- Vaidyanathan, A. et al. ABCB1 (MDR1) induction defines a common resistance mechanism in paclitaxel- and olaparib-resistant ovarian cancer cells. *Br. J. Cancer* **115**, 431–441 (2016).
- Elmeligy, M. et al. Evaluation of the effect of P-glycoprotein inhibition and induction on talazoparib disposition in patients with advanced solid tumours. *Br. J. Clin. Pharm.* **86**, 771–778 (2020).
- Shen, D. W. et al. Human multidrug-resistant cell lines: increased mdr1 expression can precede gene amplification. *Science* **232**, 643–645 (1986).
- Swisher, E. M. et al. Secondary BRCA1 mutations in BRCA1-mutated ovarian carcinomas with platinum resistance. *Cancer Res.* **68**, 2581–2586 (2008).
- Sakai, W. et al. Secondary mutations as a mechanism of cisplatin resistance in BRCA2-mutated cancers. *Nature* **451**, 1116–1120 (2008).
- Johnson, N. et al. Stabilization of mutant BRCA1 protein confers PARP inhibitor and platinum resistance. *Proc. Natl Acad. Sci. USA* **110**, 17041–17046 (2013).
- Gupta, R. et al. DNA repair network analysis reveals shieldin as a key regulator of NHEJ and PARP inhibitor sensitivity. *Cell* **173**, 972–988 e923 (2018).
- Noordermeer, S. M. et al. The shieldin complex mediates 53BP1-dependent DNA repair. *Nature* **560**, 117–121 (2018).
- Ray Chaudhuri, A. et al. Replication fork stability confers chemoresistance in BRCA-deficient cells. *Nature* **535**, 382–387 (2016).
- Quinet, A. et al. PRIMPOL-mediated adaptive response suppresses replication fork reversal in BRCA-deficient cells. *Mol. Cell* **77**, 461–474 e469 (2020).
- Barber, L. J. et al. Secondary mutations in BRCA2 associated with clinical resistance to a PARP inhibitor. *J. Pathol.* **229**, 422–429 (2013).
- Brady, S. W. et al. Combating subclonal evolution of resistant cancer phenotypes. *Nat. Commun.* **8**, 1231 (2017).
- Waks, A. G. et al. Reversion and non-reversion mechanisms of resistance to PARP inhibitor or platinum chemotherapy in BRCA1/2-mutant metastatic breast cancer. *Ann. Oncol. Off. J. Eur. Soc. Med. Oncol. ESMO* **31**, 590–598 (2020).
- Quigley, D. et al. Analysis of circulating cell-free DNA identifies multiclonal heterogeneity of BRCA2 reversion mutations associated with resistance to PARP inhibitors. *Cancer Discov.* **7**, 999–1005 (2017).
- Christie, E. L. et al. Reversion of BRCA1/2 germline mutations detected in circulating tumor DNA from patients with high-grade serous ovarian cancer. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* **35**, 1274–1280 (2017).
- Mayor, P., Gay, L. M., Lele, S. & Elvin, J. A. BRCA1 reversion mutation acquired after treatment identified by liquid biopsy. *Gynecol. Oncol. Rep.* **21**, 57–60 (2017).
- Weigelt, B. et al. Diverse BRCA1 and BRCA2 reversion mutations in circulating cell-free DNA of therapy-resistant breast or ovarian cancer. *Clin. Cancer Res.* **23**, 6708–6720 (2017).
- Lin, K. K. et al. BRCA reversion mutations in circulating tumor DNA predict primary and acquired resistance to the PARP inhibitor rucaparib in high-grade ovarian carcinoma. *Cancer Discov.* **9**, 210–219 (2019).
- Kondrashova, O. et al. Secondary somatic mutations restoring RAD51C and RAD51D associated with acquired resistance to the PARP inhibitor rucaparib in high-grade ovarian carcinoma. *Cancer Discov.* **7**, 984–998 (2017).
- Pettitt, S. J. et al. Genome-wide and high-density CRISPR-Cas9 screens identify point mutations in PARP1 causing PARP inhibitor resistance. *Nat. Commun.* **9**, 1849 (2018).
- Abrial, C. et al. Measurement of residual disease after neoadjuvant chemotherapy. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* **26**, 3094 (2008). author reply 3095.
- Shen, R. & Seshan, V. E. FACETS: allele-specific copy number and clonal heterogeneity analysis tool for high-throughput DNA sequencing. *Nucleic Acids Res.* **44**, e131 (2016).
- Carter, S. L. et al. Absolute quantification of somatic DNA alterations in human cancer. *Nat. Biotechnol.* **30**, 413–421 (2012).
- Greenblatt, M. S., Chappuis, P. O., Bond, J. P., Hamel, N. & Foulkes, W. D. TP53 mutations in breast cancer associated with BRCA1 or BRCA2 germ-line mutations: distinctive spectrum and structural distribution. *Cancer Res.* **61**, 4092–4097 (2001).
- Johannsson, O. T. et al. Tumour biological features of BRCA1-induced breast and ovarian cancer. *Eur. J. Cancer* **33**, 362–371 (1997).
- Holstege, H. et al. High incidence of protein-truncating TP53 mutations in BRCA1-related breast cancer. *Cancer Res.* **69**, 3625–3633 (2009).
- Cho, Y., Gorina, S., Jeffrey, P. D. & Pavletich, N. P. Crystal structure of a p53 tumor suppressor-DNA complex: understanding tumorigenic mutations. *Science*. **265**, 346–355 (1994).
- Hainaut, P. & Hollstein, M. p53 and human cancer: the first ten thousand mutations. *Adv. Cancer Res.* **77**, 81–137 (2000).
- Tate, J. G. et al. COSMIC: the catalogue of somatic mutations in cancer. *Nucleic Acids Res.* **47**, D941–D947 (2019).
- Adzhubei, I., Jordan, D. M. & Sunyaev, S. R. Predicting functional effect of human missense mutations using PolyPhen-2. *Curr. Protoc. Hum. Genet.* **Chapter 7**, Unit7 20 (2013).
- Ng, P. C. & Henikoff, S. SIFT: Predicting amino acid changes that affect protein function. *Nucleic Acids Res.* **31**, 3812–3814 (2003).
- Alexandrov, L. B. et al. Signatures of mutational processes in human cancer. *Nature* **500**, 415–421 (2013).
- Polak, P. et al. A mutational signature reveals alterations underlying deficient homologous recombination repair in breast cancer. *Nat. Genet.* **49**, 1476–1486 (2017).
- Poti, A. et al. Correlation of homologous recombination deficiency induced mutational signatures with sensitivity to PARP inhibitors and cytotoxic agents. *Genome Biol.* **20**, 240 (2019).

48. Gulhan, D. C., Lee, J. J., Melloni, G. E. M., Cortes-Ciriano, I. & Park, P. J. Detecting the mutational signature of homologous recombination deficiency in clinical samples. *Nat. Genet.* **51**, 912–919 (2019).
49. Rosenthal, R., McGranahan, N., Herrero, J., Taylor, B. S. & Swanton, C. DeconstructSigs: delineating mutational processes in single tumors distinguishes DNA repair deficiencies and patterns of carcinoma evolution. *Genome Biol.* **17**, 31 (2016).
50. Nolan, E. et al. Combined immune checkpoint blockade as a therapeutic strategy for BRCA1-mutated breast cancer. *Sci. Transl. Med.* **9**, <https://doi.org/10.1126/scitranslmed.aal4922> (2017).
51. Endesfelder, D. et al. Chromosomal instability selects gene copy-number variants encoding core regulators of proliferation in ER+ breast cancer. *Cancer Res.* **74**, 4853–4863 (2014).
52. Manie, E. et al. Genomic hallmarks of homologous recombination deficiency in invasive breast carcinomas. *Int. J. Cancer* **138**, 891–900 (2016).
53. Popova, T. et al. Ploidy and large-scale genomic instability consistently identify basal-like breast carcinomas with BRCA1/2 inactivation. *Cancer Res.* **72**, 5454–5462 (2012).
54. Birkbak, N. J. et al. Telomeric allelic imbalance indicates defective DNA repair and sensitivity to DNA-damaging agents. *Cancer Disco.* **2**, 366–375 (2012).
55. Chen, X. et al. Manta: rapid detection of structural variants and indels for germline and cancer sequencing applications. *Bioinformatics* **32**, 1220–1222 (2016).
56. Wang, Y. et al. The BRCA1-Delta11q alternative splice isoform bypasses germline mutations and promotes therapeutic resistance to PARP inhibition and cisplatin. *Cancer Res.* **76**, 2778–2790 (2016).
57. Maxwell, K. N. et al. BRCA locus-specific loss of heterozygosity in germline BRCA1 and BRCA2 carriers. *Nat. Commun.* **8**, 319 (2017).
58. Mirman, Z. et al. 53BP1-RIF1-shieldin counteracts DSB resection through CST- and Polalpha-dependent fill-in. *Nature* **560**, 112–116 (2018).
59. Dev, H. et al. Shieldin complex promotes DNA end-joining and counters homologous recombination in BRCA1-null cells. *Nat. Cell Biol.* **20**, 954–965 (2018).
60. Favero, F. et al. Sequenza: allele-specific copy number and mutation profiles from tumor sequencing data. *Ann. Oncol. Off. J. Eur. Soc. Med. Oncol. ESMO* **26**, 64–70 (2015).
61. Liberzon, A. et al. Molecular signatures database (MSigDB) 3.0. *Bioinformatics* **27**, 1739–1740 (2011).
62. Fang, Y. et al. Sequential therapy with PARP and WEE1 inhibitors minimizes toxicity while maintaining efficacy. *Cancer Cell* **35**, 851–867 e857 (2019).
63. Buffa, F. M., Harris, A. L., West, C. M. & Miller, C. J. Large meta-analysis of multiple cancers reveals a common, compact and highly prognostic hypoxia metagene. *Br. J. Cancer* **102**, 428–435 (2010).
64. LaGory, E. L. & Giaccia, A. J. The ever-expanding role of HIF in tumour and stromal biology. *Nat. Cell Biol.* **18**, 356–365 (2016).
65. Schaefer, C. F. et al. PID: the pathway interaction database. *Nucleic Acids Res.* **37**, D674–D679 (2009).
66. Allison Stewart, C. et al. Dynamic variations in epithelial-to-mesenchymal transition (EMT), ATM, and SLFN11 govern response to PARP inhibitors and cisplatin in small cell lung cancer. *Oncotarget* **8**, 28575–28587 (2017).
67. Larue, L., Ohsugi, M., Hirchenhain, J. & Kemler, R. E-cadherin null mutant embryos fail to form a trophoblast epithelium. *Proc. Natl Acad. Sci. USA* **91**, 8263–8267 (1994).
68. Cano, A. et al. The transcription factor snail controls epithelial-mesenchymal transitions by repressing E-cadherin expression. *Nat. Cell Biol.* **2**, 76–83 (2000).
69. Zeisberg, M. & Neilson, E. G. Biomarkers for epithelial-mesenchymal transitions. *J. Clin. Invest.* **119**, 1429–1437 (2009).
70. Chang, J. T. & Nevins, J. R. GATHER: a systems approach to interpreting genomic signatures. *Bioinformatics* **22**, 2926–2933 (2006).
71. Ye, X. et al. Distinct EMT programs control normal mammary stem cells and tumour-initiating cells. *Nature* **525**, 256–260 (2015).
72. Shibue, T. & Weinberg, R. A. EMT, CSCs, and drug resistance: the mechanistic link and clinical implications. *Nat. Rev. Clin. Oncol.* **14**, 611–629 (2017).
73. Ordonez, L. D. et al. Rapid activation of epithelial-mesenchymal transition drives PARP inhibitor resistance in Brca2-mutant mammary tumours. *Oncotarget* **10**, 2586–2606 (2019).
74. Lafon-Hughes, L., Vilchez Larrea, S. C., Kun, A. & Fernandez Villamil, S. H. VERO cells harbor a poly-ADP-ribose belt partnering their epithelial adhesion belt. *PeerJ* **2**, e617 (2014).
75. Huang, D. et al. Poly(ADP-ribose) polymerase 1 is indispensable for transforming growth factor-beta Induced Smad3 activation in vascular smooth muscle cell. *PLoS One* **6**, e27123 (2011).
76. Schacke, M. et al. PARP-1/2 inhibitor olaparib prevents or partially reverts EMT induced by TGF-beta in NMuMG Cells. *Int. J. Mol. Sci.* **20**, <https://doi.org/10.3390/ijms20030518> (2019).
77. Lonn, P. et al. PARP-1 attenuates Smad-mediated transcription. *Mol. Cell* **40**, 521–532 (2010).
78. Dahl, M. et al. Fine-tuning of Smad protein function by poly(ADP-ribose) polymerases and poly(ADP-ribose) glycohydrolase during transforming growth factor beta signaling. *PLoS One* **9**, e103651 (2014).
79. Zavadil, J. & Bottinger, E. P. TGF-beta and epithelial-to-mesenchymal transitions. *Oncogene* **24**, 5764–5774 (2005).
80. Chan, N. et al. Contextual synthetic lethality of cancer cell kill based on the tumor microenvironment. *Cancer Res.* **70**, 8045–8054 (2010).
81. Mehibel, M. et al. Modulating the tumor microenvironment to enhance efficacy of PARP inhibitors. *J. Clin. Oncol.* **37**, e14715–e14715 (2019).
82. Mehibel, M. et al. Eliminating hypoxic tumor cells improves response to PARP inhibitors in homologous recombination-deficient cancer models. *J. Clin. Invest.* **131**, <https://doi.org/10.1172/JCI146256> (2021).
83. Liu, J. F. et al. Combination cediranib and olaparib versus olaparib alone for women with recurrent platinum-sensitive ovarian cancer: a randomised phase 2 study. *Lancet Oncol.* **15**, 1207–1214 (2014).
84. Liu, J. F. et al. Overall survival and updated progression-free survival outcomes in a randomized phase II study of combination cediranib and olaparib versus olaparib in relapsed platinum-sensitive ovarian cancer. *Ann. Oncol. Off. J. Eur. Soc. Med. Oncol. ESMO* **30**, 551–557 (2019).
85. Kemp, A. Update on Phase III GY004 trial for cediranib added to Lynparza in platinum-sensitive relapsed ovarian cancer. (3/12/2020). [Press release]. <https://www.astrazeneca.com/media-centre/press-releases/2020/update-on-phase-iii-gy004-trial-for-cediranib-added-to-lynparza-in-platinum-sensitive-relapsed-ovarian-cancer.html>.
86. Kaplan, A. R. et al. Cediranib suppresses homology-directed DNA repair through down-regulation of BRCA1/2 and RAD51. *Sci. Transl. Med.* **11**, <https://doi.org/10.1126/scitranslmed.aav4508> (2019).
87. Koh, S. H. et al. The effect of PARP inhibitor on ischaemic cell death, its related inflammation and survival signals. *Eur. J. Neurosci.* **20**, 1461–1472 (2004).
88. Curtin, N. J. & Szabo, C. Therapeutic applications of PARP inhibitors: anticancer therapy and beyond. *Mol. Asp. Med* **34**, 1217–1256 (2013).
89. Szabo, C. & Dawson, V. L. Role of poly(ADP-ribose) synthetase in inflammation and ischaemia-reperfusion. *Trends Pharm. Sci.* **19**, 287–298 (1998).
90. Felsher, D. W. & Bishop, J. M. Transient excess of MYC activity can elicit genomic instability and tumorigenesis. *Proc. Natl Acad. Sci. USA* **96**, 3940–3944 (1999).
91. Vafa, O. et al. c-Myc can induce DNA damage, increase reactive oxygen species, and mitigate p53 function: a mechanism for oncogene-induced genetic instability. *Mol. Cell* **9**, 1031–1044 (2002).
92. Karlsson, A. et al. Defective double-strand DNA break repair and chromosomal translocations by MYC overexpression. *Proc. Natl Acad. Sci. USA* **100**, 9974–9979 (2003).
93. Li, Z. et al. c-Myc suppression of DNA double-strand break repair. *Neoplasia* **14**, 1190–1202 (2012).
94. Michelena, J. et al. Analysis of PARP inhibitor toxicity by multidimensional fluorescence microscopy reveals mechanisms of sensitivity and resistance. *Nat. Commun.* **9**, 2678 (2018).
95. Chen, Y. et al. c-Myc activates BRCA1 gene expression through distal promoter elements in breast cancer cells. *BMC Cancer* **11**, 246 (2011).
96. Carey, J. P. W. et al. Synthetic lethality of PARP inhibitors in combination with MYC blockade is independent of BRCA status in triple-negative breast cancer. *Cancer Res.* **78**, 742–757 (2018).
97. Patch, A. M. et al. Whole-genome characterization of chemo-resistant ovarian cancer. *Nature* **521**, 489–494 (2015).
98. Meghani, K. et al. Multifaceted impact of microRNA 493-5p on genome-stabilizing pathways induces platinum and PARP inhibitor resistance in BRCA2-mutated carcinomas. *Cell Rep.* **23**, 100–111 (2018).
99. Shen, J. et al. PARPi triggers the STING-dependent immune response and enhances the therapeutic efficacy of immune checkpoint blockade independent of BRCAness. *Cancer Res.* **79**, 311–319 (2019).
100. Pantelidou, C. et al. PARP inhibitor efficacy depends on CD8(+) T-cell recruitment via intratumoral STING pathway activation in BRCA-deficient models of triple-negative breast cancer. *Cancer Discov.* **9**, 722–737 (2019).
101. Ding, L. et al. PARP inhibition elicits STING-dependent antitumor immunity in BRCA1-deficient ovarian cancer. *Cell Rep.* **25**, 2972–2980 e2975 (2018).
102. Symmans, W. F. et al. Measurement of residual breast cancer burden to predict survival after neoadjuvant chemotherapy. *J. Clin. Oncol.: Off. J. Am. Soc. Clin. Oncol.* **25**, 4414–4422 (2007).
103. Echeverria, G. V. et al. High-resolution clonal mapping of multi-organ metastasis in triple negative breast cancer. *Nat. Commun.* **9**, 5079 (2018).
104. Chen, X. & Chang, J. T. Planning bioinformatics workflows using an expert system. *Bioinformatics* **33**, 1210–1215 (2017).
105. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).

106. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
107. DePristo, M. A. et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat. Genet* **43**, 491–498 (2011).
108. McKenna, A. et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–1303 (2010).
109. Fan, Y. et al. MuSE: accounting for tumor heterogeneity using a sample-specific error model improves sensitivity and specificity in mutation calling from sequencing data. *Genome Biol.* **17**, 178 (2016).
110. Cibulskis, K. et al. Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nat. Biotechnol.* **31**, 213–219 (2013).
111. Larson, D. E. et al. SomaticSniper: identification of somatic point mutations in whole genome sequencing data. *Bioinformatics* **28**, 311–317 (2012).
112. Kim, S. et al. Strelka2: fast and accurate calling of germline and somatic variants. *Nat. Methods* **15**, 591–594 (2018).
113. Amarasinghe, K. C. et al. Inferring copy number and genotype in tumour exome data. *BMC Genomics* **15**, 732 (2014).
114. Callari, M. et al. Intersect-then-combine approach: improving the performance of somatic variant calling in whole exome sequencing data using multiple aligners and callers. *Genome Med.* **9**, 35 (2017).
115. Wang, K., Li, M. & Hakonarson, H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* **38**, e164 (2010).
116. Cingolani, P. et al. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly. (Austin)* **6**, 80–92 (2012).
117. McLaren, W. et al. The ensembl variant effect predictor. *Genome Biol.* **17**, 122 (2016).
118. Burrell, R. A. et al. Replication stress links structural and numerical cancer chromosomal instability. *Nature* **494**, 492–496 (2013).
119. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
120. Anders, S., Pyl, P. T. & Huber, W. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166–169 (2015).
121. Korotkevich, G., Sukhov, V., Budin, N. & Sergushichev, A. Fast gene set enrichment analysis. *bioRxiv*, <https://doi.org/10.1101/060012> (2019).
122. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).

ACKNOWLEDGEMENTS

We are grateful to the patients who provided tumor biopsies for these studies. The research reported in this publication was supported by the generous philanthropic contributions to The University of Texas MD Anderson Cancer Center Moon Shots Program™, the Toomim Family Fund, and the Miriam and Sheldon Adelson Medical Research Foundation (G.B.M.). It was also supported by the National Cancer Institute of the National Institutes of Health under award numbers P50CA217685 (G.B.M.), U01CA217842 (G.B.M.), R50CA221675 (Y.L.), and U54CA209978 (J.T.C.); the Cancer Prevention and Research Institute of Texas grants RP160710 (H.P.W., S.L.M., J.T.C.) and RP170668 (J.T.C.); the Susan G. Komen Breast Cancer Foundation SAC110052 (G.B.M.); the Breast Cancer Research Foundation BCRF-19-110 (G.B.M.). H.P.W. is an ACS Research Professor.

AUTHOR CONTRIBUTIONS

B.K.A., S.L.M., G.B.M., H.P.W., J.K.L., J.T.C.; Investigation, X.L., Z.G., F.Y., A.C., S.L., Y.L., M.L., S.L.M., G.B.M., H.P.W., J.K.L., J.T.C.; Project Administration, J.B.W.; Writing – original draft, X.G., J.T.C.; Writing – reviewing and editing, X.L., X.G., S.L.M., G.B.M., H.P.W., J.K.L., J.T.C.; Supervision, H.P.W., J.K.L., J.T.C.; Funding Acquisition, B.K.A., S.L.M., G.B.M., H.P.W., J.K.L., J.T.C. Ordering of the co-first authors (X.L. and X.G.) was determined based on the contributions to the final manuscript.

COMPETING INTERESTS

G.B.M. is a board member or consultant for AstraZeneca, Chrysalis Biotechnology, GlaxoSmithKline, ImmunoMET, Ionis, Eli Lilly, PDX Pharmaceuticals, Signalchem Life Science, Symphogen, Tarveda, Turbine, Zentalis Pharmaceuticals; a shareholder of Catena Pharmaceuticals, ImmunoMet, SignalChem, and Tarveda; and has licensed technology to Myriad Genetics and Nanostring. J.K.L. has research support from Novartis, Medivation/Pfizer, Genentech, GlaxoSmithKline, EMD-Serono, Astra-Zeneca, Medimmune, and Zenith; is a board member or consultant for Astra-Zeneca, Ayala, and Pfizer (all uncompensated); and is on the speaker's bureau for MedLearning, Physician's Education Resource, Prime Oncology, Medscape, Clinical Care Options, and Medpage. Other authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41523-022-00427-9>.

Correspondence and requests for materials should be addressed to Jeffrey T. Chang.

Reprints and permission information is available at <http://www.nature.com/reprints>

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



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022