Article

Homologous recombination deficiency and hemizygosity drive resistance in breast cancer

https://doi.org/10.1038/s41586-026-10197-0

Received: 3 April 2024

Accepted: 23 January 2026

Published online: 4 March 2026

Open access

Check for updates

Anton Safonov[1,2,24], Minna Lee[3,4,24], David N. Brown[5], Luca Boscolo Bielo[1,6], Miika Mehine[5,7], Chaitanya Bandlamudi[7], Ben O'Leary[8], Hong Shao[4], Laia Vicente[9], Daniel Muldoon[5,7], Allen Zhu[4], Susana Ros[10], Antonio Marra[11], Pier Selenica[5], Ivan Bieche[12], Bradley Wubbenhorst[13], Emanuela Ferraro[1,11], Laura Courtois[12], Rania El Botty[14], Mehnaj Ahmed[1], Enrico Moiso[6], Julia Ah-Reum An[6], Mark T. A. Donoghue[7], Marie Will[1,4], Fresia Pareja[5], Emily Nizialek[10], Natalia Lukashchuk[10], Eleni Sofianopoulou[10], Yuan Liu[15], Xin Huang[15], Colombe Chappey[15], Anna D. Staniszewska[10], Dara Ross[5], Diana Mandelker[5], Marc Ladanyi[5,7], Nikolaus Schultz[4,7,16], Michael F. Berger[5,7], Maurizio Scaltriti[17], Jorge S. Reis-Filho[17], Bob T. Li[1,17], Kenneth Offit[18], Larry Norton[1], Ronglai Shen[15], Kara N. Maxwell[19,20,21], Fergus Couch[22], Susan M. Domchek[19,20,21], Elisabetta Marangoni[12], Sohrab Shah[16,23], Mark R. Albertella[10], Violeta Serra[9,10], Britta Weigelt[5], David B. Solit[1,4,7], Katherine L. Nathanson[13,19,21], Mark E. Robson[1,2,18], Nicholas C. Turner[8], Sarat Chandarlapaty[1,2,4] ✉ & Pedram Razavi[1,2,6] ✉

The co-occurrence of germline and somatic oncogenic alterations is frequently observed in breast cancer, yet their combined influence on tumour evolution and therapy resistance remains poorly defined. Through an integrated clinicogenomic analysis of more than 5,800 patients, we show that germline (g) pathogenic variants dictate the evolutionary trajectory of acquired resistance. We specifically find that g*BRCA2*-associated tumours are uniquely predisposed to develop acquired *RB1* loss-of-function alterations, resulting in poor outcomes on standard-of-care frontline CDK4/6 inhibitor (CDK4/6i) combinations. This vulnerability is driven by a dual mechanism: baseline *RB1* hemizygosity (heterozygous loss resulting in a single functional *RB1* allele), which lowers the evolutionary barrier to biallelic inactivation, and ongoing homologous recombination deficiency, which promotes acquisition of *RB1* loss-of-function alterations under the selective pressure of CDK4/6i. Preclinical models from g*BRCA2* carriers showed near-uniform resistance to CDK4/6i, with consistent post-treatment Rb loss. Across multiple independent models and in our clinical data, PARP inhibition consistently outperformed CDK4/6i. Our findings suggest that prioritizing PARP inhibition in g*BRCA2* carriers may intercept *RB1*-loss trajectories and delay resistance. More broadly, we establish a predictive framework for forecasting drug-resistant trajectories based on pre-treatment allelic configuration and mutational signatures.

Resistance to anticancer therapy often occurs through a diverse array of genomic mechanisms[1–5]. The inherent unpredictability of these events challenges the ability to develop proactive strategies to pre-empt the emergence of resistance and improve patient outcomes. The role of germline pathogenic variants in predisposition to malignancies and shaping the initial somatic landscape of tumours is well established[6–9]. This interplay is well exemplified in germline pathogenic variants in certain DNA damage repair genes[10–14] that give rise to characteristic somatic allelic configurations and genomic instability in the form of homologous recombination deficiency (HRD)[15–17], a biology that has been successfully exploited with PARP inhibitors (PARPi)[18–20]. However, despite our understanding of these initial events, the influence of germline pathogenic variants on the subsequent evolutionary life of a tumour remains poorly defined.

Clinically, the expanding landscape of approved targeted and lineage-directed therapies has introduced considerable uncertainty regarding the optimal treatment paradigm for patients with breast cancer with certain germline and somatic backgrounds. Given its dominant effect on initial cancer evolution, we posit that the germline background may equally influence tumour behaviour under therapeutic pressure, effectively directing the evolutionary trajectory of disease progression and resistance.

More broadly, defining how pre-treatment germline and somatic genomic context shapes the path of acquired therapy resistance could enable early interception strategies to prevent or delay the emergence of drug-resistant tumour clones. To elucidate the therapeutic relevance of germline–somatic interactions in a clinically meaningful setting, we performed an integrated analysis of germline and somatic genomic data

A list of affiliations appears at the end of the paper.

paired with detailed clinical annotation, including treatment response, in a large cohort of patients with breast cancer.

Clinical and genomic features of the cohort

To identify the interactions between germline pathogenic variants and somatic oncogenic alterations in breast cancer, we integrated detailed clinical annotation with prospectively collected sequencing data[21–23] based on 6,927 tumours from 5,881 patients with breast cancer (Memorial Sloan Kettering Cancer Center (MSK) cohort; Extended Data Fig. 1a and Supplementary Table 1). DNA derived from tumour tissue and blood as a source of germline DNA were each sequenced using an FDA-authorized clinical sequencing assay encompassing up to 506 cancer-associated genes, including germline analysis of 84 cancer predisposition genes[21–23]. Genes of interest for the germline analysis included canonical members of the homologous recombination pathway[10–14]: *BRCA2* (2.9%, $n = 161$ patients), *BRCA1* (2.6%, $n = 142$ patients), *CHEK2* (1.6%, $n = 87$ patients), *ATM* (1.1%, $n = 60$ patients) and *PALB2* (0.6%, $n = 35$ patients).

The clinicopathological characteristics of the germline-altered cancers and germline wild-type (gWT) cancers strongly reflected previously established patterns, suggesting that our cohort was representative of the broader population of patients with breast cancer (Table 1). Specifically, we observed a younger age of diagnosis in g*BRCA1/2* carriers than in gWT. g*BRCA1*-associated breast cancers tended to be triple-negative[24] and high-grade invasive ductal carcinomas. Meanwhile, g*BRCA2*, g*CHEK2* and g*ATM* carriers typically had hormone receptor-positive and HER2-negative ($HR^+/HER2^-$) disease (75.2%, 74.7% and 71.7%, respectively; Supplementary Table 2), consistent with previous studies[25,26].

Biallelic loss is often a necessary condition to observe a phenotypic impact in carriers of germline pathogenic variants in HRD-related genes, yet its incidence varies by gene, with higher rates of biallelic inactivation observed in high-penetrance genes[27,28]. Our results confirmed that biallelic inactivation rates varied significantly across genes, ranging from 50.6% ($n = 44$) in g*CHEK2* carriers to 77.5% ($n = 110$) and 75.8% ($n = 122$) for g*BRCA1* and g*BRCA2* carriers, respectively (Extended Data Fig. 1b). We also found lower frequency of loss of heterozygosity (LOH) in g*PALB2* carriers (51.4%, $n = 18$), but a relatively higher frequency of 'second-hit' somatic mutations resulting in biallelic loss (33.3%, $n = 6$), concurring with previous literature[29]. Confirming the associations between histological, demographic and genomic patterns with germline pathogenic variants establishes the clinical relevance of this clinicogenomic cohort.

Germline–somatic genomic interactions

We first sought to define patterns of mutual exclusivity or enrichment of somatic variants with germline pathogenic variants (germline–somatic interactions), in the context of breast cancer receptor subtype and zygosity (Fig. 1a). This analysis robustly validated previously reported enrichment of *TP53* alterations in g*BRCA1* carriers and mutual exclusivity of g*ATM* and *TP53* alterations[30,31] (Fig. 1b,c). Both findings were more pronounced in tumours exhibiting biallelic inactivation of the respective genes (Extended Data Fig. 2a,e). We further explored these observations by conducting germline–somatic interaction analyses stratified by breast cancer receptor subtypes. The g*BRCA1*–*TP53* interaction was enriched in $HR^+/HER2^-$ tumours but not in triple-negative tumours, where *TP53* variants are already highly prevalent in *BRCA1* WT tumours, supporting *TP53* loss of function (LoF) as a ubiquitous event in the oncogenesis of g*BRCA1*-driven breast cancers[32,33].

In g*BRCA2*-driven breast cancers, *RB1* somatic variants were significantly enriched, in stark contrast to their absence in g*BRCA1*-associated tumours (Fig. 1c–e and Extended Data Fig. 2b). This observation is notable, as g*BRCA1* tumours are largely triple negative, a subtype typically enriched for *RB1* alterations. Focusing on patients with $HR^+/HER2^-$ tumours (75.2% of g*BRCA2*) revealed an even higher enrichment of somatic *RB1* alterations in g*BRCA2* carriers. This analysis also uncovered an enrichment of *MYC* and *AURKA* amplifications. Of clinical relevance, *RB1*, *MYC* and *AURKA* alterations have all been implicated in resistance to CDK4/6 inhibition[34,35]. Conversely, *PIK3CA* alterations were more enriched in gWT cancers than in g*BRCA2* and g*BRCA1*-driven cancers. The receptor status and LOH-specific germline–somatic interactions are further displayed in Extended Data Fig. 2, and corresponding Oncoprints are detailed in Extended Data Fig. 3 (Supplementary Tables 3–7).

Clinical implications of g*BRCA2* status

CDK4/6i combined with endocrine therapy (CDK4/6i + ET) represents the cornerstone of treatment for patients with metastatic and high-risk early-stage HR^+/breast cancer[36–38], with *RB1* loss established as a key mechanism of resistance to CDK4/6i[39]. On the basis of our results demonstrating a significant enrichment of *RB1* alterations in g*BRCA2* carriers, we analysed the effect of g*BRCA2* status on progression-free survival (PFS) in patients with $HR^+/HER2^-$ metastatic breast cancer (MBC) treated with CDK4/6i + ET. g*BRCA2* pathogenic variants were associated with a significantly shorter PFS on CDK4/6i + ET in univariate and multivariate analyses (median PFS of 9.0 versus 15.6 months, multivariate hazard ratio (HR) = 2.17, 95% CI 1.60–2.96, $P < 0.00001$; Fig. 2a). Similar results were seen when the analysis was extended to all treatment lines, with consideration of ET partner and treatment line as covariates (HR = 1.97, 95% CI 1.55–2.51, $P < 0.00001$; Fig. 2b and Supplementary Table 8). LoF mutations in *RB1* were rare in pre-treatment samples (2%), and exclusion of these cases did not alter the results.

We next validated our findings using an independent, nationwide clinicogenomic dataset containing manually curated patient-level outcomes data from both community oncology settings and academic medical centres[40,41]. This analysis confirmed a strong association between g*BRCA2* and shorter PFS on CDK4/6i + ET (median PFS of 12.0 versus 18.3 months, HR = 1.83, 95% CI 1.36–2.48, $P < 0.00001$; Fig. 2c) among the 2,158 patients with $HR^+/HER2^-$ MBC treated with first-line CDK4/6i + ET combinations. To evaluate whether g*BRCA2* was associated with resistance specifically to CDK4/6i + ET, we expanded our analysis to assess the effect of g*BRCA2* status on PFS of other common therapeutic modalities in breast cancer (Extended Data Fig. 4a). g*BRCA2* status did not affect outcome on the vast majority of these therapies, nor did it significantly effect overall survival for patients starting on first-line CDK4/6i + ET. This demonstrates that g*BRCA2* status is not a universal determinant of outcome but instead confers context-dependent relevance to specific therapies.

We further investigated the MSK cohort focusing on patients with $HR^+/HER2^-$ MBC who received a PARPi after progression on CDK4/6i ($n = 41$). Of note, despite being administered in later lines (median line of therapy = 3), HRD-directed therapy generally resulted in superior outcomes (Fig. 2d). PARPi treatment resulted in a significantly improved PFS compared with the preceding CDK4/6i regimen (HR = 0.38, 95% CI 0.19–0.76, $P = 0.0062$), with 73.2% of patients achieving a longer PFS on PARPi than on frontline CDK4/6i. This clinical benefit was more pronounced among the patients who had failed to respond to previous CDK4/6i + ET (Fisher's exact test; OR = 8.87, 95% CI 2.84–27.98, $P = 0.00002$; Extended Data Fig. 4b). Among patients with evaluable imaging who did not discontinue therapy due to early toxicity ($n = 38$), PARPi achieved a partial or complete response in 84.2% ($n = 32$) of patients, compared with only 39.5% ($n = 15$) for previous CDK4/6i + ET (Extended Data Fig. 4c). Representative cases of rapid progression through first-line CDK4/6i + ET, followed by prolonged complete response to PARPi are highlighted in Fig. 2e,f. Collectively, these results provide a clinical rationale for prioritization of PARPi for this genomically defined subgroup of patients with breast cancer with expected poor outcomes on CDK4/6i-based combinations.

Table 1 | Patient demographics

	ATM (*n*=60)	*BRCA1* (*n*=142)	*BRCA2* (*n*=161)	*CHEK2* (*n*=87)	*PALB2* (*n*=35)	WT (*n*=5,094)	*P* value
Overall receptor status							
$HR^+/HER2^-$	43 (71.7%)	35 (24.6%)	121 (75.2%)	65 (74.7%)	22 (62.9%)	3,426 (67.3%)	<0.00001
$HER2^+$	15 (25.0%)	7 (4.9%)	15 (9.3%)	15 (17.2%)	2 (5.7%)	766 (15.0%)	
TNBC	2 (3.3%)	100 (70.4%)	25 (15.5%)	7 (8.0%)	11 (31.4%)	902 (17.7%)	
Age at diagnosis (years)							
Younger than 35	10 (16.7%)	52 (36.6%)	34 (21.1%)	14 (16.1%)	3 (8.6%)	488 (9.6%)	<0.00001
35–50	30 (50.0%)	50 (35.2%)	69 (42.9%)	33 (37.9%)	20 (57.1%)	1,952 (38.3%)	
51–60	4 (6.7%)	22 (15.5%)	39 (24.2%)	24 (27.6%)	8 (22.9%)	1,382 (27.1%)	
61–70	13 (21.7%)	12 (8.5%)	16 (9.9%)	10 (11.5%)	3 (8.6%)	929 (18.2%)	
Older than 70	3 (5.0%)	6 (4.2%)	3 (1.9%)	6 (6.9%)	1 (2.9%)	342 (6.7%)	
Unknown	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (0.0%)	
Overall histological grade							
I-low	0 (0%)	0 (0%)	0 (0%)	1 (1.1%)	1 (2.9%)	199 (3.9%)	<0.00001
II-intermediate	17 (28.3%)	11 (7.7%)	42 (26.1%)	29 (33.3%)	15 (42.9%)	1,332 (26.1%)	
III-high	31 (51.7%)	121 (85.2%)	96 (59.6%)	46 (52.9%)	14 (40.0%)	3,005 (59.0%)	
Unknown	12 (20.0%)	10 (7.0%)	23 (14.3%)	11 (12.6%)	5 (14.3%)	558 (11.0%)	
Histology							
IDC	50 (83.3%)	131 (92.3%)	132 (82.0%)	70 (80.5%)	31 (88.6%)	4,046 (79.4%)	0.0291
ILC	5 (8.3%)	4 (2.8%)	15 (9.3%)	7 (8.0%)	2 (5.7%)	607 (11.9%)	
Other, mixed or unknown	5 (8.3%)	7 (4.9%)	14 (8.7%)	10 (11.5%)	2 (5.7%)	441 (8.7%)	
Stage at diagnosis							
I	13 (21.7%)	39 (27.5%)	33 (20.5%)	20 (23.0%)	10 (28.6%)	1,265 (24.8%)	0.0213
II	15 (25.0%)	64 (45.1%)	54 (33.5%)	28 (32.2%)	11 (31.4%)	1,763 (34.6%)	
III	9 (15.0%)	14 (9.9%)	27 (16.8%)	12 (13.8%)	3 (8.6%)	822 (16.1%)	
IV	20 (33.3%)	15 (10.6%)	41 (25.5%)	23 (26.4%)	7 (20.0%)	1,002 (19.7%)	
Unknown	3 (5.0%)	10 (7.0%)	6 (3.7%)	4 (4.6%)	4 (11.4%)	242 (4.8%)	
ER IHC expression (%)							
Less than 10	1 (1.7%)	16 (11.3%)	7 (4.3%)	2 (2.3%)	0 (0%)	206 (4.0%)	<0.00001
11–40	6 (10.0%)	98 (69.0%)	27 (16.8%)	14 (16.1%)	14 (40.0%)	1,227 (24.1%)	
41–60	3 (5.0%)	2 (1.4%)	5 (3.1%)	2 (2.3%)	0 (0%)	116 (2.3%)	
61–90	4 (6.7%)	8 (5.6%)	20 (12.4%)	7 (8.0%)	4 (11.4%)	509 (10.0%)	
More than 90	46 (76.7%)	18 (12.7%)	100 (62.1%)	59 (67.8%)	16 (45.7%)	2,963 (58.2%)	
Unknown	0 (0%)	0 (0%)	2 (1.2%)	3 (3.4%)	1 (2.9%)	73 (1.4%)	
Progresterone receptor IHC expression (%)							
Less than 10	7 (11.7%)	11 (7.7%)	18 (11.2%)	8 (9.2%)	5 (14.3%)	515 (10.1%)	<0.00001
11–40	18 (30.0%)	118 (83.1%)	75 (46.6%)	35 (40.2%)	15 (42.9%)	2,146 (42.1%)	
41–60	8 (13.3%)	2 (1.4%)	18 (11.2%)	6 (6.9%)	5 (14.3%)	296 (5.8%)	
61–90	7 (11.7%)	5 (3.5%)	21 (13.0%)	6 (6.9%)	4 (11.4%)	658 (12.9%)	
More than 90	19 (31.7%)	6 (4.2%)	27 (16.8%)	29 (33.3%)	5 (14.3%)	1,371 (26.9%)	
Unknown	1 (1.7%)	0 (0%)	2 (1.2%)	3 (3.4%)	1 (2.9%)	108 (2.1%)	

Demographics and clinicopathological characteristics are reported for our cohort of patients. Of note, there were 5,579 distinct genes represented, including 13 patients who had more than one germline pathogenic variant in our genes of interest. Comparisons between germline pathogenic variant status and gWT were performed using two-sided Chi-square testing, and the corresponding *P* values are reported. There were no adjustments for multiple comparisons. For $P < 0.00001$, results are reported as <0.00001. Exact *P* values are summarized below: overall receptor status: $P = 9.7\times10^{-55}$; age at diagnosis: $P = 2.7\times10^{-22}$; menopausal status at diagnosis: $P = 1.1\times10^{-12}$; overall histological grade: $P = 1.0\times10^{-8}$; ER immunohistochemistry (IHC) expression: $P = 1.5\times10^{-32}$; and progresterone receptor IHC expression: $P = 6.7\times10^{-17}$. IDC, invasive ductal carcinoma; ILC, invasive lobular carcinoma; TNBC, triple-negative breast cancer.

Acquired *RB1* LoF in g*BRCA2*

To investigate the implications of germline status on the development of *RB1* loss, an established mechanism of CDK4/6i resistance, we assembled a large cohort of patients with paired tumour samples collected pre-CDK4/6i and post-CDK4/6i (1,312 tumour and 513 plasma circulating tumour DNA (ctDNA) samples from 546 patients). Genomic analysis of this paired pre-treatment and post-progression cohort demonstrated that acquired somatic *RB1* LoF alterations were significantly more prevalent in g*BRCA2* tumours versus g*BRCA2* WT tumours (28.6% versus 7.1%, OR = 5.17, 95% CI 2.07–13.0, $P = 0.0010$; Fig. 2g). The clinical responses and evolutionary trajectories of patients treated with CDK4/6i are highlighted by two representative patients with g*BRCA2* $HR^+/HER2^-$ MBC (Extended Data Fig. 4d,e). In both cases, acquired polyclonal *RB1* LoF variants emerged post-CDK4/6i, underscoring a strong and specific evolutionary pressure for *RB1* LoF mutations as a dominant mechanism of CDK4/6i resistance in these g*BRCA2* tumours.

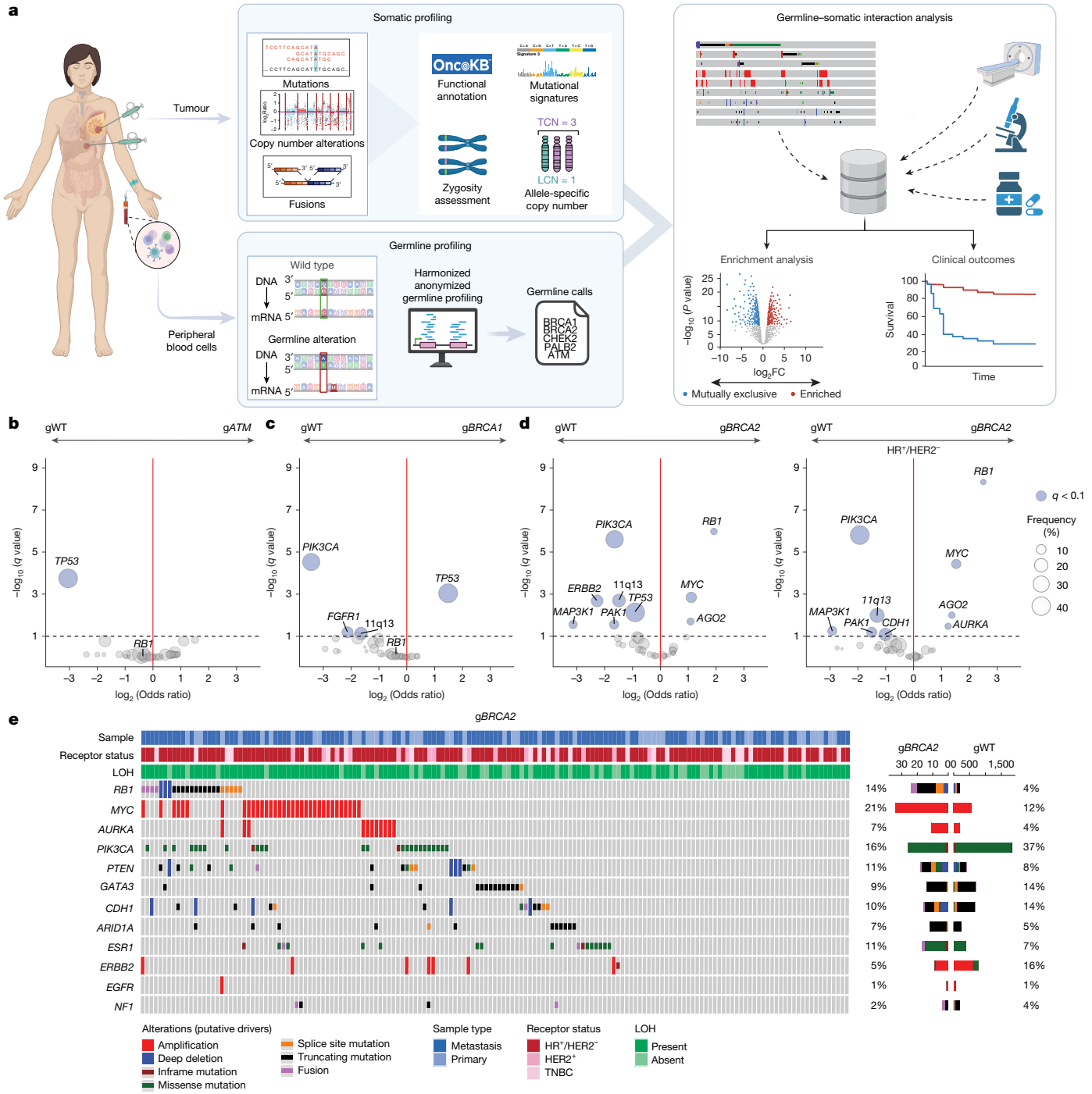


Fig. 1 | Germline–somatic interactions in breast cancer. a, Study schema. Enrichment analysis was performed to identify somatic alterations more prevalent in patients with pathogenic germline variants of homologous recombination pathway genes (compared with 'sporadic' or gWT cases). FC, fold change; LCN, lesser copy number; TCN, total copy number. Schematic created in BioRender; Razavi, P. https://biorender.com/hv0bvbr (2025). **b**, Enrichment analysis of g*ATM* ($n = 60$) compared with gWT ($n = 5{,}094$), as determined by Firth-penalized logistic regression with sample type and receptor status as covariates. Significant genes (blue) are defined by $q < 0.10$. **c**, As in panel **b**, but for comparison of g*BRCA1* ($n = 142$) with gWT ($n = 5{,}904$). g*BRCA1* tumours were enriched for somatic *TP53* variants; OR = 2.80 (95% CI 1.70–4.59, $q = 0.0009$) and mutually exclusive with *PIK3CA* variants. **d**, As in panels **b**,**c**, but for comparison of g*BRCA2* ($n = 161$) compared with gWT ($n = 5{,}094$). A subset analysis of HR$^+$/HER2$^-$ tumours ($n_{Total} = 3{,}547$, $n_{gBRCA2} = 121$ and $n_{gWT} = 3{,}426$) was then performed. Somatic *RB1* variants represented the most significant enrichment in g*BRCA2* carriers among this subgroup; OR = 5.66 (95% CI 3.34–9.60, $q = 4.7 \times 10^{-9}$). **e**, Oncoprint showing mutations, copy number deletions and fusions in the indicated genes in g*BRCA2* versus gWT. Receptor status, sample type and zygosity are annotated above.

Baseline allelic state forecasts *RB1* loss

Consideration of the genomic architecture of *BRCA2*-driven tumours provides further insight into the acquisition of *RB1* LoF alterations. *BRCA2* and *RB1* are syntenic on chromosome 13q, and previous work suggests that biallelic inactivation of *BRCA2* in g*BRCA2*-driven tumours often occurs through loss of a large chromosomal segment inclusive of WT *BRCA2* and *RB1* alleles[42–47]. Consistent with this, the pattern of *RB1* LOH was influenced by g*BRCA2* status in our cohort, demonstrating a co-occurrence of *BRCA2* and *RB1* LOH in g*BRCA2*

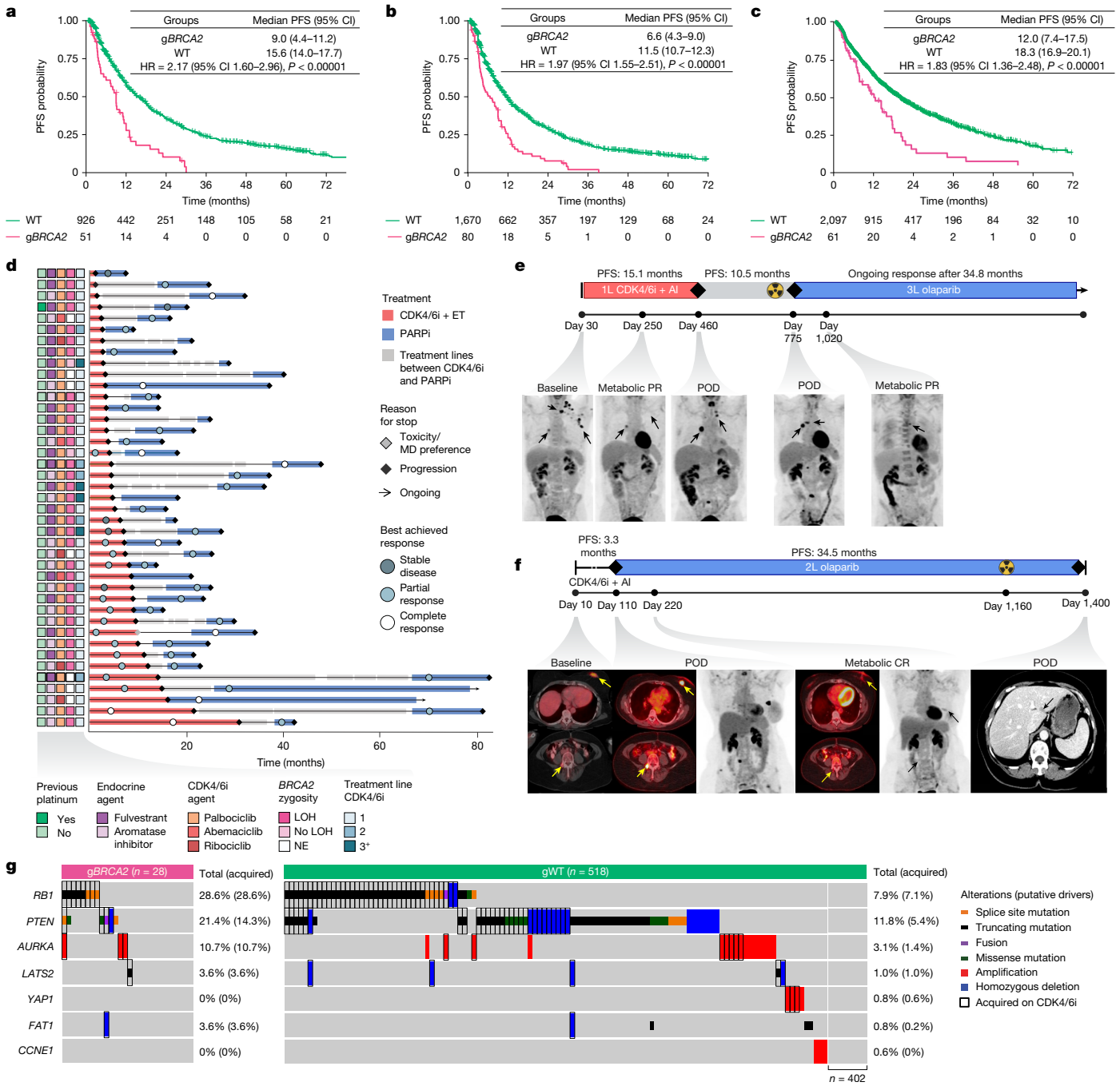


Fig. 2 | Clinical implications of g*BRCA2* status. a, PFS for patients treated with first-line CDK4/6i + ET by g*BRCA2* status. Patients with g*BRCA2* ($n = 51$) were compared with gWT ($n = 926$); $P = 8.0 \times 10^{-7}$. **b**, As in panel **a**, but depicting all lines of treatment. Patients with g*BRCA2* ($n = 80$) were compared with gWT ($n = 1{,}670$); $P = 2.9 \times 10^{-8}$. **c**, PFS on first-line CDK4/6i + ET combinations by g*BRCA2* status from the multi-institutional external validation cohort, which included 61 patients with the g*BRCA2* pathogenic variant and 2,097 with gWT. HRs in panels **a**–**c** were estimated using Cox proportional hazard models (*P* values from two-sided Wald tests). No multiple comparisons adjustment was made. **d**, Swimmer's plot of patients ($n = 41$) receiving PARPi (blue bar) after CDK4/6i + ET (red bar); HR_{PARPi} versus $HR_{CDK4/6i} = 0.38$ (95% CI 0.19–0.76, $P = 0.0062$), based on Cox proportional hazard model stratified by patient ID. Best response (complete or partial response) was compared for PARPi versus CDK4/6i + ET using a two-sided Fisher's exact test; OR = 8.87 (95% CI 2.84–27.98, $P = 0.00002$). MD, physician; NE, non-evaluable. **e**,**f**, Two representative cases demonstrating short PFS on first-line (1L) CDK4/6i + ET and subsequent metabolic complete response (CR) and durable disease control on PARPi. 2L, second line; 3L, third line; AI, aromatase inhibitor; POD, progression of disease; PR, partial response. Black and yellow arrows denote radiographic tumour activity. **g**, Paired pre-CDK4/6i or post-CDK4/6i oncoprint by g*BRCA2* status. The black rectangle denotes acquired variant. Two-sided Fisher's exact test of *RB1* LoF variant acquisition by g*BRCA2* status; OR = 5.17 (95% CI 2.07–13.0, $P = 0.0010$).

tumours compared with gWT (Fig. 3a). We validated this finding using an external cohort of 46 g*BRCA2*-associated primary breast cancers profiled with whole-exome sequencing (24 patients from the University of Pennsylvania Abramson Cancer Center and 22 from the Mayo Clinic). Consistently, 82.6% ($n = 38$) of tumours in this cohort demonstrated concurrent LOH of *BRCA2* and *RB1* ($P < 0.0001$; Extended Data Fig. 5). Although *RB1* LOH was enriched in g*BRCA2* tumours, it notably was also present in 34.9% of gWT tumours ($n = 601$ of 1,723 of evaluable cases), suggesting a broader relevance for this allelic state.

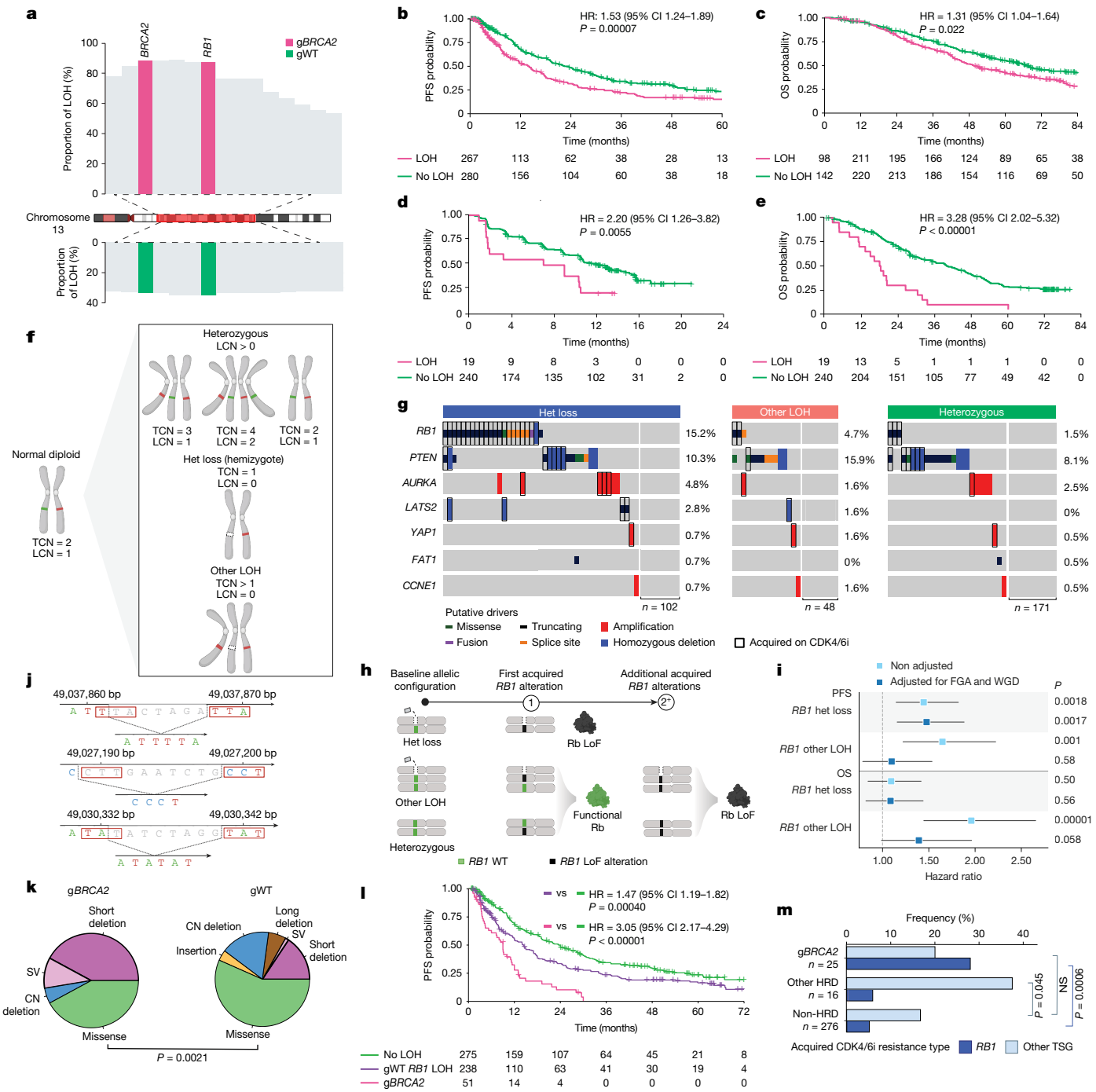


Fig. 3 | Dual effect of *RB1* LOH and HRD mutagenesis on CDK4/6i resistance. **a**, *RB1* and *BRCA2* LOH in g*BRCA2* ($n = 63$; top) versus gWT ($n = 527$; bottom); OR = 11.5 (95% CI 6.16–22.2, $P < 2.2 \times 10^{-16}$). **b**,**c**, PFS on first-line CDK4/6i + ET (**b**) and OS (**c**) by *RB1* LOH, with numbers of LOH and no LOH given below. **d**,**e**, PALOMA-3 PFS (**d**) and OS (**e**) by *RB1* LOH; OS $P = 1.5 \times 10^{-6}$. The numbers of LOH and no LOH are given below. **f**, Allelic configurations. **g**, Paired pre-CDK4/6i or post-CDK4/6i oncoprint ($n_{RB1\text{-Het loss}} = 145$, $n_{\text{Other-LOH}} = 63$ and $n_{\text{Heterozygous}} = 197$); $OR_{\text{Het loss versus heterozygous}} = 11.5$ (95% CI 3.50–46.1, $P = 1.9 \times 10^{-6}$) and $OR_{\text{Other-LOH versus heterozygous}} = 3.22$ (95% CI 0.57–18.21, $P = 0.16$). **h**, *RB1* het loss predisposes to *RB1* 'second hit'. **i**, OS and PFS on first-line CDK4/6i + ET by *RB1* ASCN ($n_{RB1\text{-Het loss}} = 180$, $n_{RB1\text{-Other-LOH}} = 83$ and $n_{\text{Heterozygous}} = 284$). Error bars denote 95% CI. FGA, fraction genome altered; WGD, whole-genome duplication. **j**, Select *RB1* variants in g*BRCA2* tumours, short indels flanked by microhomology. Schematics created in BioRender: **f**, Razavi, P. https://biorender.com/utmr5e5 (2025); **h**, Razavi, P. https://biorender.com/bewn6wg (2025); **j**, Razavi, P. https://biorender.com/gxjrb6s (2025). **k**, HRD-associated *RB1* variants (short indels and structural variants (SVs)) by g*BRCA2*; OR = 5.36 (95% CI 1.76–16.58). CN, copy number. **l**, PFS on first-line CDK4/6i + ET by *RB1* ASCN and g*BRCA2*; HR_{gBRCA2} $P = 1.2 \times 10^{-10}$; numbers are given below. **m**, Acquired *RB1* versus other TSG LoF. HRD association with non-*RB1* TSG LoF: OR = 2.99 (95% CI 1.03–9.34, $P = 0.045$). ORs by two-sided Fisher's exact test; HRs by multivariate Cox proportional hazard models (P from two-sided Wald test). No multiple comparisons adjustment were made. NS, not significant.

On the basis of our *BRCA2*–*RB1* co-LOH results, we hypothesized that *RB1* LOH is a predominant predisposing factor for CDK4/6i resistance. Hence, we extended our survival analyses to assess whether pre-treatment *RB1* zygosity was predictive of clinical benefit from CDK4/6i beyond the context of g*BRCA2* status. Specifically, we analysed 547 patients with $HR^+/HER2^-$ MBC in the MSK cohort who received

first-line CDK4/6i and had evaluable LOH status in the pre-treatment tumours. *RB1* LOH was observed in 47.8% of these patients and was predictive of both significantly shorter PFS (HR = 1.53, 95% CI 1.24–1.89, $P = 0.00007$) and overall survival (OS; HR = 1.31, 95% CI 1.04–1.64, $P = 0.022$; Fig. 3b,c and Supplementary Table 9).

To confirm these findings, we performed an analysis of pre-treatment ctDNA samples collected as part of the PALOMA-3 study[48], the pivotal randomized phase III trial of palbociclib plus fulvestrant versus fulvestrant monotherapy. *RB1* LOH was observed in 19 (7.3%) of 259 pre-treatment ctDNA samples and was associated with a significantly shorter PFS (HR = 2.20, 95% CI 1.26–3.82, $P = 0.0055$) and OS (HR = 3.28, 95% CI 2.02–5.32, $P < 0.00001$; Fig. 3d,e). These trends were more pronounced in the palbociclib plus fulvestrant arm (Extended Data Fig. 6).

We further reasoned that the specific number of functional pre-treatment *RB1* alleles influences the likelihood of acquiring a second-hit *RB1* LoF alteration under the therapeutic pressure of CDK4/6i therapy. Specifically, we hypothesized that tumours with a single functional *RB1* copy — reflecting *RB1* hemizygosity (heterozygous loss (het loss)) — would be statistically more likely to develop an *RB1* LoF variant as a second hit to achieve biallelic inactivation and resistance. To test this, we investigated the clinical relevance of pre-treatment *RB1* het loss, as compared with other allelic configurations such as loss before or after whole-genome doubling resulting in more than one remaining WT *RB1* alleles (Fig. 3f). We extended our analysis of pre-CDK4/6i and post-CDK4/6i tumours ($n = 405$ patients) to these refined *RB1* allele-specific copy number (ASCN) definitions. Consistent with our hypothesis, only pre-treatment *RB1* het loss uniquely predisposed to acquisition of *RB1* LoF upon exposure to CDK4/6i ($P < 0.00001$; Fig. 3g). Acquired *RB1* LoF events occurred almost exclusively in tumours with baseline *RB1* het loss (total, 15.2%; acquired, 14.5%), whereas they were rare in other allelic settings. This suggests that *RB1* hemizygosity lowers the evolutionary barrier to resistance, requiring only a single additional hit to achieve complete inactivation, whereas other allelic configurations require multiple independent events, making resistance via this route far less likely (Fig. 3h). To underscore *RB1* loss as a specific mechanism of resistance to CDK4/6i, we leveraged our MSK cohort to categorize each sequenced sample by previous therapeutic exposure. Among all treatment classes (including ET monotherapy, chemotherapy, antibody–drug conjugates and targeted therapies), only CDK4/6i exposure was significantly associated with the enrichment of acquired *RB1* LoF variants (Extended Data Fig. 7).

To distinguish between predictive versus prognostic properties of these changes, we repeated the outcome analyses based on the redefined pre-treatment *RB1* allelic state (Extended Data Fig. 8). Both het loss (32.7%, $n = 181$) and other LOH allelic configurations (15.0%, $n = 83$) were associated with worse PFS on CDK4/6i. As chromosomal instability is a known poor prognostic factor[49], we repeated this analysis adjusting for fraction genome altered and whole-genome doubling. Of note, *RB1* het loss was the sole significant allelic configuration associated with shorter PFS in this multivariate analysis. Conversely, other LOH configurations were uniquely associated with worse OS (HR = 1.96, 95% CI 1.45–2.66, $P = 0.00001$; Fig. 3i and Supplementary Table 8), suggesting they represent a more general prognostic marker, whereas *RB1* hemizygosity specifically serves as a predictive marker of outcomes on CDK4/6i.

Together, although the acquisition of *RB1* LoF variants conferring resistance to CDK4/6i is relatively uncommon in an unselected population[50] (8.9% in our paired pre-CDK4/6i-treated and post-CDK4/6i-treated cohort), our results demonstrate that acquired *RB1* loss is both prevalent and highly predictable in tumours with a single functional *RB1* allele before treatment. These findings have major clinical implications for the development of novel surveillance strategies and for patient selection for clinical trials of precision therapeutic strategies directed against complete *RB1* loss[51,52]. More broadly, our results suggest that the pre-treatment allelic structure can forecast not only therapeutic outcomes but also the specific molecular trajectories through which targeted therapies fail.

HRD facilitates acquired resistance

We next assessed whether g*BRCA2*-mediated HRD specifically contributes to somatic *RB1* LoF variants during CDK4/6i therapy. Hence, we analysed the types of alterations in *RB1* found in our g*BRCA2* cohort. We observed that somatic *RB1* variants identified in g*BRCA2* tumours predominantly exhibited the characteristic pattern of microhomology flanking a short deletion, a hallmark of back-up DNA repair mechanism through microhomology-mediated end joining in HRD-driven tumours[53] (Fig. 3j). Indeed, we observed significant enrichment for short indels or structural variants within the coding sequence of *RB1* in the g*BRCA2* setting, indicative of *BRCA2* deficiency-mediated lack of competent DNA repair (Fig. 3k).

Integration of our genomic analysis with treatment response data also suggested that HRD-mediated mutagenesis was probably responsible for the exceptionally rapid progression observed in g*BRCA2* carriers treated with CDK4/6i combinations. Although *RB1* LOH was associated with a significantly shorter PFS in gWT cases (HR = 1.47, 95% CI 1.19–1.82, $P = 0.0004$), in the context of g*BRCA2*, *RB1* LOH resulted in an even shorter PFS (HR = 3.05, 95% CI 2.17–4.29, $P < 0.00001$; Fig. 3l). Thus, g*BRCA2* breast cancers with *RB1* LOH appear to exploit the imperfections of back-up DNA repair mechanisms to escape therapeutic pressure. This phenomenon resembles the concept of 'exaptation', which refers to an evolutionary co-option of a trait to serve a purpose distinct from its original role[54].

We expanded our analysis to assess the role of HRD (of either germline or somatic origin) and clinical outcomes on CDK4/6i. The presence of a HRD-associated signature in pre-treatment tumours of patients with $HR^+/HER2^-$ MBC conferred a significantly worse outcome on CDK4/6i + ET (HR = 1.78, 95% CI 1.27–2.48, $P = 0.00075$), even outside the context of g*BRCA2* (Extended Data Fig. 9). Having established that this phenomenon extends beyond *BRCA2*, we next investigated whether a similar HRD-driven second-hit pattern was evident across other tumour suppressor genes (TSGs) implicated in CDK4/6i resistance. Non-*BRCA2* HRD tumours showed a significantly increased tendency to acquire resistance via TSGs other than *RB1* (37.5% of cases, OR = 2.99, 95% CI 1.03–9.34, $P = 0.045$) relative to non-HRD cases. The majority of cases with these acquired TSG LoF alterations were preceded by het loss of the respective gene in the pre-treatment setting (Fig. 3m).

In summary, the data suggest that both an increased prevalence in *RB1* LOH and ongoing HRD-driven mutagenesis, which hastens the acquisition of *RB1* LoF mutations, contribute to the poor outcomes of patients with breast cancer with g*BRCA2* mutations treated with CDK4/6i + ET. Our findings demonstrate that this concept is not limited to g*BRCA2* and *RB1*, but rather reflects a broader principle linking HRD biology and baseline allelic configurations to vulnerabilities across TSGs implicated in therapy resistance.

Validation in xenograft models

To directly validate the role of Rb loss as a driver of CDK4/6i resistance in g*BRCA2*-associated tumours, we generated several patient-derived xenografts (PDXs) from *BRCA2* carriers with $HR^+/HER2^-$ breast cancer. The first set of PDXs was derived from a 30-year-old g*BRCA2* pathogenic variant carrier with de novo MBC treated at MSK (Fig. 4a). Clinically, this patient received abemaciclib in combination with pembrolizumab in the second-line setting as part of a trial, followed by standard-of-care CDK4/6i + ET with rapid disease progression. A biopsy of the rapidly progressing chest wall lesion revealed an *RB1* LoF variant, which was absent in the pre-CDK4/6i metastatic biopsy. This patient subsequently received radiation therapy to the progressive

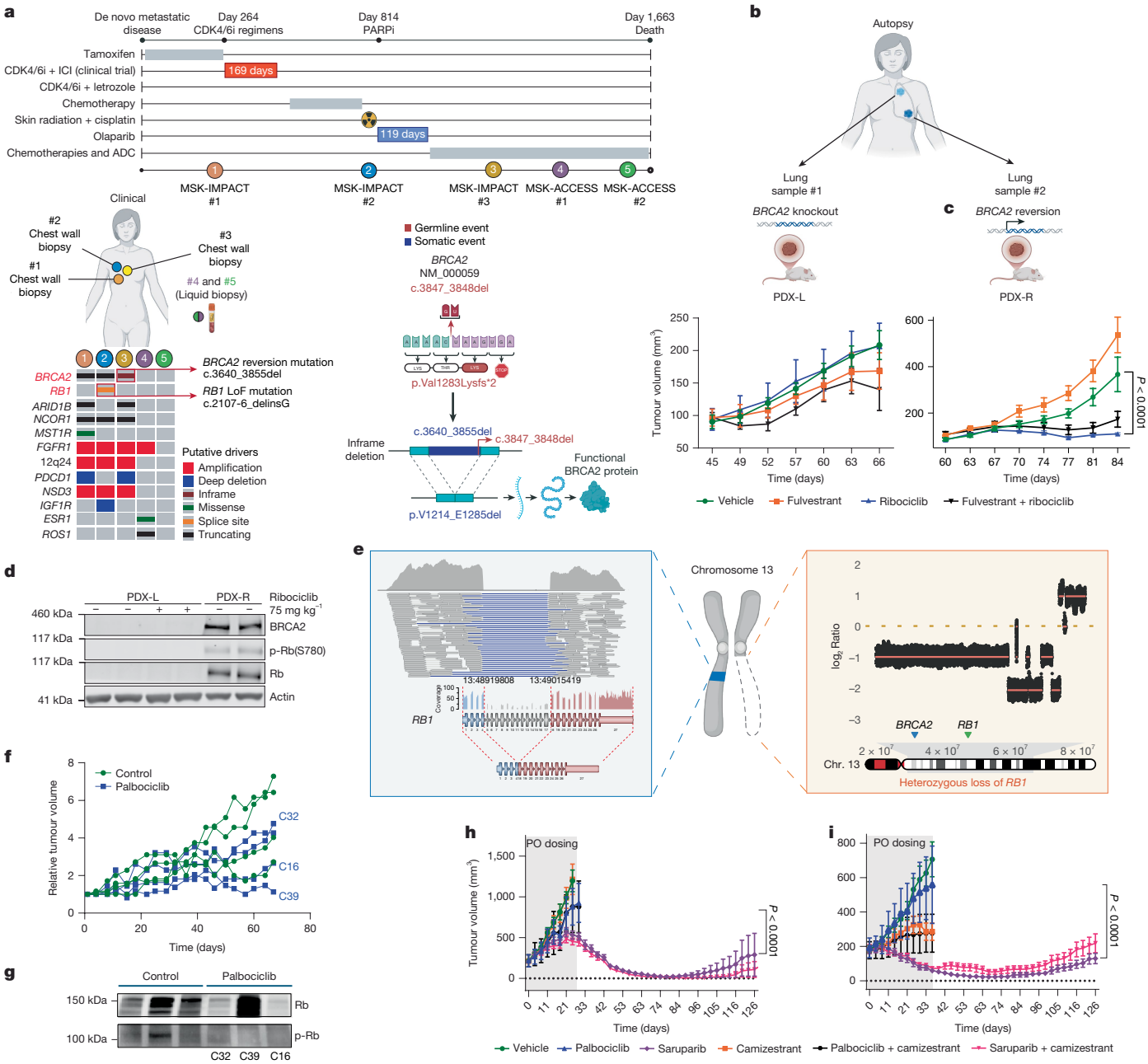


Fig. 4 | Mechanistic validation of Rb loss as a driver of CDK4/6i resistance in g*BRCA2* tumours. a, PDXs derived from two metastatic deposits in a rapid autopsy study of an MSK patient with g*BRCA2*: PDX-R and PDX-L. ADC, antibody–drug conjugate; ICI, immune checkpoint inhibitor. Schematic created in BioRender; Razavi, P. https://biorender.com/gjwhomb (2025). **b**,**c**, In vivo efficacy study of PDX-L ($n = 5$; **b**) and PDX-R ($n = 4$; **c**) treated with vehicle, fulvestrant (200 mg kg^{-1} subcutaneous, twice weekly), ribociclib (75 mg kg^{-1} PO 5 days per week), or fulvestrant and ribociclib. PO, oral dosing. $P = 1.29 \times 10^{-13}$, as calculated by a linear mixed-effect model. Data are mean ± s.e.m. **d**, Western blot analysis of PDX-L and PDX-R with vehicle and ribociclib ($n = 3$ replicates). Actin was used as a loading control. **e**, Schematic representing the *RB1* structural variant from whole-genome sequencing of a CDK4/6i-resistant PDX-L sample, resulting in deletion of exons 5–17. Chr. chromosome. **f**, PDX-C was derived from a patient with g*BRCA2*. Vehicle was compared with palbociclib (75 mg kg^{-1} PO 5 days per week). Data of individual mice are shown to demonstrate that tumours developing relative resistance to palbociclib also harboured Rb loss, as demonstrated in the western blot in panel **g** ($n = 3$ replicates). **h**, In vivo efficacy study of PDX-P1. PDXs were treated with vehicle, camizestrant (10 mg kg^{-1} PO daily), palbociclib (50 mg kg^{-1} PO daily), saruparib (1 mg kg^{-1} PO daily) or combination (camizestrant + palbociclib and camizestrant + saruparib). PO dosing was discontinued following takedown of the vehicle arm. Experiments included $n = 5$ mice in the palbociclib ± camizestrant arms and $n = 10$ mice in all other arms. $P = 1.55 \times 10^{-8}$. **i**, In vivo efficacy study of PDX-P2, with arms as defined in panel **h**. Experiments included $n = 7$ mice in the camizestrant monotherapy arm and $n = 8$ mice in all other arms. $P = 2.72 \times 10^{-77}$. For panels **h**,**i**, data are mean ± s.e.m.; P values were determined by two-way repeated measures analysis of variance (ANOVA) followed by Sidak's multiple-comparisons test. Dotted horizontal line indicates tumour volume = 0.

chest wall disease concurrently with systemic platinum, resulting in a major response. This patient then received treatment with PARPi, until resistance emerged through acquisition of a *BRCA2* reversion variant (c.3640_3855del). Despite several additional lines of cytotoxic chemotherapy, the patient eventually succumbed to the disease and participated in the MSK Last Wish Program (MSK IRB Protocol #15-021), enabling a rapid autopsy and post-mortem tissue collection and analysis.

Multi-site sequencing of autopsy samples did not reveal any evidence of previously detected *RB1* somatic variant by deep targeted

next-generation sequencing panel testing, indicating that this specific subclone was probably eradicated by subsequent therapies received after CDK4/6i exposure (Extended Data Fig. 10a and Supplementary Table 10). However, PDXs generated from two separate metastatic deposits in the lungs provided a unique opportunity to interrogate the evolutionary consequences of active HRD in the context of g*BRCA2* and *RB1* LOH. One PDX demonstrated *BRCA2* reversion variant (PDX-R), whereas the other had persistent biallelic *BRCA2* loss and no evidence of reversion variant on deep sequencing (PDX-L; Extended Data Fig. 10b). Otherwise, PDX-L and PDX-R displayed virtually identical copy number patterns, reflecting a strong genomic concordance between the two samples (Extended Data Fig. 10c).

When challenged with therapy, PDX-R tumours responded to ribociclib-based therapy, whereas PDX-L tumours were uniformly resistant (Fig. 4b,c). In the resistant PDX-L tumours, we observed no BRCA2 or Rb expression, whereas expression of both proteins was retained in sensitive PDX-R tumours (Fig. 4d). Whole-genome sequencing of CDK4/6i-resistant tumours from PDX-L identified a large structural variant in *RB1*, characterized by deletion of exons 5–17 (Fig. 4e). This acquired structural variant, combined with pre-existing *RB1* het loss, resulted in complete loss of *RB1*. Such large deletions are a recognized hallmark of HRD-mediated instability[55], highlighting the role of HRD in facilitating *RB1* loss and validating our clinicogenomic observations in a rigorously controlled preclinical setting.

We next validated these findings in another PDX derived from a g*BRCA2* carrier (PDX-C). The majority of samples developed resistance to palbociclib monotherapy (Fig. 4f), and Rb loss of expression was again observed in post-treatment tumours that were resistant to CDK4/6i, but not in CDK4/6i-sensitive tumours (Fig. 4g). The convergence of our clinicogenomic observations and the evidence from preclinical models implicating HRD in CDK4/6i resistance led us to compare the differential therapeutic response patterns between CDK4/6i and PARPi. In two additional independent g*BRCA2* PDXs (PDX-P1 and PDX-P2), we demonstrated clear resistance in each CDK4/6i ± ET condition (palbociclib or camizestrant monotherapy and combination therapy), whereas sensitivity to selective PARP1 inhibition was maintained (saruparib monotherapy as well as in combination with camizestrant and fulvestrant; Fig. 4h,i).

Overall, these preclinical models validate our clinicogenomic findings, demonstrating that acquired *RB1* loss underlies resistance to CDK4/6i in g*BRCA2* tumours with active HRD. In this context, biallelic *RB1* inactivation and Rb loss emerge under CDK4/6i selective pressure, whereas sensitivity to PARPi is maintained.

Discussion

We have described disease-specific and subtype-specific patterns of germline–somatic interactions and their clinical implications in a large clinicogenomic cohort of patients with breast cancer. Patients harbouring g*BRCA2* mutations or HRD arising from other causes experience a significantly worse outcome on combination CDK4/6i + ET, the current frontline standard of care in the metastatic $HR^+/HER2^-$ setting. This finding was validated using a large, independent, multi-institutional real-world dataset representative of the general population of patients with MBC treated in both community and academic settings.

We have demonstrated a proclivity for g*BRCA2* tumours to develop *RB1* biallelic loss as a mechanism of resistance to CDK4/6i-based combination therapy and propose two separate characteristics of g*BRCA2* tumours that facilitate the acquisition of *RB1* loss. First, g*BRCA2*-driven breast cancers are more likely to harbour *RB1* het loss before the start of CDK4/6i; we showed that this allelic configuration independently predisposes the tumour to loss of the unaltered *RB1* allele. Second, the mutagenic processes active in HRD-driven tumours promote the acquisition of *RB1* LoF variants as a mechanism of biallelic inactivation (a second hit). Consistent with these clinicogenomic findings, our preclinical models provide direct evidence mechanistically linking HRD and g*BRCA2* to the acquisition of *RB1* inactivation under the therapeutic pressure of CDK4/6i. In multiple g*BRCA2* PDXs, acquired *RB1* loss emerged as the dominant mode of CDK4/6i resistance in the context of active HRD. Conversely, sensitivity to PARPi was preserved, underscoring persistent vulnerability despite the emergence of CDK4/6i resistance.

These observations raise provocative questions regarding the optimal therapy sequence for g*BRCA2*-driven breast cancers. PARPi represent a proven targeted strategy for patients with HRD tumours[20,56]. However, given the long-standing use of CDK4/6i combinations as frontline standard of care for HR^+ breast cancers, use of PARPi is typically relegated to later lines of therapy following the emergence of resistance to CDK4/6i + ET. Our data reveal a potentially greater benefit for PARPi therapy before CDK4/6i + ET in this molecularly defined subset of patients with breast cancer. The dominant mechanism of resistance to CDK4/6i + ET may reduce the utility of several therapies typically used later in the disease course that rely on an intact G1 checkpoint. As the development of *RB1* loss appears to be partially mediated by HRD, earlier use of a PARPi may suppress this specific route of drug resistance resulting in a longer duration of response to subsequent CDK4/6i combinations. Specifically, reversion mutations that restore homologous recombination proficiency represent a widely prevalent and established mechanism of resistance to PARPi in g*BRCA2* carriers[57]. Therefore, tumour cells with somatic reversion mutations in *BRCA2* may be less likely to develop HRD-driven, *RB1* LoF mutations under the selective pressure of subsequent CDK4/6i.

The insights gleaned from our analysis should prompt a re-evaluation of conventional clinical practice in which the sequence of therapies is dictated by the clinical benefit observed in a biomarker unselected population. Most relevant to our current study, our results question whether frontline PARPi therapy in patients with pathogenic g*BRCA2* variants or HRD $HR^+/HER2^-$ tumours would not only improve frontline outcomes but also establish the genomic prerequisites for more successful subsequent lines of therapy. Our data and this hypothesis serve as the impetus for the EvoPAR-Breast01 study (NCT06380751), a randomized international phase III clinical trial, which will compare frontline saruparib (a PARP1-selective inhibitor) plus camizestrant (an oral selective ER degrader) versus standard-of-care CDK4/6i + ET in patients with $HR^+/HER2^-$ MBC harbouring germline or somatic alterations in selected HRD genes[58].

More generally, we have demonstrated that pre-treatment *RB1* het loss (hemizygosity) predicts the development of a somatic LoF variant involving the WT allele under the selective pressure of CDK4/6i. This proposed mechanism is the obverse of Knudson's two-hit model, which seminally demonstrated that copy number loss occurs as a necessary second hit for oncogenesis in patients with hereditary retinoblastoma[9]. By contrast, in our proposed model, pre-existing *RB1* het loss serves as the first hit increasing the likelihood of developing biallelic *RB1* loss via acquired *RB1* LoF variants as the second hit. Further studies will be needed to better refine the genomic and transcriptomic contexts predictive of CDK4/6i resistance resulting from biallelic *RB1* inactivation.

Our study has potential limitations. First, we focused our analysis on exonic somatic variants that have previously demonstrated functional significance, which may fail to capture potentially consequential interactions, such as intronic variants[59] or epigenetic silencing[60] of *RB1*. Second, given the real-world nature of this retrospective analysis, the timing between a particular event (start of treatment or disease progression) and a particular biopsy varied. In addition, our paired pre-CDK4/6i and post-CDK4/6i analysis may have failed to capture clonal expansion events that were extinguished by further therapy. Despite these caveats, we were able to conclusively demonstrate that pre-treatment *RB1* allelic configuration and HRD predict for the development of a second hit resulting in complete *RB1* loss and our preclinical models provide strong evidence for biological plausibility of the proposed mechanism.

From a translational perspective, this work contributes to ongoing efforts to hasten the identification and interception of acquired resistance mechanisms. The ability to predict the precise mechanism of acquired resistance from genomic characteristics detectable in the pre-treatment sample would be a valuable tool in the precision oncology arsenal. Although our work principally illustrates this concept through *RB1* hemizygosity, we hypothesize that the same framework may extend to other acquired tumour suppressors implicated in resistance. Currently defined 'targetable' genomic lesions encompass mutations, copy number homozygous deletions and amplifications, and structural variants, which in their present genomic configuration have already disrupted a specific oncogenic pathway. With this work, we expand the paradigm of the 'actionable genome' to states of allelic imbalance that have no biologic activity in their current allelic configuration but are predictive of therapeutic failure, and thus could be used to guide early interception strategies designed to delay the onset of resistance.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-026-10197-0.

1. Holohan, C., Van Schaeybroeck, S., Longley, D. B. & Johnston, P. G. Cancer drug resistance: an evolving paradigm. *Nat. Rev. Cancer* **13**, 714–726 (2013).
2. Vasan, N., Baselga, J. & Hyman, D. M. A view on drug resistance in cancer. *Nature* https://doi.org/10.1038/s41586-019-1730-1 (2019).
3. Will, M., Liang, J., Metcalfe, C. & Chandarlapaty, S. Therapeutic resistance to anti-oestrogen therapy in breast cancer. *Nat. Rev. Cancer* **23**, 673–685 (2023).
4. Razavi, P. et al. The genomic landscape of endocrine-resistant advanced breast cancers. *Cancer Cell* **34**, 427–438.e6 (2018).
5. Garraway, L. A. & Jänne, P. A. Circumventing cancer drug resistance in the era of personalized medicine. *Cancer Discov.* **2**, 214–226 (2012).
6. Li, F. P. & Fraumeni, J. F. Soft-tissue sarcomas, breast cancer, and other neoplasms. A familial syndrome? *Ann. Intern. Med.* **71**, 747–752 (1969).
7. Malkin, D. et al. Germ line p53 mutations in a familial syndrome of breast cancer, sarcomas, and other neoplasms. *Science* **250**, 1233–1238 (1990).
8. Alexandrov, L. B. et al. Signatures of mutational processes in human cancer. *Nature* https://doi.org/10.1038/nature12477 (2013).
9. Knudson, A. G. Mutation and cancer: statistical study of retinoblastoma. *Proc. Natl Acad. Sci. USA* **68**, 820–823 (1971).
10. Miki, Y. et al. A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. *Science* **266**, 66–71 (1994).
11. Wooster, R. et al. Identification of the breast cancer susceptibility gene BRCA2. *Nature* **378**, 789–792 (1995).
12. Rahman, N. et al. PALB2, which encodes a BRCA2-interacting protein, is a breast cancer susceptibility gene. *Nat. Genet.* **39**, 165–167 (2007).
13. Swift, M., Reitnauer, P. J., Morrell, D. & Chase, C. L. Breast and other cancers in families with ataxia-telangiectasia. *N. Engl. J. Med.* **316**, 1289–1294 (1987).
14. Cybulski, C. et al. CHEK2 is a multiorgan cancer susceptibility gene. *Am. J. Hum. Genet.* **75**, 1131–1135 (2004).
15. Moynahan, M. E., Pierce, A. J. & Jasin, M. BRCA2 is required for homology-directed repair of chromosomal breaks. *Mol. Cell* **7**, 263–272 (2001).
16. Moynahan, M. E., Chiu, J. W., Koller, B. H. & Jasint, M. Brca1 controls homology-directed DNA repair. *Mol. Cell* **4**, 511–518 (1999).
17. Lord, C. J. & Ashworth, A. BRCAness revisited. *Nat. Rev. Cancer* **16**, 110–120 (2016).
18. Farmer, H. et al. Targeting the DNA repair defect in BRCA mutant cells as a therapeutic strategy. *Nature* **434**, 917–921 (2005).
19. Pommier, Y., O'Connor, M. J. & De Bono, J. Laying a trap to kill cancer cells: PARP inhibitors and their mechanisms of action. *Sci. Transl. Med.* **8**, 362ps17 (2016).
20. Robson, M. et al. Olaparib for metastatic breast cancer in patients with a germline BRCA mutation. *N. Engl. J. Med.* https://doi.org/10.1056/NEJMoa1706450 (2017).
21. Mandelker, D. et al. Mutation detection in patients with advanced cancer by universal sequencing of cancer-related genes in tumor and normal DNA vs guideline-based germline testing. *JAMA* **318**, 825–835 (2017).
22. Cheng, D. T. et al. Memorial Sloan Kettering-Integrated Mutation Profiling of Actionable Cancer Targets (MSK-IMPACT). *J. Mol. Diagn.* https://doi.org/10.1016/j.jmoldx.2014.12.006 (2015).
23. Zehir, A. et al. Mutational landscape of metastatic cancer revealed from prospective clinical sequencing of 10,000 patients. *Nat. Med.* **23**, 703–713 (2017).
24. Krammer, J. et al. Breast cancer detection and tumor characteristics in BRCA1 and BRCA2 mutation carriers. *Breast Cancer Res. Treat.* **163**, 565–571 (2017).
25. Goodwin, P. J. et al. Breast cancer prognosis in BRCA1 and BRCA2 mutation carriers: an international prospective breast cancer family registry population-based cohort study. *J. Clin. Oncol.* **30**, 19–26 (2012).
26. Brekelmans, C. T. M. et al. Tumour characteristics, survival and prognostic factors of hereditary breast cancer from BRCA2-, BRCA1- and non-BRCA1/2 families as compared to sporadic breast cancer cases. *Eur. J. Cancer* **43**, 867–876 (2007).
27. Maxwell, K. N. et al. BRCA locus-specific loss of heterozygosity in germline BRCA1 and BRCA2 carriers. *Nat. Commun.* **8**, 319 (2017).
28. Riaz, N. et al. Pan-cancer analysis of bi-allelic alterations in homologous recombination DNA repair genes. *Nat. Commun.* **8**, 857 (2017).
29. Li, A. et al. Homologous recombination DNA repair defects in PALB2-associated breast cancers. *npj Breast Cancer* **5**, 23 (2019).
30. Weigelt, B. et al. The landscape of somatic genetic alterations in breast cancers from atm germline mutation carriers. *J. Natl Cancer Inst.* **110**, 1030–1034 (2018).
31. Holstege, H. et al. High incidence of protein-truncating TP53 mutations in BRCA1-related breast cancer. *Cancer Res.* **69**, 3625–3633 (2009).
32. Na, B. et al. Therapeutic targeting of BRCA1 and TP53 mutant breast cancer through mutant p53 reactivation. *npj Breast Cancer* **5**, 14 (2019).
33. Brugarolas, J. & Jacks, T. Double indemnity: p53, BRCA and cancer. p53 mutation partially rescues developmental arrest in Brca1 and Brca2 null mice, suggesting a role for familial breast cancer genes in DNA damage repair. *Nat. Med.* **3**, 721–722 (1997).
34. Kim, S. et al. Sequential activation of E2F via Rb degradation and c-Myc drives resistance to CDK4/6 inhibitors in breast cancer. *Cell Rep.* **42**, 113198 (2023).
35. Wander, S. A. et al. The genomic landscape of intrinsic and acquired resistance to cyclin-dependent kinase 4/6 inhibitors in patients with hormone receptor-positive metastatic breast cancer. *Cancer Discov.* **10**, 1174–1193 (2020).
36. Hortobagyi, G. N. et al. Ribociclib as first-line therapy for HR-positive, advanced breast cancer. *N. Engl. J. Med.* **375**, 1738–1748 (2016).
37. Cristofanilli, M. et al. Fulvestrant plus palbociclib versus fulvestrant plus placebo for treatment of hormone-receptor-positive, HER2-negative metastatic breast cancer that progressed on previous endocrine therapy (PALOMA-3): final analysis of the multicentre, double-blind, phase 3 randomised controlled trial. *Lancet Oncol.* **17**, 425–439 (2016).
38. Goetz, M. P. et al. MONARCH 3: abemaciclib as initial therapy for advanced breast cancer. *J. Clin. Oncol.* **35**, 3638–3646 (2017).
39. Li, Z. et al. Loss of the FAT1 tumor suppressor promotes resistance to CDK4/6 inhibitors via the Hippo pathway. *Cancer Cell* **34**, 893–905.e8 (2018).
40. Ma, X. et al. Comparison of population characteristics in real-world clinical oncology databases in the US: Flatiron Health, SEER, and NPCR. Preprint at *medRXiv* https://doi.org/10.1101/2020.03.16.20037143 (2023).
41. Database characterization guide. *Flatiron Health* https://flatiron.com/database-characterization (accessed 27 December 2025)
42. Kainu, T. et al. Somatic deletions in hereditary breast cancers implicate 13q21 as a putative novel breast cancer susceptibility locus. *Proc. Natl Acad. Sci. USA* **97**, 9603–9608 (2000).
43. Rouault, A. et al. Deletion of chromosomes 13q and 14q is a common feature of tumors with BRCA2 mutations. *PLoS ONE* **7**, e52079 (2012).
44. Tsukamoto, K. et al. Two distinct commonly deleted regions on chromosome 13q suggest involvement of BRCA2 and retinoblastoma genes in sporadic breast carcinomas. *Cancer* **78**, 1929–1934 (1996).
45. Annala, M. et al. Treatment outcomes and tumor loss of heterozygosity in germline DNA repair-deficient prostate cancer. *Eur. Urol.* **72**, 34–42 (2017).
46. Mandigo, A. C. & Knudsen, K. E. Double trouble: concomitant RB1 and BRCA2 depletion evokes aggressive phenotypes. *Clin. Cancer Res.* **26**, 1784–1786 (2020).
47. Chakraborty, G. et al. Significance of BRCA2 and RB1 co-loss in aggressive prostate cancer progression. *Clin. Cancer Res.* **26**, 2047–2064 (2020).
48. Turner, N. C. et al. Overall survival with palbociclib and fulvestrant in advanced breast cancer. *N. Engl. J. Med.* **379**, 1926–1936 (2018).
49. Bielski, C. M. et al. Genome doubling shapes the evolution and prognosis of advanced cancers. *Nat. Genet.* **50**, 1189–1195 (2018).
50. André, F. et al. Pooled ctDNA analysis of MONALEESA phase III advanced breast cancer trials. *Ann. Oncol.* **34**, 1003–1014 (2023).
51. Soria-Bretones, I. et al. The spindle assembly checkpoint is a therapeutic vulnerability of CDK4/6 inhibitor-resistant ER^+ breast cancer with mitotic aberrations. *Sci. Adv.* **8**, eabq4293 (2022).
52. Gong, X. et al. Aurora A kinase inhibition is synthetic lethal with loss of the RB1 tumor suppressor gene. *Cancer Discov.* **9**, 248–263 (2019).
53. Ahrabi, S. et al. A role for human homologous recombination factors in suppressing microhomology-mediated end joining. *Nucleic Acids Res.* **44**, 5743–5757 (2016).
54. Gould, S. J. & Vrba, E. S. Exaptation — a missing term in the science of form. *Paleobiology* https://doi.org/10.1017/S0094837300004310 (1982).
55. Nguyen, L., Martens, J. W. M., Van Hoeck, A. & Cuppen, E. Pan-cancer landscape of homologous recombination deficiency. *Nat. Commun.* **11**, 5584 (2020).
56. 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).
57. 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).
58. Razavi, P. et al. Abstract P2-10-17: a randomized phase III study of first-line saruparib (AZD5305) + camizestrant vs CDK4/6i plus physician's choice endocrine therapy or + camizestrant in patients w/ BRCA1/BRCA2/PALB2 mutations & HR^+/$HER2^-$ advanced breast cancer (EvoPAR-Breast01). *Clin. Cancer Res.* **31**, P2-10-17 (2025).
59. Febres-Aldana, C. et al. Rb tumor suppressor in small cell lung cancer: combined genomic and IHC analysis with a description of a distinct Rb-proficient subset. *Clin. Cancer Res.* **28**, 4702–4713 (2022).
60. Youn, J. I. et al. Epigenetic silencing of retinoblastoma gene regulates pathologic differentiation of myeloid cells in cancer. *Nat. Immunol.* **14**, 211–220 (2013).

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-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, 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 licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence 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 licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

© The Author(s) 2026

[1]Department of Medicine, Memorial Sloan Kettering Cancer Center, New York, NY, USA. [2]Weill Cornell Medical College, New York, NY, USA. [3]Department of Surgery, Memorial Sloan Kettering Cancer Center, New York, NY, USA. [4]Human Oncology and Pathogenesis Program, Memorial Sloan Kettering Cancer Center, New York, NY, USA. [5]Department of Pathology and Laboratory Medicine, Memorial Sloan Kettering Cancer Center, New York, NY, USA. [6]Translational Oncology Partnership Program, Memorial Sloan Kettering Cancer Center, New York, NY, USA. [7]Marie-Josée and Henry R. Kravis Center for Molecular Oncology, Memorial Sloan Kettering Cancer Center, New York, NY, USA. [8]Royal Marsden Hospital and Institute of Cancer Research, London, UK. [9]Experimental Therapeutics Group, Vall d'Hebron Institute of Oncology, Barcelona, Spain. [10]AstraZeneca, Cambridge, UK. [11]Early Drug Development, European Institute of Oncology, IRCCS, Milan, Italy. [12]Department of Genetics, Institut Curie, Paris, France. [13]Division of Translational Medicine and Human Genetics, Department of Medicine, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, PA, USA. [14]Laboratory of Preclinical Investigation, Institut Curie, Paris, France. [15]Pfizer, New York, NY, USA. [16]Department of Epidemiology and Biostatistics, Memorial Sloan Kettering Cancer Center, New York, NY, USA. [17]AstraZeneca, Gaithersburg, MD, USA. [18]Clinical Genetics Service, Memorial Sloan Kettering Cancer Center, New York, NY, USA. [19]Abramson Cancer Center, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, PA, USA. [20]Division of Hematology-Oncology, Department of Medicine, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, PA, USA. [21]Basser Center for BRCA, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, PA, USA. [22]Experimental Pathology and Laboratory Medicine, Mayo Clinic, Rochester, MN, USA. [23]Computational Oncology Service, Memorial Sloan Kettering Cancer Center, New York, NY, USA. [24]These authors contributed equally: Anton Safonov, Minna Lee. ✉e-mail: chandars@mskcc.org; razavip@mskcc.org

Methods

Study cohort and prospective sequencing

The study cohort comprised 6,927 tumour samples from 5,881 patients with breast cancer. All patients underwent prospective clinical tumour and normal DNA sequencing as part of their clinical care (February 2014 to September 2021). The present study was approved by the MSK Institutional Review Board (IRB) and all patients provided written informed consent for tumour and paired normal DNA sequencing and review of medical records for clinical annotations. Genomic sequencing was performed on tumour DNA extracted from formalin-fixed, paraffin-embedded tissue and normal DNA extracted from mononuclear cells from peripheral blood in all patients as previously described[22]. Patient samples were sequenced in a CLIA-compliant laboratory using one of several versions of the MSK-Integrated Mutation Profiling of Actionable Cancer Targets (IMPACT) targeted sequencing panel, which interrogates exonic and selected intronic regions of 341, 410, 468 or 506 genes depending on the assay version, with somatic mutation calling performed using an extensively validated pipeline followed by manual review, as previously described[22,23]. Tumours were obtained from the primary site in 50.1% (n = 3,470) of samples and from a metastatic site in 49.8% (n = 3,457) of samples.

Anonymized germline variant calling was performed using a sequence analysis pipeline validated for clinical use in a CLIA-compliant laboratory performing clinical sequencing of patient tumours and matched normal blood specimens as a part of routine clinical care[21]. Germline variants with a population allele frequency of less than 2% in gnomAD exomes (v2.0.1) and genomes (v2.1.1) were assessed for pathogenicity following ACMG guidelines. Variants predicted by Variant Effect Predictor to have high-impact consequences were considered putative LoF variants. Variants were designated as pathogenic if classified as pathogenic or likely pathogenic in-house or by ClinVar or predicted to be LoF in a TSG. We excluded variants flagged as potentially arising from clonal haematopoiesis or circulating tumour cells as previously described[61].

We manually reviewed variants with discordant interpretations between ClinVar and in-house classifications, as well as variants with conflicting interpretations in ClinVar. Novel LoF variants were considered pathogenic unless located in the terminal exon, where such variants were considered pathogenic only if the same gene harboured previously designated pathogenic LoF variants located further downstream in the coding sequence. Low-risk variants in *CHEK2* such as c.470T > C were excluded from the analysis.

Population frequencies were obtained from gnomAD exomes (v2.0.1) and gnomAD genomes (v2.1.1). Variants were annotated for pathogenicity interpretations using ClinVar (accessed December 2024) and in-house classifications (as of August 2022) by expert clinical geneticists. Ensembl Variant Effect Predictor was used to annotate variants for predicted functional consequences.

The external cohort consisted of tumour samples from g*BRCA2* primary tumours, which underwent whole-exome sequencing (WES) from samples collected at the University of Pennsylvania and Mayo Clinic[5]. Patients gave written informed consent for research use of germline DNA and tumour specimens under IRB-approved protocol at the University of Pennsylvania and Mayo Clinic. As previously described[62], somatic tumour DNA was extracted from formalin-fixed, paraffin-embedded primary breast cancer specimens using standard laboratory deparaffinization, whereas germline DNA was extracted from whole blood or saliva. Tumour DNA libraries were prepared using the NEBNext formalin-fixed, paraffin-embedded repair mix and NEBNext Ultra II DNA library prep kit (New England Biolabs), per the manufacturer's instructions. Germline DNA libraries were prepared using the NEBNext Ultra DNA library prep kit (New England Biolabs), per the manufacturer's instructions. DNA libraries were pooled and hybridized using SureSelect Target Enrichment System for Illumina Multiplex Sequencing (Agilent) and associated protocols. For WES, tumour and germline libraries were hybridized to SureSelect All Exon v5, SureSelect All Exon v6+COSMIC and SureSelect All Exon v7 captures (Agilent). WES was performed using an Illumina HiSeq 4000, and targeted sequencing was performed using an Illumina NovaSeq 6000. All sequencing was performed with 150 paired-end reads by the University of Pennsylvania Next Generation Sequencing Core.

FASTQ files from WES and targeted sequencing were aligned to the hg19 build of the human genome using the Burrows–Wheeler Aligner (BWA v0.7.17-r1188)[63]. Various bam file processing operations were performed using Samtools/htslib/bcftools (v1.11). The resulting bam files were processed according to Genome Analysis Toolkit (GATK v3.7) best practices (picardtools v2.20.7).

Zygosity inference

We inferred somatic zygosity for all germline pathogenic variants using locus-specific copy number and ASCN inference utilizing FACETS as previously described[64]. All ASCN solutions (FACETS outputs) from tumours with germline pathogenic variants were manually reviewed to ensure that the optimal solution was selected. Purity estimates were similarly inferred from the FACETS. We incorporated ASCN, purity and variant allele frequency (VAF) into a previously described framework[65], allowing statistical inference of heterozygous, biallelic (loss of WT allele) or loss of mutant allele. The zygosity of the germline variant was considered indeterminate and excluded from zygosity analyses if the: (1) variant was homozygous in the germline; (2) read depth of coverage in the normal blood specimen was less than 50; (3) FACETS-derived total and minor copy numbers were not evaluable at the corresponding locus; or (4) an optimal solution could not be identified, most commonly due to low tumour purity. Using these criteria, out of the 472 total cases with germline pathogenic variant in the genes of interest, somatic zygosity status was evaluable for 447 (94.7%) of tumours. We excluded cases where copy number analysis and VAF were suggestive of loss of mutant allele, rather than WT (n = 22, 4.6%) from further analyses.

To initially determine whether a given germline variant was in allelic imbalance in the corresponding tumour specimen, we evaluated consistency between observed somatic VAF and expected VAF. The latter value was calculated as a function of ASCN and purity as previously defined[65]. Germline variants were considered heterozygous if their observed VAF was either (1) consistent with the expected VAF (within its 95% binomial CI) given balanced heterozygosity (total copy number (TCN) and lesser copy number (LCN) of either 2 and 1 or 4 and 2 in diploid and genome doubled tumours, respectively), or (2) less than the lower bound of the 95% CI of the expected VAF corresponding to a TCN and LCN of 3 and 1, respectively, which was either single copy gain of the mutant or WT allele. Germline variants in allelic imbalance of any kind were those with an observed VAF that was either within or greater than the 95% CI of the expected VAF corresponding to a copy number state other than balanced heterozygosity. For allelically imbalanced germline variants, loss of the WT was determined as those with an observed VAF within the 95% CI (or greater than the lower bound of the 95% CI) of the expected VAF corresponding to a LCN equal to 0 (observed VAF is concordant with the expected VAF when the lesser allele has a copy number of 0).

Data anonymization

All patients (n = 5,881) consented to an IRB-approved protocol allowing analysis of somatic and clinical data (NCT01775072). A subset of 2,896 patients had additionally consented for identified analysis of germline variants under this protocol, whereas the remaining patients (n = 2,985) consented only to somatic analyses.

For analyses involving germline pathogenic variant status, we anonymized data into a unique anonymized ID (A-ID-#####), enabling germline calls to be conducted on matched normal samples for all patients regardless of germline consent status in accordance

with MSK IRB guidelines for anonymized germline–somatic analyses, as previously described[65,66]. For all patients, data were binned to avoid unique clinical or somatic alteration values and thereby prevent re-identification[65,66]. In brief, for patients who did not provide consent for identified germline analysis, genomic data were anonymized with a deterministic one-way hash function. In these patients, germline variant calling was performed using the clinically validated pipeline described above[67], and PFS and OS times were rounded to the nearest month. For analyses in which germline data were not required, a separate study-specific unique identifier (S-ID-#####) and fully anonymized clinical data were used for all the patients regardless of germline consent status. Continuous clinical (for example, time to progression) and somatic genomic data were therefore able to be used for analyses that did not incorporate germline data.

Clinical data annotation

For the MSK cohort, clinical data were obtained from the validated Breast Translational Program's clinical database, which contains comprehensive, longitudinal information on patient demographics, pathology, treatment regimens and clinical outcomes[4,39,68–71]. Structured data were systematically curated using standardized clinical data annotation processes by a dedicated team of expert clinical data annotators and were maintained under rigorous quality-control procedures by experienced data managers. The data lock for the analysis was April 2025.

Each patient in the cohort was assigned a single receptor status. Recognizing potential intertumoural heterogeneity, we sought a unified definition as follows: (1) in cases in which any metastatic biopsy was sequenced, receptor status was defined by treating clinician interpretation at the time of assigned first-line treatment; (2) in cases in which only a primary tumour is sequenced, receptor status was defined by receptor status of the sequenced primary. We excluded cases in which sequencing was obtained for external consultation (and therefore, lacked clinical or pathological data, $n = 310$), cases where the diagnosis was ductal carcinoma in situ (DCIS), with no evidence of invasive breast cancer during the clinical course ($n = 1$) or cases in which multi-site sequencing demonstrated no intrapatient genomic overlap and multiple distinct receptor statuses consistent with multiple primary tumours ($n = 4$; Extended Data Fig. 1a). Using these definitions, the MSK cohort consisted of 3,703 patients (66.5%) with $HR^+/HER2^-$ receptor status and 1,043 patients (18.7%) with triple-negative breast cancer, with the remainder (820, 14.7%) being HER2 amplified ($HER2^+$).

Progression events were defined as (1) a radiographic or clinical disease progression prompting change in systemic therapy or recommendation for ablative local therapy directed at a site (or sites) of progressive disease; or (2) clinician assessment detailing radiographic and/or clinical progression, after which it was documented that the patient and physician decided to continue the same therapy post-progression. In such cases, the time of progression was defined as the date of documented progression rather than the date of therapy discontinuation.

Enrichment analysis

We compiled mutations, fusions and copy number alterations predicted to be functionally significant (oncogenic or likely oncogenic) by the OncoKB precision oncology variant database[72]. All putative *RB1* homozygous deletions were manually re-reviewed. Cases in which the putative homozygous deletion spanned beyond the size of an amplicon (that is, an event that would be interpreted as incompatible with tumour survival) were not considered a functionally significant variant.

In cases in which a patient had multiple samples sequenced, we compiled the total somatic variants called from either the sequenced primary sample or all sequenced metastatic samples (omitting the primary), to avoid duplicate samples and to ensure that each set of variants was assigned either 'primary' or 'metastasis' as a covariate. For the purposes of this analysis, local recurrence samples were considered 'primary'. Upon excluding samples with indeterminable receptor status (as described in the 'clinical data annotation' section), 5,566 patients were eligible for Firth-penalized logistic regression.

Receptor status was also defined on a 'per-patient' basis as described above. Genes with alterations in less than 2% of the cohort were omitted from the analysis. For each remaining combination of germline gene X and somatic gene Y, we performed a Firth-penalized logistic regression to account for the sparsity of the dataset. Receptor status and sample type (metastatic versus primary) were included as covariates. The analysis was also repeated for each receptor status subtype, as well as repeated for biallelic versus mono-allelic germline variants. The latter step was only performed for samples in which zygosity was evaluable ($n = 5,516$). Putative P values were adjusted for multiple hypothesis testing using the Benjamini–Hochberg method; $q < 0.10$ was deemed to be statistically significant.

ASCN definitions

In the MSK cohort, somatic LOH events of *RB1* were defined based on manual FACETS review of pre-CDK4/6i treatment samples. Pre-CDK4/6i ASCN analysis was performed on samples collected before first-line CDK4/6i treatment or as part of a matched pre-treatment and post-treatment CDK4/6i pair. Of 922 patients meeting this criteria, 196 samples were excluded due to low purity or other technical limitations (such as 'waterfalling' artefact), or paucity of heterozygous single-nucleotide polymorphisms allowing for confident lesser copy number inference. *RB1* LOH was defined as LCN = 0, irrespective of TCN, whereas heterozygous state was any LCN > 0.

In the PALOMA-3 cohort, baseline ctDNA samples from the palbociclib combination arm were sequenced using a 1,729 amplicon custom AmpliSeq panel, which included 119 single-nucleotide polymorphisms located within the *RB1* gene. ctDNA-based LOH analysis was conducted using a bespoke pipeline[48]; *RB1* LOH was defined as previously described[73].

In the MSK cohort, we further separated the *RB1* LOH (LCN = 0) group into (1) het loss, defined as a state with TCN = 1 and LCN = 0; and (2) other LOH, defined as a state with TCN > 1 and LCN = 0. Fraction genome altered was also calculated for each pre-treatment sample and defined as the fraction of $\log_2$ copy number variation (gain or loss) of more than 0.2 divided by the total size of the copy number profiled region.

HRD inference analysis

To study the implications of HRD on clinical outcomes and mechanism of resistance, we inferred HRD from two orthogonal methods, which have been validated for use with targeted next-generation sequencing data. IMPACT-HRD quantifies genomic scars associated with HRD by analysing ASCN alterations determined with the FACETS algorithm (v0.5.14) and computing those genomic scars with the impact-hrd package (https://github.com/mskcc/facets-suite/blob/master/R/copy-number-scores.R). All IMPACT-HRD assessments were completed using R (v4.1.2). In particular, three metrics were evaluated: number of telomeric allelic imbalances, large-scale transitions and losses of heterozygosity. The overall HRD phenotype is defined as the unweighted sum of these three metrics.

Clinical outcome analysis

We determined the association between genomic alterations and PFS with disease progression on therapy with CDK4/6i or patient death. Disease progression was defined as the date of the radiological study or clinical assessment that established progression of disease and prompted a change in systemic treatment, intervention with locally directed therapy (for example, radiation therapy), or otherwise an annotation in the chart documenting progression of disease. We categorized CDK4/6i regimens based on their ET partner (aromatase inhibitor versus selective oestrogen receptor degrader). Patients with ablation of only known sites of disease with radiotherapy or surgical resection

before initiation of CDK4/6i therapy were excluded, as were patients who discontinued therapy due to toxicity within 2 weeks. When ET or CDK4/6i changed to another ET and CDK4/6i, respectively, for reasons other than disease progression (for example, toxicity, patient preference or insurance coverage), the time on successive regimens was combined to more accurately capture real-world PFS on the CDK4/6i + ET combination.

We used both univariate and multivariate Cox proportional hazard models (adjusted for ET partner (fulvestrant versus aromatase inhibitors) and treatment line, where applicable). For patients with multiple lines of therapy from the same class of treatment, only the first treatment line from that class that was started after the MSK-IMPACT biopsy was included in the analysis. For analyses pertaining to ASCN, fraction genome altered and whole-genome duplication were used as additional covariates. These are recognized poor prognostic factors and may be a confounding factor given increased likelihood for tumours with measures of copy number instability to harbour LOH of any specific region.

For OS analysis, we implemented a left-truncated model to account for the immortal time from diagnosis of metastatic disease (time zero) to enrolment on sequencing protocol. Similarly to the univariate analyses, we used univariate and multivariate Cox proportional hazard models. In addition to ET partner, age at metastatic diagnosis was also included as a covariate. We rejected the null hypotheses with a two-sided $\alpha = 0.05$.

For the matched-pairs analysis, we included all patients with available paired pre-CDK4/6i and post-CDK4/6i sequencing data. Pre-CDK4/6i samples consisted exclusively of tumour specimens sequenced using MSK-IMPACT with available ASCN and zygosity assessment for *RB1*. Post-treatment samples included post-progression CDK4/6i tumour specimens sequenced using MSK-IMPACT as well as ctDNA sequenced using either MSK-ACCESS[74] or Guardant360 (ref. 75). For analyses comparing *RB1* heterozygous loss with other allelic configurations, we focused specifically on drivers of resistance rather than subclonal events. We therefore excluded post-treatment alterations in which the VAF was less than 0.30 of the maximum allele frequency of high-confidence variants present in the particular sample of interest.

For assessment of acquired tumour suppressor loss in the HRD versus *BRCA2* versus non-HRD group, we considered tumour suppressor genes that have been implicated in CDK4/6i resistance: *RB1* (refs. 35,39,76), *PTEN*[77], *LATS2* (ref. 39), *FAT1* (ref. 39), *TP53* (ref. 71), *ARID1A*[78], *LATS1* (ref. 39) and *NF1* (ref. 79). We excluded samples in which the ASCN was not evaluable in all these genes of interest, or in which there was already a biallelic LoF of one of the genes predicted to confer immediate resistance (*RB1*, *PTEN* and *NF1*). We included patients with baseline *TP53* loss, as it has been shown to facilitate cell cycle re-entry and is therefore associated with acquired resistance on an intermediate timescale[71]. We defined pre-treatment samples as either (1) gBRCA2, (2) non-*BRCA2* HRD (either harbouring a germline variant in *BRCA1*, *PALB2* or classified as HRD-positive by the HRD-IMPACT assay) versus (3) non-HRD. All clinical outcome analyses were conducted with R software (v4.5.1) and the survival and exact2x2 packages.

In vivo PDX models

Targeted sequencing of post-mortem and PDX studies. Post-mortem tissue samples were selected for DNA extraction, library preparation and targeted sequencing. Up to 30 mg frozen tissue was digested with 40 µl of proteinase K (600 mAU ml^{-1}) in 360 µl Buffer ATL at 56 °C. Genomic DNA was isolated using the DNAeasy Blood & Tissue Kit (69504, Qiagen) according to the manufacturer's protocol, including treatment with RNase A. DNA was eluted in 60 µl 0.5X Buffer AE heated to 55 °C.

After PicoGreen quantification and quality control using an Agilent BioAnalyzer, 100 ng of genomic DNA was used to prepare libraries using the KAPA Hyper Prep Kit (07961901001, Roche) with eight cycles of PCR. Of each barcoded library, 100–135 ng was captured by hybridization in a pool of 9 samples using the IMPACT assay (IDT), designed to capture all protein-coding exons and select introns of 505 commonly implicated oncogenes, tumour suppressor genes and members of pathways deemed actionable by targeted therapies. Captured pools were sequenced using an Illumina NovaSeq 6000 in PE100 run mode using the NovaSeq 6000 S4 Reagent Kit (200 cycles). All experiments were carried out at MSK's Integrated Genomics Organization.

The demultiplexed FASTQ files from the post-mortem samples were aligned to the human genome reference GRCh37/hg19 using bwa mem (v0.7.17-r1188)[63] and deduplicated using Picard MarkDuplicates (v2.21.8). Quality-control metrics of the alignments included (1) unique passing filter (PF)-aligned read pairs, (2) mean target coverage, (3) mean insert size, and (4) major or minor contamination.

Variant calling was performed using a previously described pipeline[80]. In brief, structural nucleotide variants were detected in the tumour–normal pairs using Mutect (v1.1.6)[81], whereas indels were detected using a consensus of Varscan 2 (v2.4.6)[82], Strelka (v2.9.10)[83], Scalpel (v0.5.4)[84] and Platypus (v0.8.1.2)[85]. Variants found with more than 0% global allele frequency in the 1000 Genomes database (phase III) or more than 0.01% across any population in the ExAC database (release 0.3.1) or that were covered by 10 reads in the tumour or 5 reads in the germline were filtered out. Variants for which the tumour variant allele fraction was more than five times than that of the normal variant allele fraction were filtered out. The aggregated set of variants identified in the tissues and xenografts were re-genotyped in all samples using SAMtools mpileup (v1.19.2)[86]. Copy number alterations were detected using Facets (v0.6.2)[64]. In addition, off-target reads were used to estimate $\log_2$ ratios using CNVkit (v0.9.8)[87]. Structural variants were detected using the consensus of Manta (v1.6.0)[88], SvABA (v1.1.0)[89] and GRIDSS (v2.13.2)[90]. The aggregated set of structural variants identified were re-genotyped in all samples using Paragraph (v2.3)[91] and annotation of the structural variants was done using vcf2maf (v1.6.22; https://github.com/mskcc/vcf2maf) and AnnotSV (v3.5.3)[92]. Reversion mutations and structural variants affecting *BRCA2* were further classified using aardvark (v0.35)[93,94].

MSK PDXs (PDX-L and PDX-R)

Animal studies. Mouse studies were conducted through the MSK antitumour core facility in compliance with institutional guidelines under an Institutional Animal Care and Use Committee-approved protocol (MSK IRB 12-10-016). PDXs were established by implanting freshly collected autopsy samples from a patient in MSK's Last Wish Program (MSK IRB 15-021), which enables patients to donate their bodies post-mortem for research. The samples were collected at MSK under approved IRB biospecimen protocols (MSK IRB 12-245 and 06-107).

Animals were maintained in accordance with the Guide for the Care and Use of Laboratory Animals in an AAALAC-accredited facility. All procedures outlined in the study were approved by the MSK Cancer Center Institutional Animal Care and Use Committee. Animals were housed in individually ventilated caging systems (Thoren Caging Systems), on autoclaved aspen chip bedding (PJ Murphy Forest Products) and were provided a γ-irradiated commercial diet (PicoLab Rodent Diet 20, 5053 LabDiet, PMI Nutrition International), and acidified water (pH 2.5–2.8) ad libitum. Mice were housed at a population density that ranged from 1 to 5 mice per cage in an environment providing a temperature of 21.1–22.2 °C (70–72 °F), 30–70% humidity, 10–15 fresh air exchanges hourly and a 12–12-h light–dark cycle (lights on, 06:00–18:00). For PDX studies, the tumours were expanded by serial subcutaneous transplantation.

In vivo studies. The 0.18 mg/90-day-release oestrogen pellets were implanted into 6-week-old female NSG mice 5 days before tumour implantation. When xenografts reached 100 mm^3, mice were randomized to treatment arms of vehicle (saline), fulvestrant (200 mg kg^{-1} subcutaneous, twice weekly), ribociclib (75 mg kg^{-1} PO, 5 days per week) or combination therapy. Tumour size was measured twice a week. The

animals were euthanized at the end of the experiment and tumours were collected for histological and biochemical analyses. The maximum allowed tumour size was 2,000 mm^3. The sample size for PDX experiments was calculated based on previous experience with this model and drug response. No statistical method was used to predetermine sample size. The investigators were not blinded to allocation during experiments and outcome assessment.

For analysis, a linear mixed-effect model was used for comparing the growth curves between the treatment conditions. In detail, the model included the tumour volume as the dependent variable, individual mouse ID as a random intercept, and day, treatment and the interaction term between day and treatment as fixed effects. In PDX-R, the model comparing ribociclib to vehicle failed to converge. This was attributable to one extreme outlier mouse in each group at multiple timepoints identified using the Tukey method (tumour volume exceeding quantile 3 + 1.5 × interquartile range). Once excluding these outliers, the mixed effect model successfully converged and satisfied the model assumptions. A sensitivity analysis confirmed that the exclusion of these two mice did not alter the biological conclusion regarding the significance. Complete tumour raw data are included in the supplement.

Immunoblotting. Frozen PDX tumours were thawed on ice, cut into small pieces and placed in Lysing Matrix tubes (6910100, MP Biomedicals). SDS lysis buffer was added, and the sample was homogenized for 40 s and then boiled at 100 °C for 5 min. The supernatant was transferred and subjected to sonication at 40–45 amp for 30 s, repeated twice. After sonication, lysates were boiled again and then centrifuged. Protein concentration was quantified with BCA protein assay (23225, Thermo Scientific). Of protein, 25 µg was loaded for 3–8% Tris-acetate gel (NuPage) electrophoresis and transferred to nitrocellulose membranes. Blots were blocked with 5% non-fat milk in TBST for 1 h at room temperature and then incubated with primary antibody at 4 °C overnight. The following primary antibodies were used at 1:1,000 dilution: BRCA2 (123491, Abcam), pRB S780 (8180S, Cell Signaling Technology) and Rb (9313S, Cell Signaling Technology). Secondary antibodies conjugated with fluorescence (#926-32211 and #926-68070, LI-COR Bioscience) were incubated for 1 h at room temperature. Blots were imaged by Odyssey Clx Imaging System (LI-COR Biosciences); raw images are included in Supplementary Fig. 1.

Whole-genome sequencing of PDX samples. Post-CDK4/6i PDX samples were selected for DNA extraction, library preparation and whole-genome sequencing. The tissue samples were homogenized in 500 µl MagMAX DNA Cell and Tissue Extraction Buffer (A45469, Thermo Fisher) for up to 40 s and DNA from lysate was extracted using the MagMAX DNA Multi-Sample Ultra 2.0 Kit (A36570, Thermo Fisher) on the KingFisher Apex System (Thermo Fisher) according to the manufacturer's protocol. The samples were eluted in 80 µl elution solution.

After PicoGreen quantification and quality control using an Agilent TapeStation, 500 ng of genomic DNA were sheared using a LE220-plus Focused-ultrasonicator (500569, Covaris) and sequencing libraries were prepared using the KAPA EvoPrep Kit (10212250702, Roche) with modifications. The libraries were subjected to a 0.5× size selection using aMPure XP beads (A63882, Beckman Coulter) after post-ligation cleanup. The libraries were not amplified by PCR and were pooled at equal volume. The samples were sequenced using an Illumina NovaSeq X in PE150 run mode using the NovaSeq X 25B Reagent Kit. All experiments were carried out at MSK's Integrated Genomics Organization.

The demultiplexed FASTQ files were aligned to a chimeric genome reference comprising the human reference GRCh37/hg19 and the mouse GRCm38/mm10 using bwa mem (v0.7.17-r1188)[63]. Read pairs where at least one end (R1 and/or R2) had a primary alignment the mouse genome were filtered out and the remaining read pairs were re-aligned to the human reference GRCh37/hg19 as described above. Quality control of the alignments was done as previously described above in which the percent of mouse content was quantified as the number of unique PF read pairs aligned to the mouse genome relative to the total unique PF aligned read pairs. Copy number alterations were detected using Facets (v0.6.2)[64]. Structural variants were detected using the consensus of Manta (v1.6.0)[88], SvABA (v1.1.0)[89] and GRIDSS (v2.13.2)[90]. The aggregated set of structural variants identified in the xenografts were re-genotyped in all samples using Paragraph (v2.3)[91], and annotation of the structural variants was done using vcf2maf (v1.6.22; https://github.com/mskcc/vcf2maf) and AnnotSV (v3.5.3)[92].

Institut Curie PDXs. PDX-C (HBCx-118) models of ER^+ MBC were obtained by engrafting biopsies from spinal bone metastases of patients with ER^+ breast cancer progressing under ET. Specifically, this model was derived from a patient with ER^+ MBC previously exposed to combination fluorouracil, epirubicin and cyclophosphamide (FEC) chemotherapy and tamoxifen in the adjuvant setting, and aromatase inhibitor and paclitaxel in the metastatic setting. The protocol was approved by the Institut Curie Hospital committee (Comité de Revue Institutionnel). Bone metastasis biopsies were engrafted with informed consent from the patient into the interscapular fat pad of female Swiss nude mice (Charles River Laboratories), which were maintained under specific pathogen-free conditions. Their care and housing were in accordance with institutional guidelines and the rules of the French Ethics Committee: CEEA-IC (Comité d'Ethique en matière d'expérimentation animale de l'Institut Curie, National registration number: #118). The project authorisation no. is 02163.02. The housing facility was kept at 22 °C (± 2 °C) with a relative humidity of 30–70%. The light–dark cycle was 12 h light–12 h dark.

For efficacy studies, tumour fragments were transplanted into female 8-week-old Swiss nude mice. When tumours reached a volume comprised between 100 and 200 mm^3, xenografts were randomly assigned to the different treatment groups of vehicle and palbociclib 75 mg kg^{-1} PO 5 days per week. Tumour size was measured with a manual caliper twice per week. Tumour volumes were calculated as $V = a \times b2/2$, a being the largest diameter, by the smallest. Tumour volumes were then reported to the initial volume as relative tumour volume. Means (and s.d.) of relative tumour volume in the same treatment group were calculated, and growth curves were established as a function of time. For each tumour, the percent change in volume was calculated as ($Vf - V0$/V0)/100, $V0$ being the initial volume (at the beginning of treatment) and Vf the final volume (at the end of treatment).

PDX-P1 and PDX-P2. All animal work was conducted according to AstraZeneca's Global Bioethics Policy (https://www.astrazeneca.com/content/dam/az/Sustainability/Bioethics_Policy.pdf), in accordance with the PREPARE guidelines and reported in line with the ARRIVE guidelines.

Studies with PDX-P1 (ST4316B) were performed under contract with XenoStart at AAALAC-accredited facilities and performed in accordance with protocols approved by the START 'Institutional Animal Care and Use Committee' and AstraZeneca's 'Platform for Animal Research Tracking and External Relationships' (PARTNER) group. Mice were acclimated for a minimum of 24 h and housed on irradiated corncob bedding (Teklad) in individual HEPA-ventilated cages (Sealsafe Plus, Techniplast USA) on a 12-h light–dark cycle at 21–23 °C and 40–60% humidity. Animals were fed water ad libitum (reverse osmosis, 2 ppm Cl_2) and an irradiated standard rodent diet (Teklad). Xenografts were established by subcutaneous surgical implantation of an approximately 70 mg tumour fragment into the right flanks of 6–12-week-old female athymic nude animals. Tumours reached 0.15–0.3 cm^3 before the animals were randomized into groups. Tumour volume (mm^3) was calculated as $width^2 \times length \times 0.52$.

Studies with PDX-P2 (HBCx-22) were performed under contract with Xentech under authorization by the 'Direction Départementale de la Protection des Populations, Ministère de l'Agriculture et de l'Alimentation', France and in accordance with protocols approved by

Xentech along with AstraZeneca's PARTNER group. Mice were delivered to the facility at least 7 days before the experiment for acclimatizing to environmental conditions. Mice were housed in polysulfone plastic (PSU) individually ventilated cages (213 mm width × 362 mm diameter × 185 mm height) bedded with sterilized and dust-free bedding cobs. Animals had controlled light–dark cycle (14-h circadian cycle of artificial light) at 20–24 °C and 40–75% humidity. Each mouse was offered a complete pellet diet (150-SP-25, SAFE) and filtered, sterilized tap water ad libitum throughout the study.

Xenografts were established by subcutaneous surgical implantation of approximately 20 mm^3 into the flank of female athymic nude -Foxn1nu mice. Tumours reached 0.1–0.3 cm^3 before animals were randomly assigned into treatment groups. Tumour volume (mm^3) was calculated as [length × $width^2$]/2.

Animals were randomized into treatment groups, saruparib 1 mg kg^{-1}, camizestrant 10 mg kg^{-1} and palbociclib 50 mg kg^{-1} dosed PO daily, according to the tumour size criteria outlined above to obtain treatment arms with homogeneous geomean volumes. For both studies, tumours were measured twice weekly. Changes in tumour volume and growth inhibition were determined by bilateral Vernier caliper measurement (length × width). Length was the longest diameter across the tumour, and width was the corresponding perpendicular. Conscious animals were euthanized by cervical dislocation with secondary confirmation at the end of the study or for welfare condition. For analysis, two-way ANOVA followed by Sidak's multiple comparisons test was used to compare tumour volumes up to day 35 (last day of evaluable tumour volume among all groups) between the treatment conditions saruparib and palbociclib.

Reporting summary

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

Data availability

The assembled prospective somatic and germline mutational data from tumours needed to replicate our figures for the entire cohort are provided as Supplementary Tables. Deidentified clinical outcomes have been provided in the Supplementary Tables. BAM files from whole-genome sequencing of the post-CDK4/6i sample from PDX-L, as well as BAM files from deep-targeted sequencing from rapid autopsy have been uploaded to the European Genome Archive (accession EGAC50000000858). Otherwise, our analyses may be replicated with the supplementary files provided.

Code availability

The custom code and scripts for the current study are available on GitHub (https://github.com/antonmsafonov/germline-somatic-RB1/).

61. Schrader, K. A. et al. Germline variants in targeted tumor sequencing using matched normal DNA. *JAMA Oncol*. **2**, 104–111 (2016).
62. Shah, J. B. et al. Analysis of matched primary and recurrent BRCA1/2 mutation-associated tumors identifies recurrence-specific drivers. *Nat. Commun.* **13**, 6728 (2022).
63. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
64. 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).
65. Jonsson, P. et al. Tumour lineage shapes BRCA-mediated phenotypes. *Nature* https://doi.org/10.1038/s41586-019-1382-1 (2019).
66. Srinivasan, P. et al. The context-specific role of germline pathogenicity in tumorigenesis. *Nat. Genet.* **53**, 1577–1585 (2021).
67. Cheng, D. T. et al. Comprehensive detection of germline variants by MSK-IMPACT, a clinical diagnostic platform for solid tumor molecular oncology and concurrent cancer predisposition testing. *BMC Med. Genomics* **10**, 33 (2017).
68. Gupta, A. et al. APOBEC3 mutagenesis drives therapy resistance in breast cancer. *Nat. Genet*. **57**, 1452–1462 (2025).
69. Razavi, P. et al. Alterations in PTEN and ESR1 promote clinical resistance to alpelisib plus aromatase inhibitors. *Nat. Cancer* **1**, 382–393 (2020).
70. Razavi, P. et al. High-intensity sequencing reveals the sources of plasma circulating cell-free DNA variants. *Nat. Med.* **25**, 1928–1937 (2019).
71. Kudo, R. et al. Long-term breast cancer response to CDK4/6 inhibition defined by TP53-mediated geroconversion. *Cancer Cell* **42**, 1919–1935.e9 (2024).
72. Chakravarty, D. et al. OncoKB: a precision oncology knowledge base. *JCO Precis. Oncol.* **1**, PO.17.00011 (2017).
73. O'leary, B. et al. Circulating tumor DNA markers for early progression on fulvestrant with or without palbociclib in ER^+ advanced breast cancer. *J. Natl Cancer Inst.* **113**, 309–317 (2021).
74. Rose Brannon, A. et al. Enhanced specificity of clinical high-sensitivity tumor mutation profiling in cell-free DNA via paired normal sequencing using MSK-ACCESS. *Nat. Commun.* **12**, 3770 (2021).
75. Lanman, R. B. et al. Analytical and clinical validation of a digital sequencing panel for quantitative, highly accurate evaluation of cell-free circulating tumor DNA. *PLoS ONE* **10**, e0140712 (2015).
76. Fry, D. W. et al. Specific inhibition of cyclin-dependent kinase 4/6 by PD 0332991 and associated antitumor activity in human tumor xenografts. *Mol. Cancer Ther.* **3**, 1427–1437 (2004).
77. Costa, C. et al. Pten loss mediates clinical cross-resistance to CDK4/6 and PI3Kα inhibitors in breast cancer. *Cancer Discov.* **10**, 72–85 (2020).
78. Li, Q. et al. INK4 tumor suppressor proteins mediate resistance to CDK4/6 kinase inhibitors. *Cancer Discov.* **12**, 356–371 (2022).
79. Lloyd, M. R. et al. A bedside-to-bench translational analysis of NF1 alterations and CDK4/6 inhibitor resistance in hormone receptor-positive metastatic breast cancer. *EBioMedicine* **118**, 105828 (2025).
80. Shi, W. et al. Reliability of whole-exome sequencing for assessing intratumor genetic heterogeneity. *Cell Rep.* **25**, 1446–1457 (2018).
81. Cibulskis, K. et al. Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nat. Biotechnol.* **31**, 213–219 (2013).
82. Koboldt, D. C. et al. VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Res.* **22**, 568–576 (2012).
83. Saunders, C. T. et al. Strelka: accurate somatic small-variant calling from sequenced tumor-normal sample pairs. *Bioinformatics* **28**, 1811–1817 (2012).
84. Fang, H. et al. Indel variant analysis of short-read sequencing data with Scalpel. *Nat. Protoc.* **11**, 2529–2548 (2016).
85. Rimmer, A. et al. Integrating mapping-, assembly- and haplotype-based approaches for calling variants in clinical sequencing applications. *Nat. Genet.* **46**, 912–918 (2014).
86. Danecek, P. et al. Twelve years of SAMtools and BCFtools. *Gigascience* **10**, giab008 (2021).
87. Talevich, E., Shain, A. H., Botton, T. & Bastian, B. C. CNVkit: genome-wide copy number detection and visualization from targeted DNA sequencing. *PLoS Comput. Biol.* **12**, e1004873 (2016).
88. Chen, X. et al. Manta: rapid detection of structural variants and indels for germline and cancer sequencing applications. *Bioinformatics* **32**, 1220–1222 (2016).
89. Wala, J. A. et al. SvABA: genome-wide detection of structural variants and indels by local assembly. *Genome Res.* **28**, 581–591 (2018).
90. Cameron, D. L. et al. GRIDSS: sensitive and specific genomic rearrangement detection using positional de Bruijn graph assembly. *Genome Res.* **27**, 2050–2060 (2017).
91. Chen, S. et al. Paragraph: a graph-based structural variant genotyper for short-read sequence data. *Genome Biol.* **20**, 291 (2019).
92. Geoffroy, V. et al. AnnotSV: an integrated tool for structural variations annotation. *Bioinformatics* **34**, 3572–3574 (2018).
93. Moreno, T., Magana, J. & Quigley, D. A. AARDVARK: an automated reversion detector for variants affecting resistance kinetics. *Bioinformatics* **39**, btad509 (2023).
94. 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).

Acknowledgements We acknowledge the following funding sources: the Komen Career Catalyst Research Award (to P.R.); the MBCure Foundation (to P.R.); the Breast Cancer Research Foundation (BCRF; to P.R., S.C., K.L.N. and K.O.); the National Cancer Institute grants P50 CA247749 (to S.C. and P.R.) and R35 CA305347 (to S.C.); the Cancer Couch Fund (to S.C.); the Foss Family Foundation (to S.C. and P.R.); the Shen Family (to S.C.); the ASCO Conquer Cancer Foundation Young Investigator Award (to A.S.); the Fondazione Gianni Bonadonna International Post-Doctoral Research Fellowship Program (to L.B.B.); the Gray Foundation (to K.L.N. and M.E.R.); the Basser Center for BRCA (K.L.N.); and the Niehaus Center for Integrated Cancer Genomics (to K.O.). We thank our patients and their families for their participation in our studies; and to patient H.R. for participating in the Last Wish Program, whose invaluable contribution enabled the development of preclinical models and genomic evidence that were central to validating our findings.

Author contributions A.S., S.C. and P.R. conceived the study. A.S. led the clinicogenomic, genomic and statistical analyses under the guidance of P.R., with contributions from L.B.B., A.M., D.N.B., D. Muldoon, M.M., C.B. and M.T.A.D., and critical input from R.S., L.N., B.T.L., S.S., D.B.S. and S.C. M. Lee, A.Z., S.R., I.B., L.V., V.S., M.R.A., M.W., R.E.B., E. Marangoni, L.C., M.S., A.D.S. and S.C. conducted the preclinical experiments. A.S., D.N.B., C.B., P.S., R.S. and J.S.R.-F. developed the analytical pipelines for ASCN and whole-genome sequencing analyses. B.O., E.S., B. Wubbenhorst, E.N., N.L., C.C., Y.L., X.H., K.N.M., F.C., S.M.D., N.C.T. and K.L.N. contributed to the clinicogenomic validation analyses. M.M., C.B., D. Mandelker, M.T.A.D., M. Ladanyi, M.E.R., K.O., M.F.B., N.S. and D.B.S. developed the MSK-IMPACT germline and somatic analytical platforms and the germline variant pathogenicity annotation pipeline. J.A.-R.A. and H.S. contributed to data integration and project coordination. A.S., E.F., L.B.B., A.M., M.A., E. Moiso and P.R. contributed to clinical data curation. B. Weigelt, H.S., D.R., F.P., D. Mandelker and J.S.R.-F. contributed to the pathological analyses. A.S., M. Lee, D.N.B., S.C. and P.R. wrote the manuscript with input from all authors. P.R. and S.C. supervised the overall study.

Competing interests A.S. has received travel support from Novartis. J.S.R.-F., M.S., E.N., N.L., E.S., S.R., M.R.A., A.D.S. and V.S. are currently employed by AstraZeneca. Y.L. and X.H. are currently employed by Pfizer. M.F.B. has received consultation from AstraZeneca, Eli Lilly and Paige AI; has received research support from Boundless Bio; and declares intellectual property interests with SOPHiA Genetics. B.T.L. is a current employee and stockholder of AstraZeneca, and the research was conducted while he served as faculty at MSK; he has served as an uncompensated advisor and consultant to Amgen, AstraZeneca, Boehringer Ingelheim, Bolt Biotherapeutics, Daiichi Sankyo, Genentech and Lilly; has received research grants to his institution from Amgen, AstraZeneca, Bolt Biotherapeutics, Daiichi Sankyo, Genentech, Lilly, Nuvalent and Revolution Medicines; is an inventor on three institutional patents at MSK (US62/685,057, US62/514,661 and US63/424,813); and has intellectual property rights as a book author at Karger Publishers and Shanghai Jiao Tong University Press. M. Lee reports equity and professional services and activites from Biologica. S.C. has received institutional grants or funding from Daiichi Sankyo and AstraZeneca; share options in Totus Medicines; and consultation, ad board or honoraria from Daiichi Sankyo, AstraZeneca, Lilly, Casdin Capital, Merck and Pathos AI. P.R. has received institutional grants and/or funding from Grail, Novartis, AstraZeneca, Invitae, Biothernostics, Tempus, Neogenomics, Guardant Health, Personalis, Myriad Genetics, Foresight Diagnostics, Natera, Biodesix, SAGA Diagnostics, SOPHiA Genetics, Haystack and Roche; and consultation, ad board or honoraria from Novartis, AstraZeneca, Lilly/Loxo, Stemline Therapeutics, Prelude Therapeutics, Neogenomics, Regor Pharmaceuticals, SAGA Diagnostics, SOPHiA Genetics, Tempus, Myriad Genetics, Foresight Diagnostics, Natera, Pathos AI and BioNTech. The other authors declare no competing interests.

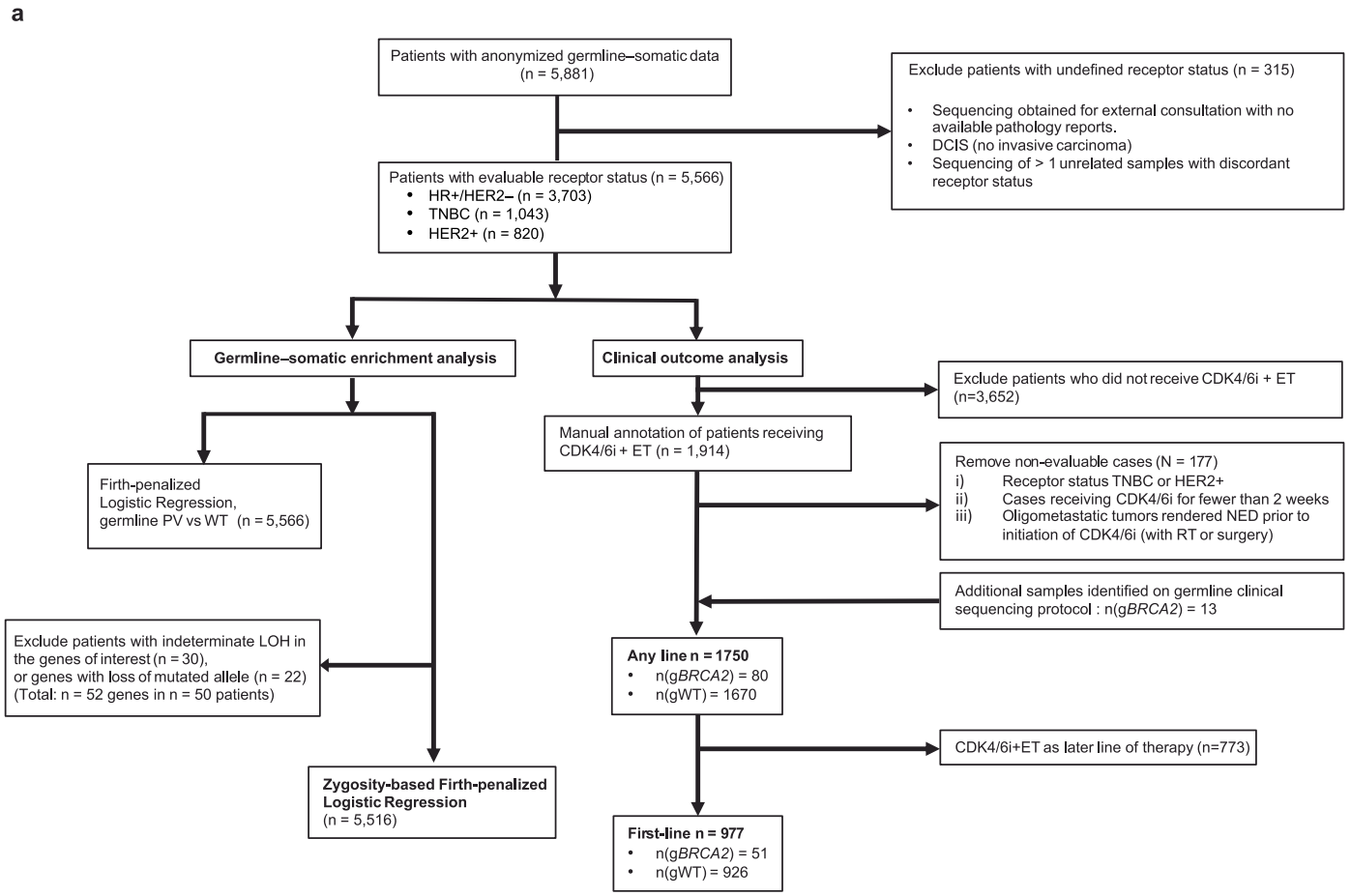
Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41586-026-10197-0.

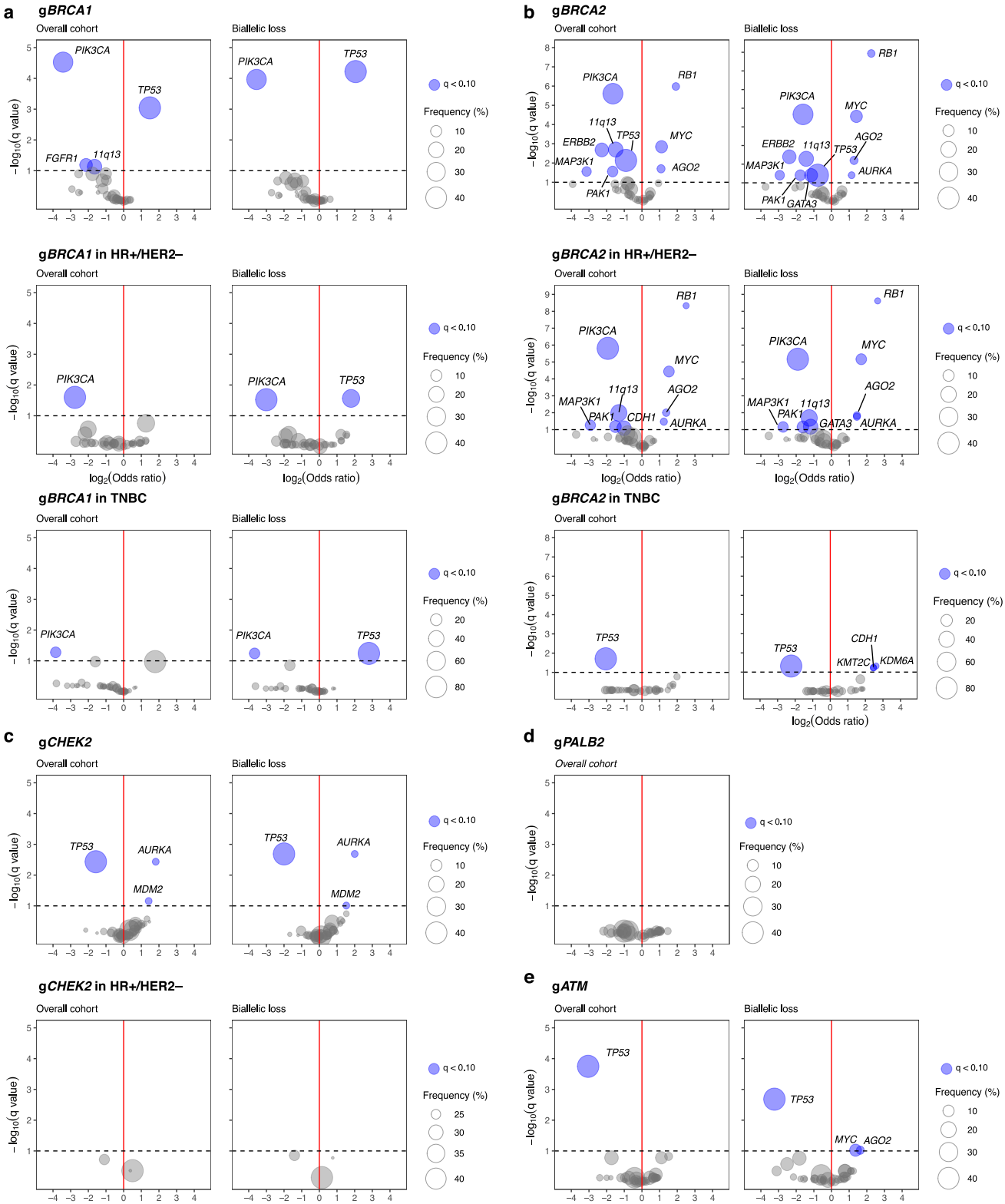
Correspondence and requests for materials should be addressed to Sarat Chandarlapaty or Pedram Razavi.

Peer review information *Nature* thanks the anonymous reviewers for their contribution to the peer review of this work.

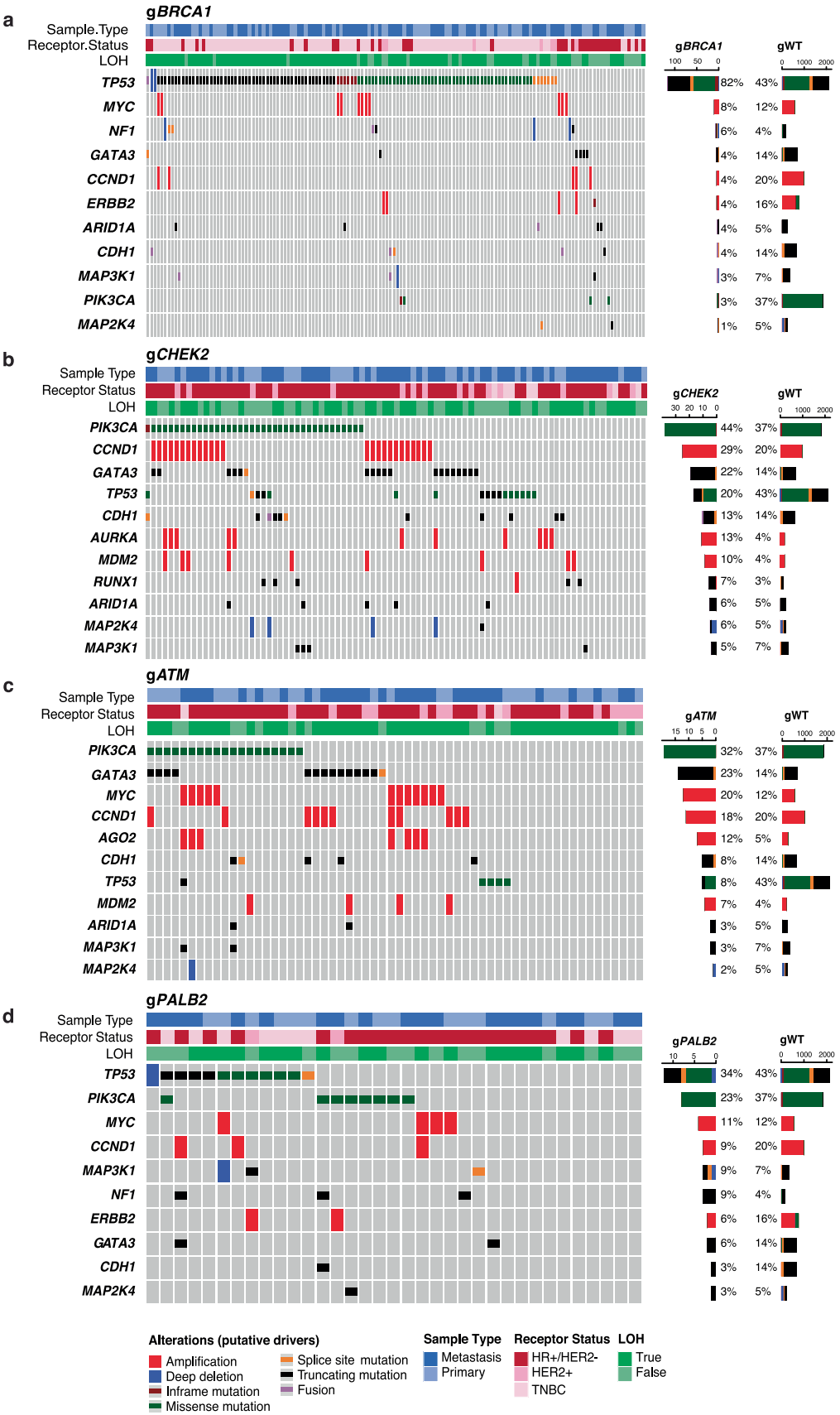
Reprints and permissions information is available at http://www.nature.com/reprints.



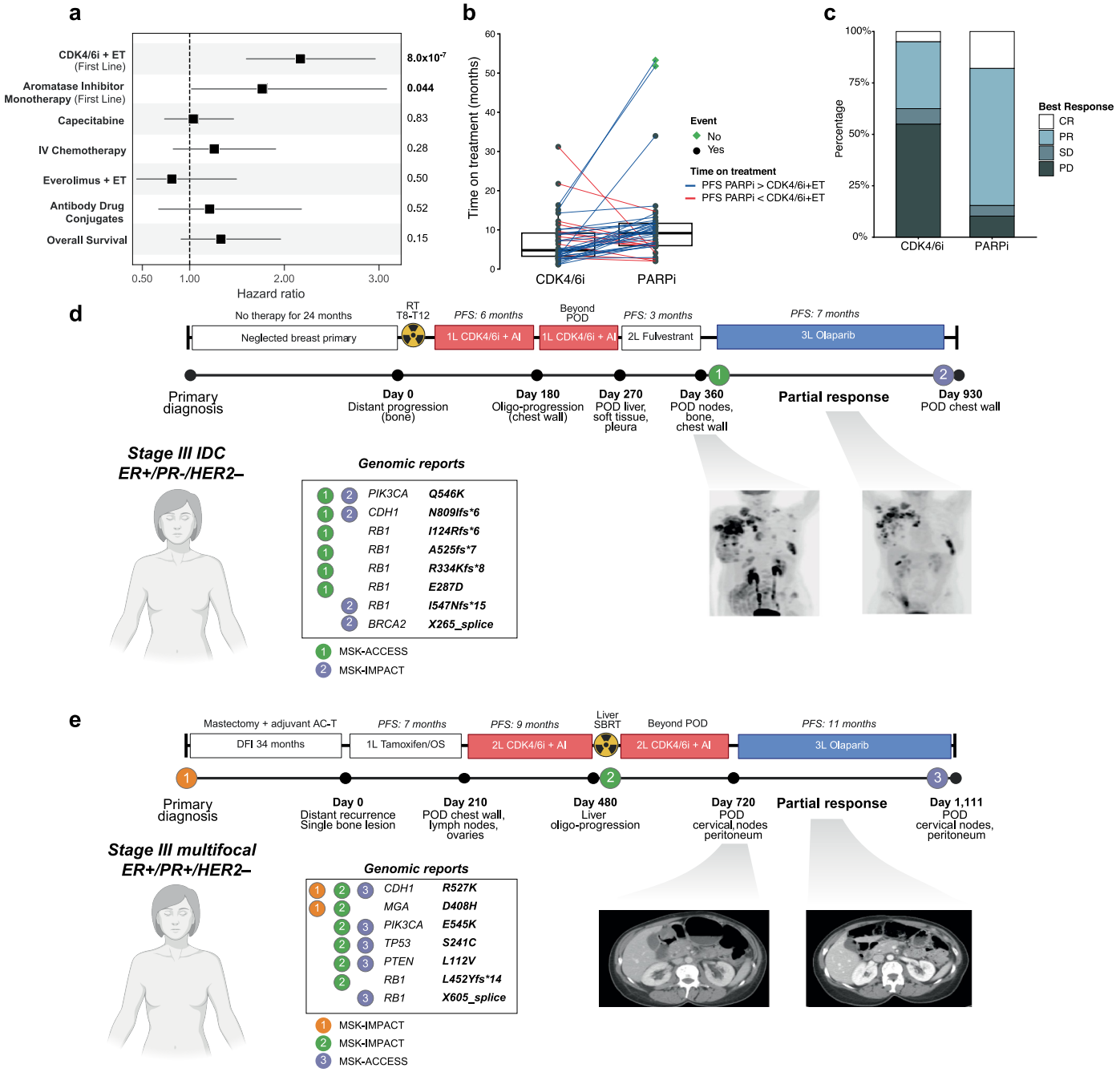
Extended Data Fig. 1 | CONSORT diagram and zygosity status. a, CONSORT diagram outlining inclusion criteria for germline–somatic analysis and CDK4/6i outcome analysis comparing g*BRCA2* to g*WT*. Cases were excluded from further clinicogenomic analysis if clinical or pathological data were not available, if the patient did not have invasive breast cancer (i.e. ductal carcinoma in situ was sequenced), and if the patient had multiple evolutionarily distinct tumors with discordant receptor status (n = 315 excluded total). Patients were excluded if zygosity could not be ascertained due to technical reasons (including indeterminate LOH in the genes of interest); patients with loss of mutated allele were also excluded from analysis. Patients were excluded from downstream analysis of CDK4/6i + ET and endocrine therapy outcomes for the following reasons if: i) initial receptor status was TNBC or HER2 + , ii) the patients received CDK4/6i for <2 weeks prior to toxicity, iii) if all the known sites of disease were ablated prior to initiation of CDK4/6i + ET (e.g., oligometastatic disease rendered to have "no evidence of disease" prior to initiation of systemic after radiotherapy). **b**, Zygosity status of germline PVs involved in homologous recombination pathway. Descriptive summary of zygosity status of each gene included in our study, distinguishing between mechanism of loss of WT allele (LOH or somatic second hit LoF mutation). Loss of the mutated allele (Loss PV) is also labeled. Biallelic inactivation rates varied across genes. As this is a descriptive figure, no statistical analysis was performed.



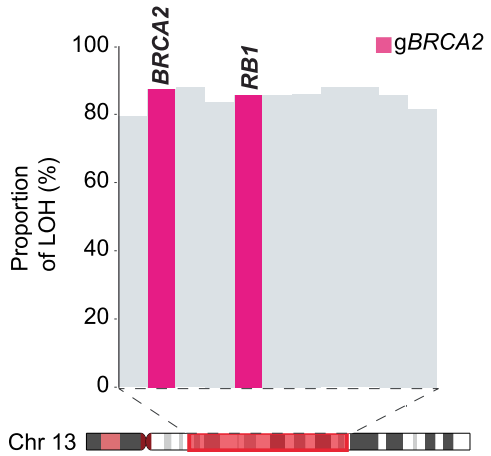
Extended Data Fig. 2 | Germline–somatic interactions, stratified by zygosity status. a-e, Enrichment analysis of germline PVs compared to gWT. Analyses for g*BRCA1* (**a**), g*BRCA2* (**b**), g*CHEK2* (**c**), g*PALB2* (**d**); g*ATM* (**e**). Overall results were generated by iteratively applying Firth penalized regression across all genes with a somatic variant frequency >3% in the cohort. Receptor status and zygosity were employed as covariates. For each analysis, *q* values were calculated, and significant somatic–germline interactions were defined as $q < 0.10$. Where indicated, overall cohort analysis is followed by stratification by allelic status and receptor status.



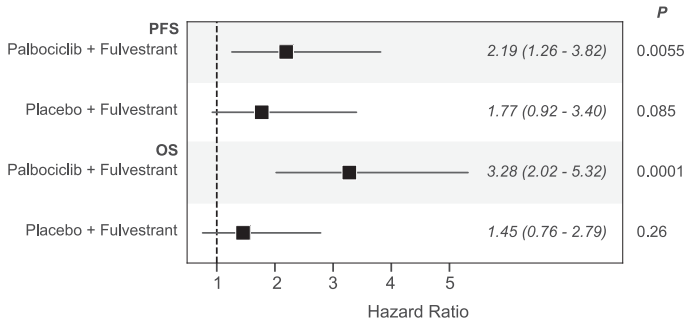
Extended Data Fig. 3 | Oncoprints of enriched germline–somatic interactions. Oncoprints showing somatic alteration types across patients (columns) and genes (rows). Somatic mutation types, copy number alterations, and fusions in the indicated genes are annotated; only genes identified as significant in the germline–somatic interaction analyses are included. Receptor status, sample type, and zygosity are annotated above the oncoprint. **a**, *gBRCA1* vs. gWT. **b**, g*CHEK2* vs. gWT. **c**, g*ATM* vs. gWT. **d**, g*PALB2* vs. gWT.



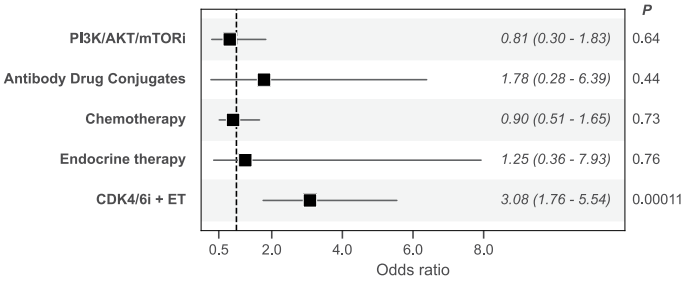
Extended Data Fig. 4 | Impact of g*BRCA2* status on therapeutic response across systemic modalities. a, Forest plot depicting the PFS by g*BRCA2* status on the outcome of other therapeutic modalities in MBC, as well as OS of g*BRCA2* status of the CDK4/6i cohort ($n_{gBRCA2} = 51$, $n_{WT} = 925$). Additional therapeutic settings studied included endocrine monotherapy ($n_{gBRCA2} = 15$, $n_{WT} = 670$), capecitabine ($n_{gBRCA2} = 46$, $n_{WT} = 1{,}173$), intravenous chemotherapy ($n_{gBRCA2} = 36$, $n_{WT} = 696$), mTOR inhibitor ($n_{gBRCA2} = 21$, $n_{WT} = 620$), and antibody–drug conjugates ($n_{gBRCA2} = 13$, $n_{WT} = 347$). Analysis was performed by Cox proportional hazard model. Data is reported as HRs, with bars representing 95% CIs. **b**, Ladder plot demonstrating time on treatment between CDK4/6i and subsequent-line PARPi for n = 41 independent patients. Progression is indicated by black dots, and ongoing treatment by green triangles. The upper and lower hinges of the box plots represent the 75th and 25th percentile of treatment duration in each group, with the central line representing the median. **c**, Comparison of best radiographic response between CDK4/6i and subsequent-line PARPi. **d,e**, Two representative g*BRCA2* patients with rapid progression on CDK4/6i + ET. In both cases, post-CDK4/6i samples harbored multiple *RB1* LoF alterations, suggestive of convergent evolution. Both patients showed relatively longer responses on subsequent-line PARPi. Schematics in panels **d**,**e** created in BioRender; Razavi, P. https://biorender.com/c2o6wfb (2025).



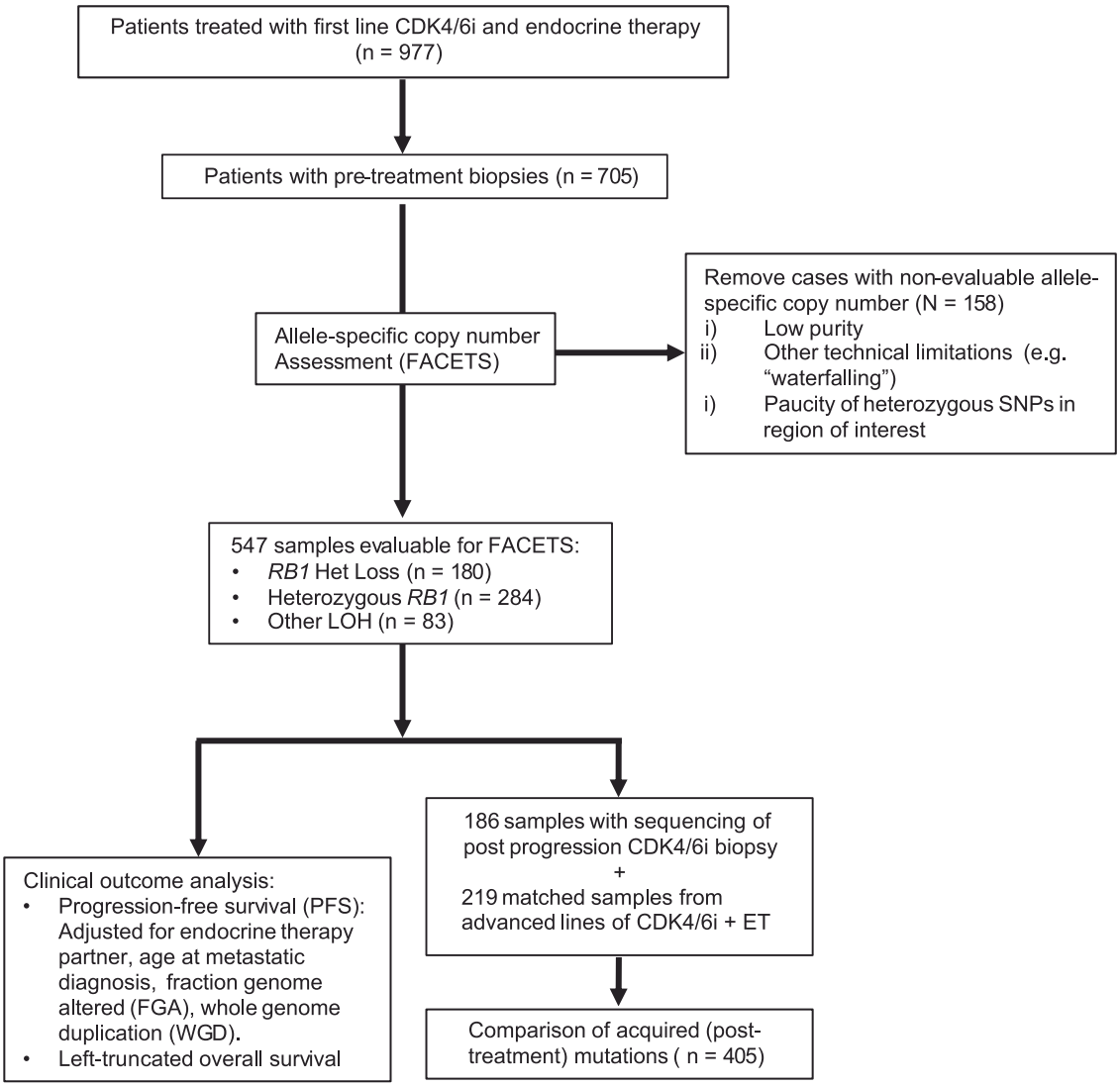
Extended Data Fig. 5 | Validation of concurrent *RB1* and *BRCA2* LOH in external g*BRCA2* WES cohort. FACETS was performed on whole-exome sequencing of g*BRCA2* breast cancer samples from Abramson Cancer Center at the University of Pennsylvania (n = 24), and the Mayo Clinic (n = 22). Of these samples, 38 (82.6%) demonstrated concurrent LOH of *BRCA2* and *RB1*. The association between *RB1* LOH and *BRCA2* LOH was statistically significant by two-sided Fisher's exact test (OR = Inf, 95% CI: 11.12 – Inf, $P < 0.00001$).



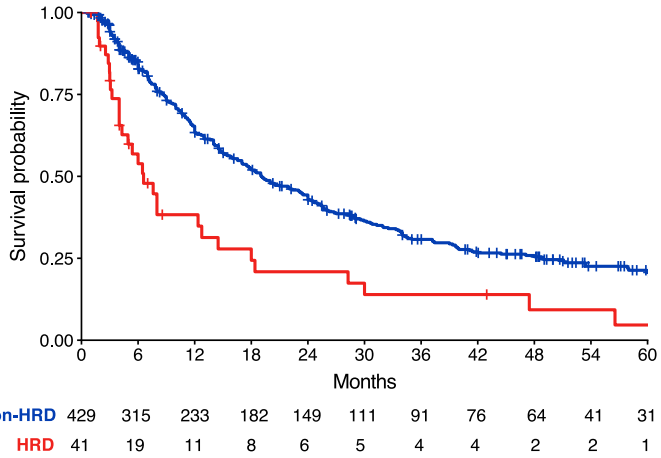
Extended Data Fig. 6 | Progression-free and overall survival in PALOMA-3. PFS and OS results from the experimental arm (fulvestrant plus palbociclib) and the placebo arm of PALOMA-3 (fulvestrant plus placebo) by baseline *RB1* LOH (Experimental arm: $n_{RB1\text{-LOH}} = 19$, $n_{\text{No-}RB1\text{-LOH}} = 240$, Placebo arm: $n_{RB1\text{-LOH}} = 11$, $n_{\text{No-}RB1\text{-LOH}} = 131$). The PFS and OS results from the experimental arm are replicated from Fig. 3d,e. HRs were estimated using Cox proportional hazard models (*P* from two-sided Wald test). No multiple comparisons adjustment. Data are reported as HRs, with bars representing 95% CIs.



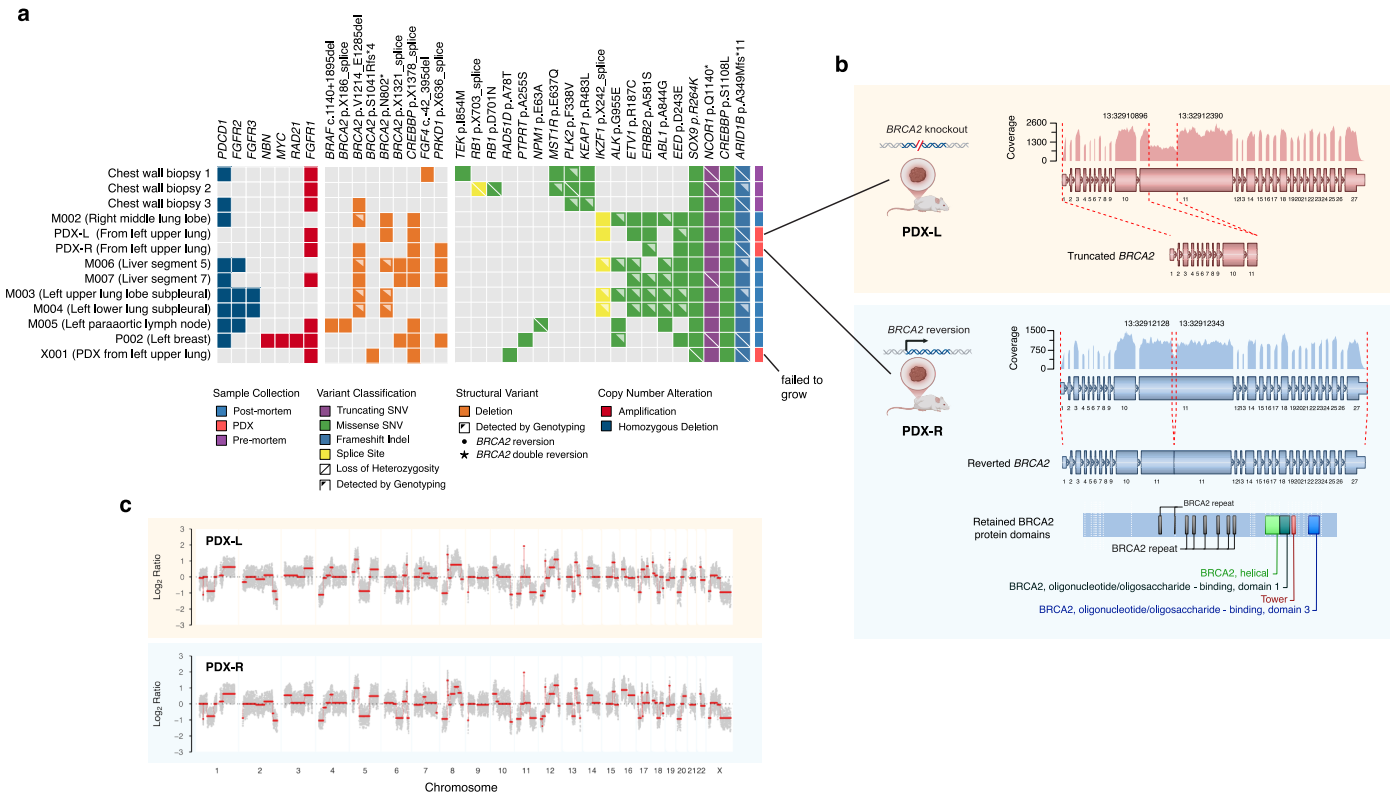
Extended Data Fig. 7 | Association between preceding treatment and *RB1* loss-of-function variant. Sequenced tumor samples from patients with HR+/HER2− MBC were categorized by preceding therapeutic exposure: endocrine monotherapy (n = 1,818 samples sequenced after therapy), CDK4/6i (n = 926), chemotherapy (n = 1,395), antibody–drug conjugate (n = 92), other targeted therapy (PI3K, mTOR or AKT inhibitors, n = 277). A multivariable logistic regression model was constructed with preceding treatment type as independent variable and presence of *RB1* LoF as dependent variable. Data are reported as ORs, with bars representing 95% CIs. Two-sided Wald tests were used for inference. Only prior CDK4/6i exposure was significantly associated with *RB1* LoF (OR = 3.08, 95% CI: 1.76 – 5.54, $P = 0.00011$). *RB1* LoF was not enriched after endocrine monotherapy (OR = 1.25, 95% CI: 0.36 – 7.93, $P = 0.76$), PI3K/AKT/mTORi (OR = 0.81, 95% CI: 0.30 – 1.83, $P = 0.64$), antibody–drug conjugate therapy (OR = 1.78, 95% CI: 0.28 – 6.39, $P = 0.45$), or chemotherapy (OR = 0.90, 95% CI: 0.51 – 1.65, $P = 0.74$).



Extended Data Fig. 8 | CONSORT diagram for case selection for first-line CDK4/6i + ET outcomes stratified by allele-specific copy number of *RB1*. ASCN analysis was performed on tumor samples collected prior to initiation of first-line CDK4/6i + ET. Samples were excluded (n = 158) for technical reasons including low purity or paucity of heterozygous SNPs in the region of interest. The CONSORT diagram describes the clinical outcomes analysis on the 547 samples remaining, as well as the matched pairs analysis on the 405 samples with corresponding post progression biopsies.



Extended Data Fig. 9 | Implications of HRD signature on CDK4/6i PFS in g*BRCA2* WT breast cancers. HRD inference was performed with an orthogonal method (HRD-IMPACT, see Methods). This included calculation of large state transitions (LST), telomeric allelic imbalance (TAI) and LOH from samples that passed purity and ASCN imputation QC metrics. The HR was derived using a Cox proportional hazard model (*P* from two-sided Wald test). Among g*BRCA2* WT patients, HRD was associated with a significantly shorter PFS (HR = 1.78, 95% CI: 1.27 – 2.48, *P* = 0.00075).



Extended Data Fig. 10 | Multi-site genomic profiling from rapid autopsy of a g*BRCA2* carrier. a, Oncoprint summarizing targeted sequencing (MSK-IMPACT) of tumors collected at rapid autopsy as well as prior clinical biopsies from the same patient. Each column represents an individual tumor sample, grouped by clinical versus autopsy samples. Of note, an *RB1* splice variant was detected exclusively in chest wall biopsy 2, performed in the clinical setting. This alteration was not identified in subsequent clinical or autopsy samples, despite deep targeted sequencing and genotyping of the sequencing reads. **b**, PDXs were created from two distinct metastatic deposits, both in the left upper lung. One sample demonstrated a large deletion (spanning from 32,910,896 to 32,912,390), resulting in a truncating mutation and conferring *BRCA2* protein loss (PDX-L). Another sample demonstrated an in-frame deletion which encompassed the site of the original PV, resulting in restoration of the reading frame and a predicted *BRCA2* reversion variant (PDX-R). **c**, Aside from the SV shown in **b**, PDX-L and PDX-R exhibited overlapping copy number patterns, reflecting a strong genomic concordance between the two samples.

nature portfolio

Corresponding author(s): Pedram Razavi
Sarat Chandarlapaty

Last updated by author(s): 1/12/2026

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒ A description of all covariates tested
☐	☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection

No software was used for data collection.

Data analysis

Study cohort and prospective sequencing
The study cohort comprised 6,927 tumor samples from 5,881 breast cancer patients. All patients underwent prospective clinical tumor and normal DNA sequencing as part of their clinical care (February 2014 to September 2021). The present study was approved by the Memorial Sloan Kettering Cancer Center (MSK) Institutional Review Board (IRB) and all patients provided written informed consent for tumor and paired normal DNA sequencing and review of medical records for clinical annotations. Genomic sequencing was performed on tumor DNA extracted from formalin-fixed, paraffin-embedded tissue and normal DNA extracted from mononuclear cells from peripheral blood in all patients as previously described. Patient samples were sequenced in a CLIA-compliant laboratory using one of several versions of the MSK-IMPACT targeted sequencing panel, which interrogates exonic and selected intronic regions of 341, 410, 468 or 506 genes depending on the assay version, with somatic mutation calling performed using an extensively validated pipeline followed by manual review, as described previously. Tumors were obtained from the primary site in 50.1% (n = 3,470) of samples and from a metastatic site in 49.8% (n = 3,457) of samples.

Anonymized germline variant calling was performed using a sequence analysis pipeline validated for clinical use in a CLIA-compliant laboratory performing clinical sequencing of patient tumors and matched normal blood specimens as a part of routine clinical care21. Germline variants with a population allele frequency <2% in gnomAD exomes (v2.0.1) and genomes (v2.1.1) were assessed for pathogenicity following ACMG guidelines. Variants predicted by VEP to have high-impact consequences were considered putative LoF variants. Variants were designated as pathogenic if classified as pathogenic or likely pathogenic in-house or by ClinVar or predicted to be LoF in a TSG. We excluded variants flagged as potentially arising from clonal hematopoiesis or circulating tumor cells as previously described.

We manually reviewed variants with discordant interpretations between ClinVar and in-house classifications, as well as variants with

conflicting interpretations in ClinVar. Novel LoF variants were considered pathogenic unless located in the terminal exon, where such variants were considered pathogenic only if the same gene harbored previously designated pathogenic LoF variants located further downstream in the coding sequence. Low-risk variants in CHEK2 such as c.470T>C were excluded from the analysis.

Population frequencies were obtained from gnomAD exomes (v2.0.1) and gnomAD genomes (2.1.1). Variants were annotated for pathogenicity interpretations using ClinVar (accessed December 2024) and in-house classifications (as of August 2022) by expert clinical geneticists. Ensembl Variant Effect Predictor was used to annotate variants for predicted functional consequences.

The external cohort consisted of tumor samples from gBRCA2 primary tumors which underwent whole-exome sequencing (WES) from samples collected at University of Pennsylvania and Mayo Clinic. Patients gave written informed consent for research use of germline DNA and tumor specimens under IRB-approved protocol at University of Pennsylvania and Mayo Clinic. As previously described, somatic tumor DNA was extracted from FFPE primary breast cancer specimens using standard laboratory deparaffinization, while germline DNA was extracted from whole blood or saliva. Tumor DNA libraries were prepared using the NEBNext FFPE Repair mix and NEBNext Ultra II DNA library prep kit (New England Biolabs), per manufacturer's instructions. Germline DNA libraries were prepared using the NEBNext Ultra DNA library prep kit (New England Biolabs), per manufacturer's instructions. DNA libraries were pooled and hybridized using SureSelect Target Enrichment System for Illumina Multiplex Sequencing (Agilent) and associated protocols. For WES, tumor and germline libraries were hybridized to SureSelect All Exon v5, SureSelect All Exon v6+COSMIC, and SureSelect All Exon v7 captures (Agilent). WES was performed using an Illumina HiSeq 4000 and targeted sequencing was performed using an Illumina NovaSeq 6000. All sequencing was performed with 150 paired-end reads by the University of Pennsylvania Next Generation Sequencing Core.

FASTQ files from whole-exome and targeted sequencing were aligned to the hg19 build of the human genome using the Burrows-Wheeler Aligner (BWA v.0.7.17-r1188). Various bam file processing operations were performed using Samtools/htslib/bcftools v1.11. The resulting bam files were processed according to Genome Analysis Toolkit (GATK v3.7) best practices (picardtools v2.20.7).

Zygosity inference

We inferred somatic zygosity for all germline PVs using locus- and allele-specific copy number (ASCN) inference utilizing FACETS as previously described. All ASCN solutions (FACETS outputs) from tumors with germline PVs were manually reviewed to ensure that the optimal solution was selected. Purity estimates were similarly inferred from the FACETS. We incorporated ASCN, purity and variant allele frequency into a previously described framework, allowing statistical inference of heterozygous, biallelic (loss of WT allele), or loss of mutant allele. The zygosity of the germline variant was considered indeterminate and excluded from zygosity analyses if the: (1) variant was homozygous in the germline; (2) read depth of coverage in the normal blood specimen was <50; (3) FACETS-derived total and minor copy number were not evaluable at the corresponding locus, or (4) An optimal solution could not be identified, most commonly due to low tumor purity. Using these criteria, out of the 627 total cases with germline PV in the genes of interest, somatic zygosity status was evaluable for 586 (93.4%) of tumors. We excluded cases where copy number analysis and variant allele frequency was suggestive of loss of mutant allele, rather than WT (n = 30, 5.1%) from further analyses.

To initially determine whether a given germline variant was in allelic imbalance in the corresponding tumor specimen, we evaluated consistency between observed somatic VAF and expected VAF. The latter value was calculated as a function of ASCN and purity as defined previously. Germline variants were considered heterozygous if their observed VAF was either (1) consistent with the expected VAF (within its 95% binomial CI) given balanced heterozygosity (tcn and lcn of either 2 and 1 or 4 and 2 in diploid and genome doubled tumors, respectively), or (2) less than the lower bound of the 95% CI of the expected VAF corresponding to a tcn and lcn of 3 and 1, respectively, which was either single copy gain of the mutant or WT allele. Germline variants in allelic imbalance of any kind were those with an observed VAF that was either within or greater than the 95% CI of the expected VAF corresponding to a copy-number state other than balanced heterozygosity. For allelically imbalanced germline variants, loss of the WT was determined as those with an observed VAF within the 95% CI (or greater than the lower bound of the 95% CI) of the expected VAF corresponding to a lcn equal to 0 (observed VAF is concordant with the expected VAF when the lesser allele has a copy number of 0).

Data anonymization

All patients (n = 5,881) consented to an IRB-approved protocol allowing analysis of somatic and clinical data (NCT01775072). A subset of 2,896 patients had additionally consented for identified analysis of germline variants under this protocol whereas the remaining patients (n = 2,985) consented only to somatic analyses.

For analyses involving germline PV status, we anonymized data into a unique anonymized ID (A-ID-#####), enabling germline calls to be conducted on matched normal samples for all patients regardless of germline consent status in accordance with MSK IRB guidelines for anonymized germline–somatic analyses, as previously described. For all patients, data were binned to avoid unique clinical or somatic alteration values and thereby prevent re-identification. Briefly, for patients who did not provide consent for identified germline analysis, genomic data was anonymized with a deterministic one-way hash function. In these patients, germline variant calling was performed using the clinically validated pipeline described above, and PFS and OS times were rounded to the nearest month. For analyses in which germline data were not required, a separate study-specific unique identifier (S-ID-#####) and fully anonymized clinical data were used for all the patients regardless of germline consent status. Continuous clinical (e.g. time to progression) and somatic genomic data were therefore able to be used for analyses which did not incorporate germline data.

Allele-Specific Copy Number Definitions

In the MSK cohort, somatic loss of heterozygosity events of RB1 were defined based on manual FACETS review of pre-CDK4/6i treatment samples. Pre-CDK4/6i ASCN analysis was performed on samples collected prior to first-line CDK4/6i treatment or as part of a matched pre- and post-treatment CDK4/6i pair. Of 922 patients meeting this criteria, 196 samples were excluded due to low purity or other technical limitations (such as "waterfalling" artifact), or paucity of heterozygous SNPs allowing for confident lesser copy number inference. RB1 loss of heterozygosity (LOH) was defined as lesser copy number of zero (LCN = 0), irrespective of total copy number (TCN), while heterozygous state was any lesser copy number greater than or equal to one (LCN > 0).

In the PALOMA-3 cohort, baseline ctDNA samples from the palbociclib combination arm were sequenced using a 1,729 amplicon custom AmpliSeq panel, which included 119 SNPs located within the RB1 gene. ctDNA-based LOH analysis was conducted using a bespoke pipeline; RB1 LOH was defined as previously described in O'Leary et al.

In the MSK cohort, we further separated the RB1 LOH (LCN = 0) group into i) heterozygous loss ("Het Loss"), defined as a state with total copy number of one and lesser copy number of zero (TCN = 1 & LCN = 0) and ii) "other LOH" defined as a state with total copy number greater than

one and lesser copy number of zero (TCN > 1 & LCN = 0). Fraction genome altered (FGA) was also calculated for each pre-treatment sample and defined as the fraction of log2 copy number variation (gain or loss) > 0.2 divided by the total size of the copy number profiled region.

Homologous Recombination Inference Analysis
To study the implications of HRD on clinical outcomes and mechanism of resistance, we inferred HRD from two orthogonal methods which have been validated for use with targeted NGS data. IMPACT-HRD quantifies genomic scars associated with HRD by analyzing ASCN alterations determined with the FACETS algorithm (version 0.5.14) and computing those genomic scars with the impact-hrd package (https://github.com/mskcc/facets-suite/blob/master/R/copy-number-scores.R). All IMPACT-HRD assessments were completed using R Version 4.1.2. In particular, three metrics were evaluated: number of telomeric allelic imbalances (NtAI), large-scale transitions (LST), and losses of heterozygosity (HRD-LOH). The overall HRD phenotype is defined as the unweighted sum of these three metrics (HRD-sum).

Targeted sequencing of post-mortem and PDX studies
Post-mortem tissue samples were selected for DNA extraction, library prep and targeted sequencing. Up to 30 mg frozen tissue were digested with 40 µl of Proteinase K (600 mAU/ml) in 360 µl Buffer ATL at 56°. Genomic DNA (gDNA) was isolated using the DNAeasy Blood & Tissue Kit (QIAGEN catalog # 69504) according to the manufacturer's protocol, including treatment with RNase A. DNA was eluted in 60 µL 0.5X Buffer AE heated to 55°C.

After PicoGreen quantification and quality control using an Agilent BioAnalyzer, 100 ng of gDNA were used to prepare libraries using the KAPA Hyper Prep Kit (Roche catalog # 07961901001) with 8 cycles of PCR. 100–135 ng of each barcoded library were captured by hybridization in a pool of 9 samples using the IMPACT (Integrated Mutation Profiling of Actionable Cancer Targets) assay (IDT), designed to capture all protein-coding exons and select introns of 505 commonly implicated oncogenes, tumor suppressor genes, and members of pathways deemed actionable by targeted therapies. Captured pools were sequenced using an Illumina NovaSeq 6000 in PE100 run mode using the NovaSeq 6000 S4 Reagent Kit (200 cycles). All experiments were carried out at MSK's Integrated Genomics Organization.
The demultiplexed FASTQ files from the post-mortem samples were aligned to the human genome reference GRCh37/hg19 using bwa mem (version 0.7.17-r118863) and deduplicated using Picard MarkDuplicates (version 2.21.8). Quality control metrics of the alignments included (1) Unique PF aligned read pairs, (2) Mean target coverage, (3) Mean insert size and (4) Major/minor contamination.

Variant calling was performed using a previously described pipeline. Briefly, SNVs were detected in the tumor–normal pairs using Mutect (version 1.1.6) whilst indels were detected using a consensus of Varscan 2 (version 2.4.6), Strelka (version 2.9.10), Scalpel (version 0.5.4) and Platypus (version 0.8.1.2). Variants found with >0% global AF in the 1000 Genomes database (phase 3) or >0.01% across any population in the ExAC database (release 0.3.1) or that were covered by 10 reads in the tumor or 5 reads in the germline were filtered out. Variants for which the tumor variant allele fraction was <5 times than that of the normal variant allele fraction were filtered out. The aggregated set of variants identified in the tissues and xenografts were re-genotyped in all samples using SAMtools mpileup (version 1.19.2). Copy number alterations were detected using Facets (version 0.6.2). In addition, off-target reads were used to estimate Log2 Ratios using CNVkit (version 0.9.8). Structural variants were detected using the consensus of Manta (version 1.6.0), SvABA (version 1.1.0) and GRIDSS (version 2.13.2). The aggregated set of structural variants identified were re-genotyped in all samples using Paragraph (version 2.3) and annotation of the structural variants was done using vcf2maf (version 1.6.22; https://github.com/mskcc/vcf2maf) and AnnotSV (version 3.5.3). Reversion mutations and structural variants affecting BRCA2 were further classified using aardvark (version 0.35).

Whole-genome sequencing of patient-derived xenograft samples
Post-CDK4/6i patient-derived xenograft samples were selected for DNA extraction, library prep and whole genome sequencing. The tissue samples were homogenized in 500µL MagMAX DNA Cell and Tissue Extraction Buffer (ThermoFisher catalog # A45469) for up to 40s and DNA from lysate was extracted using the MagMAX DNA Multi-Sample Ultra 2.0 Kit (ThermoFisher catalog # A36570) on the KingFisher Apex System (ThermoFisher) according to the manufacturer's protocol. The samples were eluted in 80 µL elution solution.

After PicoGreen quantification and quality control using an Agilent TapeStation, 500 ng of gDNA were sheared using a LE220-plus Focused-ultrasonicator (Covaris catalog # 500569) and sequencing libraries were prepared using the KAPA EvoPrep Kit (Roche catalog # 10212250702) with modifications. The libraries were subjected to a 0.5x size selection using aMPure XP beads (Beckman Coulter catalog # A63882) after post-ligation cleanup. The libraries were not amplified by PCR and were pooled at equal volume. The samples were sequenced using an Illumina NovaSeq X in PE150 run mode using the NovaSeq X 25B Reagent Kit. All experiments were carried out at MSK's Integrated Genomics Organization.

The demultiplexed FASTQ files were aligned to a chimeric genome reference comprising the human reference GRCh37/hg19 and the mouse GRCm38/mm10 using bwa mem (version 0.7.17-r118863) Read pairs where at least one end (R1 and/or R2) had a primary alignment the mouse genome were filtered out and the remaining read pairs were re-aligned to the human reference GRCh37/hg19 as described above. Quality control of the alignments was done as previously described above whereby the % of mouse content was quantified as number of Unique PF read pairs aligned to the mouse genome relative to the total Unique PF aligned read pairs. Copy number alterations were detected using Facets (version 0.6.2). Structural variants were detected using the consensus of Manta (version 1.6.0), SvABA (version 1.1.0) and GRIDSS (version 2.13.2). The aggregated set of structural variants identified in the xenografts were re-genotyped in all samples using Paragraph (version 2.3) and annotation of the structural variants was done using vcf2maf (version 1.6.22; https://github.com/mskcc/vcf2maf) and AnnotSV (version 3.5.3).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

The assembled prospective anonymized somatic and germline mutational data needed to replicate our figures for the entire cohort are provided as Supplementary Tables. Deidentified clinical outcomes have been provided in the Supplementary Tables. Whole genome sequencing from gBRCA2 PDX-L as well as multi-site sequencing from rapid autopsy (sequenced with targeted next generation sequencing) is provided through the European Genome Archive: Accession EGAC50000000858

Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

Reporting on sex and gender	Male patients composed 1.0% (n = 59) of this breast cancer cohort (otherwise composed of n = 5,881 total patients, 5,822 females). Hence, we did not separately include sex as a covariate in clinical outcome or genomic enrichment analyses. This demographic data is included in Table 1 in summary form, and in Supplementary Table 1 in a disaggregated format.
Reporting on race, ethnicity, or other socially relevant groupings	We did not report race or ethnicity characteristics in this version of this manuscript.
Population characteristics	The study cohort comprised 6,927 tumor samples from 5,881 patients who had a diagnosis of breast cancer. All patients underwent prospective clinical tumor and germline sequencing as part of their clinical care (February 2014 to September 2021). Further clinical information, including germline genotype, select demographics and breast cancer receptor status, is summarized in Table 1.
Recruitment	The study cohort comprised 6,927 tumor samples from 5,881 breast cancer patients. All patients underwent prospective clinical tumor and germline sequencing as part of their clinical care (February 2014 to September 2021).
Ethics oversight	The present study was approved by the Memorial Sloan Kettering Cancer Center (MSK) Institutional Review Board (IRB, biospecimen research protocol 21-149, which serves as a retrospective analysis of data generated from prospective clinical sequencing protocol (NCT01775072) Mouse studies were conducted through the MSK antitumor core facility in compliance with institutional guidelines under an IACUC approved protocol (MSK IRB 12-10-016). PDXs were established by implanting freshly collected autopsy samples from a patient in MSK's Last Wish Program which enables patients at their end of their life to donate their bodies for research. The samples were collected at MSK under approved IRB biospecimen protocols (MSK IRB 12-245 and 06-107). Animals were maintained in accordance with the Guide for the Care and Use of Laboratory Animals in an AAALAC-accredited facility. All procedures outlined in the study were approved by the Memorial Sloan-Kettering Cancer Center IACUC

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	As this was a retrospective study, sample size was not predetermined. This cohort consisted of all patients sequenced through the dates of the study. To illustrate the adequate power of this retrospective study, consider the following: With regard to the investigation of the effect of germline BRCA2 status on progression free survival (PFS) in patients receiving CDK4/6i combination therapy :

Considering a sample size of 1,231 patients receiving CDK4/6i with a frequency of gBRCA2 mutation of 3.5%. There were 965 events noted during this retrospective study. Therefore, our study had 97% power to detect a HR of 2.0 at alpha level of 0.05.

Considering a a sample size of 359 patients receiving first line CDK4/6i with a frequency of RB1 het loss of 30%. There were 260 events noted during this retrospective study. Therefore, our study had 99% power to detect a HR of 1.8 at alpha level of 0.05.

Sample size for PDX experiments was calculated based on previous experience with this model and drug response. No statistical method was used to predetermine sample size.

Data exclusions

Patients were excluded from the initial germline-somatic sequencing cohort for the following reasons:
i) Sequencing was obtained for external consultation; no pathology reports or clinical reports provided
ii) The only tissue specimen sequenced was a ductal carcinoma in situ
iii) Multiple samples which appeared to be evolutionarily unrelated, with discordant receptor status

From the clinical CDK4/6i and ET analysis, samples were additionally excluded for the following reasons:
i) Receptor status TNBC or HER2+
ii) Cases receiving CDK4/6i for fewer than 2 weeks
iiI) Oligometastatic tumors for which the only visible metastatic sites underwent local ablation with radiation therapy or surgery prior to initiation of CDK4/6i.

For analysis of allele-specific copy number analysis and clinical outcome, cases were excluded due to low purity or other technical limitations (such as “waterfalling” artifact), or paucity of heterozygous SNPs allowing for confident lesser copy number inference.

Replication

We sought to replicate our finding of germline BRCA2 pathogenic variant predisposing to more rapid progression on CDK4/6i and endocrine therapy by utilizing the Flatiron Health database. This is an independent, nationwide clinicogenomic dataset containing manually curated patient-level outcomes data from both community oncology settings and academic medical centers (n = 2,185 patients with HR+/HER2- metastatic breast cancer treated with first line CDK4/6i and endocrine therapy.

Lastly, we sought to validate our observation that germline BRCA2-associated breast cancers frequently harbor co-occuring loss of heterozygosity (LOH) of both RB1 and BRCA2. This was accomplished with whole exome sequencing from patients at University of Penn/ Abramson Cancer Center/ Basser Center as well as Mayo clinic.

We next sought to replicate our finding that RB1 LOH status predisposes to more rapid progression and decreased overall survival on CDK4/6i and endocrine therapy. We performed analysis of baseline plasma cell free DNA samples from PALOMA-3, the pivotal randomized phase III trial of palbociclib plus fulvestrant versus fulvestrant monotherapy in patients with HR+/HER2- MBC. RB1 LOH in the PALOMA-3 cohort was associated with a significantly shorter PFS, as detailed in our manuscript.

To replicate PDX experiments, several different germline BRCA2 models were utilized (PDX-L vs PDX-R, PDX-P1 and PDX-P2, as well as PDX-C1), with consistent findings demonstrating CDK4/6i resistance. Each Western blot analysis (PDX-C, PDX-L, and PDX-R) was repeated three times , with consistent findings.

Randomization

For clinicogenomic analysis, randomization was not applicable as this was not an intervention-based study. During the PDX experiments, mice were randomly assigned to the treatment arms.

Blinding

For clinicogenomic analysis, blinding was not applicable to this non-interventional study. All analyses were performed on pre-existing, de-identified genomic datasets obtained from independent cohorts. The investigators were not blinded to allocation during experiments and outcome assessment during the PDX experiments.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☐	☒ Animals and other organisms
☐	☒ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	The following lists the antibodies used as described in the Methods in their respective sections: rabbit anti-BRCA2 (Abcam; cat# AB123491; lot# 1002818-19; clone N/A ; 1:1000); rabbit anti-Phospho-Rb Ser780 (Cell Signaling Technology; cat# 8180s; clone D59B7; lot# 7; 1:1000); rabbit anti-Rb (Cell Signaling Technology; cat# 9313s; clone D20; lot# 7; 1:1000); mouse anti-Rb (Cell Signaling Technology; cat# 9309s; clone 4H1; lot# 18; 1:1000); rabbit anti-beta-Actin (Cell Signaling Technology; cat# 4970s; clone 13E5; lot# 19; 1:1000); IRDye® 680RD Goat anti-Mouse IgG Secondary Antibody (LICORbio; cat# 926-68070; lot# D41105-03; 1:5000); IRDye® 800CW Goat anti-Rabbit IgG Secondary Antibody (LICORbio; cat# 926-32211; lot# D50805-03; 1:5000).
Validation	We used commercial antibodies which were validated by the manufacturer. Validation can be found on the manufacturer's website: abcam.com, cellsignal.com, licorbio.com. Validation of Abcam anti-BRCA2 was performed internally through immunoblotting of lysates from cells with or without BRCA2.

Animals and other research organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research, and Sex and Gender in Research

Laboratory animals	For PDX-P1 (ST4316B) and PDX-P2 (HBCx-22): 6- to 12-week-old Female athymic Nude animals. PDX-P1: Mice are acclimated for a minimum of 24 hours and housed on irradiated corncob bedding (Teklad) in individual HEPA ventilated cages (Sealsafe® Plus, Techniplast USA) on a 12-hour light-dark cycle at 21-23°C and 40-60% humidity. Animals are fed water ad libitum (reverse osmosis, 2 ppm Cl2) and an irradiated standard rodent diet (Teklad). P DX-P2: mice are delivered to the facility at least 7 days before the experiment for acclimatizing to environmental conditions. Mice are housed in Polysulfone plastic (PSU) individually ventilated cages (IVC) (mm 213 W x 362 D x 185 H, Allentown, USA) bedded with sterilized and dust-free bedding cobs. Animals have controlled light-dark cycle (14-hour circadian cycle of artificial light) at 20°C - 24° C and 40% - 75% humidity. Each mouse is offered a complete pellet diet (150-SP-25, SAFE) and filtered, sterilized tap water ad libitum throughout the study. For PDX-L and PDX-R: 6 week old female NSG mice. Animals were maintained in accordance with the Guide for the Care and Use of Laboratory Animals in an AAALAC-accredited facility. All procedures outlined in the study were approved by the Memorial Sloan-Kettering Cancer Center IACUC. Animals were housed in individually ventilated caging systems (Thoren Caging Systems, Hazleton, PA), on autoclaved aspen chip bedding (PJ Murphy Forest Products, Montville, NJ) and were provided a γ-irradiated commercial diet (PicoLab Rodent Diet 20, 5053 LabDiet, PMI Nutrition International, St Louis, MO), and acidified water (pH 2.5 to 2.8) ad libitum. Mice were housed at a population density that ranged from 1 to 5 mice per cage in an environment providing a temperature of 21.1 to 22.2 °C (70 to 72 °F), 30% to 70% humidity, 10 to 15 fresh air exchanges hourly, and a 12:12-h light:dark cycle (lights on, 0600 to 1800). For PDX-C: 8 week old Swiss nude mice. Their care and housing were in accordance with institutional guidelines and the rules of the French Ethics Committee: CEEA-IC (Comité d'Ethique en matière d'expérimentation animale de l'Institut Curie, National registration number: #118). The project authorisation no. is 02163.02. The housing facility was kept at 22 °C (±2 °C) with a relative humidity of 30–70%. The light–dark cycle was 12 h light/12 h dark.
Wild animals	No wild animals were used in the study
Reporting on sex	The PDX models used in this study were established from tumors derived from female patients with breast cancer. All engraftments were performed in female immunodeficient mice to maintain hormonal context relevant to HR+ disease. Sex was therefore controlled consistently across experiments. While sex-based biological differences were not a variable under study, we recognize their importance and have minimized confounding by using female hosts throughout.
Field-collected samples	No field collected samples were used in the study
Ethics oversight	Studies with PDX-L and PDX-R were conducted through the MSK antitumor core facility in compliance with institutional guidelines under an IACUC approved protocol (MSK IRB 12-10-016). Animals were maintained in accordance with the Guide for the Care and Use of Laboratory Animals in an AAALAC-accredited facility. All procedures outlined in the study were approved by the Memorial Sloan-Kettering Cancer Center IACUC Studies with PDX-P1 (ST4316B) were performed under contract with XenoStart at AAALAC-accredited facilities and performed in accordance with protocols approved by the START 'Institutional Animal Care and Use Committee' (IACUC) and AstraZeneca's 'Platform for Animal Research Tracking and External Relationships' (PARTNER) group Studies with PDX-P2 (HBCx-22) were performed under contract with Xentech under authorization by the 'Direction Départementale de la Protection des Populations, Ministère de l'Agriculture et de l'Alimentation', France and in accordance with protocols approved by Xentech along with AstraZeneca's PARTNER group. For PDX-C: care and housing were in accordance with institutional guidelines and the rules of the French Ethics Committee: CEEA-IC (Comité d'Ethique en matière d'expérimentation animale de l'Institut Curie, National registration number: #118)

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Clinical data

Policy information about clinical studies

All manuscripts should comply with the ICMJE guidelines for publication of clinical research and a completed CONSORT checklist must be included with all submissions.

Clinical trial registration	NCT01775072.
Study protocol	The study protocol describing MSK-IMPACT is available at: https://clinicaltrials.gov/study/NCT01775072
Data collection	Participants were recruited from Memorial Sloan Kettering Cancer Center (New York, NY); the germline-somatic analysis included patients who underwent prospective clinical tumor sequencing as part of their clinical care (February 2014 to September 2021). For clinical outcome and matched pre and post CDK4/6i inhibitors, we expanded this cohort to include CDK4/6i outcomes and samples sequenced up to a data cut off date of April 2025. For the MSK cohort, clinical data were obtained from the validated Breast Translational Program's clinical database, which contains comprehensive, longitudinal information on patient demographics, pathology, treatment regimens, and clinical outcomes. Structured data were systematically curated using standardized clinical data annotation processes by a dedicated team of expert clinical data annotators and were maintained under rigorous quality-control procedures by experienced data managers. The data lock for the analysis was April 2025.
Outcomes	The primary outcome - to elucidate the somatic enrichments and mutual exclusivities of breast cancers from germline variants in the HRD pathway - was pre-defined during the initial study design phase. This was assessed by comprehensive analysis of germline variants, as well as somatic mutations, copy number alterations and structural variants. This is further delineated in the data analysis section. The secondary outcome - to investigate the associations between germline-somatic traits and clinical outcomes was predefined as statistical analysis correlating genomic metrics corresponding to top germline-somatic findings (e.g. germline status, allele-specific copy number) with progression -free and overall survival. This was assessed using Kaplan-Meier survival analysis, multivariate Cox regression, and correlation tests, with validation in external datasets (FLATIRON, PALOMA-3), where available. Specific definitions for the clinical outcome analysis are noted below: Progression events were defined as i) a radiographic or clinical disease progression prompting change in systemic therapy or recommendation for ablative local therapy directed at site(s) of progressive disease; or ii) clinician assessment detailing radiographic and/or clinical progression, after which it was documented that patient and physician decided to continue same therapy post progression. In such cases, the time of progression was defined as the date of documented progression rather than the date of therapy discontinuation. We determined the association between genomic alterations and PFS with disease progression on therapy with CDK4/6i or patient death. Disease progression was defined as the date of the radiology study or clinical assessment that established progression of disease and prompted a change in systemic treatment, intervention with locally directed therapy (e.g. radiation therapy), or otherwise an annotation in the chart documenting progression of disease. We categorized CDK4/6i regimens based on their ET partner (aromatase inhibitor vs. selective estrogen receptor degrader/ SERD). Patients with ablation of only known sites of disease with radiotherapy or surgical resection prior to initiation of CDK4/6i therapy were excluded, as were patients who discontinued therapy due to toxicity within two weeks. When ET or CDK4/6i changed to another ET and CDK4/6i, respectively, for reasons other than disease progression (e.g. toxicity, patient preference or insurance coverage), the time on successive regimens was combined to more accurately capture real-world PFS on CDK4/6i + ET combination. We employed both univariate and multivariate Cox proportional hazard models (adjusted for ET partner [i.e. fulvestrant vs. aromatase inhibitors], and treatment line, where applicable). For patients with multiple lines of therapy from the same class of treatment, only the first treatment line from that class that was started after the MSK-IMPACT biopsy was included in the analysis. For analyses pertaining to ASCN, fraction genome altered and whole-genome duplication were employed as additional covariates. These are recognized poor prognostic factors and may be a confounding factor given increased likelihood for tumors with measures of copy number instability to harbor LOH of any specific region. For overall survival analysis, we implemented a left-truncated model to account for the immortal time from diagnosis of metastatic disease (time zero) to enrollment on sequencing protocol. Similarly to the univariate analyses, we employed univariate and multivariate Cox proportional hazard models. In addition to ET partner, age at metastatic diagnosis was also included as a covariate. We rejected the null hypotheses with a two-sided $\alpha = 0.05$. For the matched-pairs analysis, we included all patients with available paired pre- and post-CDK4/6i sequencing data. Pre-CDK4/6i samples consisted exclusively of tumor specimens sequenced using MSK-IMPACT with available ASCN and zygosity assessment for RB1. Post-treatment samples included post-progression CDK4/6i tumor specimens sequenced using MSK-IMPACT as well as ctDNA sequenced using either MSK-ACCESS or Guardant360. For analyses comparing RB1 heterozygous loss with other allelic configurations, we focused specifically on drivers of resistance rather than subclonal events. We therefore excluded post-treatment alterations where the variant allele frequency was less than 0.30 of the maximum allele frequency of high-confidence variants present in the particular sample of interest. For assessment of acquired tumor suppressor loss in the HRD vs. BRCA2 vs. non-HRD group, we considered tumor suppressor genes that have been implicated in CDK4/6i resistance (RB1, PTEN, LATS2, FAT1, TP53, ARID1A, LATS1, NF1). We excluded samples where the ASCN was not evaluable in all these genes of interest, or where there was already a biallelic loss of function of one of the genes predicted to confer immediate resistance (RB1, PTEN, NF1). We included patients with baseline TP53 loss, as it has been shown to facilitate cell cycle re-entry and is therefore associated with acquired resistance on an intermediate timescale. We defined pre-

treatment samples as either i) gBRCA2, ii) non-BRCA2 HRD (either harboring a germline variant in BRCA1, PALB2, or classified as HRD-positive by the HRD-IMPACT assay), versus iii) non-HRD. All clinical outcome analyses were conducted with R software version 4.5.1 and the survival and exact2x2 packages.

Plants

Seed stocks	NA
Novel plant genotypes	NA
Authentication	NA