Tracing back primed resistance in cancer via sister cells

Authors: Jun Dai[1,†], Shuyu Zheng[1,†], Matías M. Falco[1], Jie Bao[1], Johanna Eriksson[1], Sanna Pikkusaari[1], Sofia Forstén[1], Jing Jiang[1], Wenyu Wang[1], Luping Gao[1], Fernando Perez-Villatoro[1], Olli Dufva[2], Khalid Saeed[2], Yinyin Wang[1], Ali Amiryousefi[1], Anniina Färkkilä[1], Satu Mustjoki[2], Liisa Kauppi[1], Jing Tang[1]*, Anna Vähärautio[1]*

Affiliations:

[1] Research Program in Systems Oncology, Research Programs Unit, Faculty of Medicine, University of Helsinki, Helsinki, Finland.

[2] Research Program in Translational Immunology, Research Programs Unit, Faculty of Medicine, University of Helsinki, Helsinki, Finland.

† These authors contributed equally to this work.

* These authors jointly supervised this work.

[*] Corresponding authors: Email: jing.tang@helsinki.fi; anna.vaharautio@helsinki.fi

Abstract:

Exploring non-genetic evolution of cell states during cancer treatments has become attainable by recent advances in lineage-tracing methods. However, transcriptional changes that drive cells into resistant fates may be subtle, necessitating high resolution analysis. We developed ReSisTrace that uses shared transcriptomic features of synchronised sister cells to predict the states that prime treatment resistance. We applied ReSisTrace in ovarian cancer cells perturbed with olaparib, carboplatin or natural killer (NK) cells. The pre-resistant phenotypes were defined by cell cycle and proteostatic features, reflecting the traits enriched in the upcoming subclonal selection. Furthermore, DNA repair deficiency rendered cells susceptible to both DNA damaging agents and NK killing in a context-dependent manner. Finally, we leveraged the pre-resistance profiles to predict and validate small molecules driving cells to sensitive states prior to treatment. In summary, ReSisTrace resolves pre-existing transcriptional features of treatment vulnerability, facilitating both molecular patient stratification and discovery of synergistic pre-sensitizing therapies.

Main

Recent studies have highlighted the role of non-genetic heterogeneity driving cancer treatment resistance[1–4]. Mechanistically, such heterogeneity accumulates continuously via inherent stochasticity in gene expression, and is further strengthened upon each cell division due to uneven partitioning of biomolecules and organelles[5]. As a result, even isogenic cancer cells exhibit variable drug responses, associated with transient changes in their chromatin or expression states before the treatment[1–4]. Multiple methodologies have been developed to explore these pre-existing resistant states, involving physical separation of cell populations[6] or predefined markers[7]. More recently, labelling lineages by random barcodes that are readable by single-cell RNA-sequencing (scRNA-seq) has enabled a fate-coupled analysis of full transcriptomic profiles. They have been applied to study induced reprogramming[8], hematopoietic differentiation[9] and lately also cancer resistance. For example, ClonMapper allows physical separation based on barcode hybridisation with scRNA-seq readout[10]. However, ClonMapper analyses fate-coupled transcriptomes on very large progenies, and is thus optimal in finding only the structures with clear and stable transcriptional differences. The Watermelon system further couples lineage labels to proliferation rate but was able to determine specific transcriptomics patterns associated with resistant fates only during, not prior to the drug treatment[11]. Furthermore, the published studies do not strictly control the number of cell divisions, therefore leading to large transcriptional divergence within the progenies, nor do they estimate the level of this divergence for distinct transcripts and thus are unable to assess the accuracy of fate-associated expression patterns across the transcriptome.

We addressed these limitations by developing ReSisTrace, which leverages the transcriptional similarity of sister cells to enable a high-resolution prediction of primed resistance. ReSisTrace also allows exploring asymmetric transcriptomic features between sister cells, which accumulate phenotypic heterogeneity upon each generation of cancer cells. We further analysed primed resistance in the subclonal context to address the interplay between genetic and non-genetic traits. We applied ReSisTrace against chemotherapy, targeted therapy, and immune cytotoxicity, all in the context of high-grade serous ovarian cancer (HGSOC). The improved resolution of ReSisTrace enables the detection and targeting of distinct transcriptomic patterns that precede treatment resistance, allowing us to discover cell states with increased vulnerability to each applied treatment, as well as novel pre-sensitizing agents that drive cells to these states.

Results

ReSisTrace reveals primed resistance via sister cell inference

To reveal the cell states that are primed to treatment resistance, uniquely labelled cells are synchronised and allowed to divide once, after which half of the cells are analysed by scRNA-seq, while the other half undergoes anti-cancer treatment (Fig. 1a). Surviving cells were allowed to recover, and then analysed by scRNA-seq to identify resistant lineages and their

gene expression profiles. We determined the transcriptomes of these resistant lineages at the pre-treatment stage based on sister-cell coupling, and annotated them as pre-resistant cells. For simplicity, cells with labels missing from post-treatment samples were annotated as 'pre-sensitive' whilst they inherently also contain false negative pre-resistant cells due to inevitable cell loss.

We applied ReSisTrace in a HGSOC cell line Kuramochi[12] to address primed resistance against carboplatin chemotherapy, the PARP inhibitor olaparib, and anti-tumor immunity represented by natural killer (NK) cells. Carboplatin is part of standard-of-care for HGSOC patients, while PARP inhibitors are used to treat patients with homologous recombination deficiency (HRD)[13,14]. Kuramochi harbours a heterozygous *BRCA2* mutation and is functionally HRD[15,16], although showing lower sensitivity to PARP inhibitors than homozygous *BRCA1/2* mutant cell lines[16]. NK cells are one of the key players in anti-cancer immunity[17], yet - unlike CD8+ T cells - do not need to be autologous. We also included a non-treatment control that mimicked the splitting, re-plating, and growth conditions of the drug treatment experiments, allowing us to deconvolute the confounding effect of experiment-specific growth fitness from primed drug resistance.

Cells were labelled uniquely by using lentiviral constructs with random barcodes of 20 bases, and then sub-sampled from a large pool, analogous to the concepts used in unique molecular identifiers (UMIs)[18]. The ReSisTrace construct was incorporated to the Perturb-seq[19] vector backbone (Fig. 1a), achieving high detection efficiency both before and after the treatments, in approximately 90% of cells (Fig. S1a). On average, 87.49% of the barcodes were unique in each pre-treatment sample (Fig. S1b). As expected, the majority of colonies before splitting the samples consisted of two cells (Fig. S1c), whereas after splitting, most lineages were represented by only one cell (Fig. S1d). Mathematical modelling and computer simulations showed that more than 80% of pre-sensitive cells captured by ReSisTrace were truly positive (Fig. S1, e and f, and Table S1) at the killing rate selected for the experiment (Fig. S1g). This suggests that the true pre-resistant cells can be captured with four to five-fold enrichment in the observed pre-resistant population compared to the observed pre-sensitive population, despite experimental uncertainties such as cell loss. Furthermore, we performed all the assays in two replicates, achieving highly consistent transcriptomic changes between the pre-resistant and pre-sensitive populations (Fig. S1h). In the non-treatment control experiment, robust transcriptomic changes were also observed between pre-surviving and pre-extinct cells, suggesting that the experimental procedures themselves pose a non-random selection pressure which should be corrected.

To validate the use of sister cells as proxies, we confirmed that cells with shared labels, i.e. putative sister cells, have significantly more similar transcriptomes than random pairs of cells (Fig. 1b). We further compared the genes that showed significantly higher similarity within sister cell pairs, termed 'sister-concordant genes', against the other genes, termed 'sister-discordant genes' (Fig. 1c). The sister-discordant genes were expressed at lower levels, along with increased drop-out noise ($P < 2.2 \times 10^{-16}$, two-tailed t-test) (Fig. S2a). Sister-discordant genes also showed higher relative degradation rates ($P = 4.2 \times 10^{-10}$, two-tailed t-test) and

slightly higher splicing rates relative to transcription (P = 0.000038, two-tailed t-test) (Fig. 1d and Fig. S2, b and c), suggesting that transcripts to be degraded show higher expressional variation. While sister-concordant genes were not enriched in any pathway, the sister-discordant genes were highly enriched in Inositol phosphate metabolism and Phosphatidylinositol signalling system (Fig. 1e). These pathways provide substrates for Phosphatidylinositol 3-kinase/AKT/mTOR signalling, which has been shown to drive asymmetric cell division in divergent contexts including cancer[20–22].

Subclonal enrichment contributes to pre-resistance signatures

We next explored the structure of cell populations in the transcriptomic space. A UMAP embedding showed no distinct separation between the pre-resistant and pre-sensitive cells, while the post-treatment cells were shifted from the pre-treatment cells (Fig. 2a and Fig. S3a). As expected, all the post-treatment samples showed larger clone sizes (Fig. S3b) ($P < 2.2\times10^{-16}$, two-tailed Kolmogorov-Smirnov test) and smaller lineage diversity compared to the pre-treatment samples (Fig. S3c) ($P = 3.80 \times 10^{-5}$, paired t-test, n = 8). To assess the putative role of genetic heterogeneity, we inferred copy number variations (CNVs) from the scRNA-seq data, and identified five subclones that closely matched the clusters in the UMAP (Fig. 2b and Fig. S4a). The subclonal identities of lineages were stable during the treatments (Fig. S4b), resulting in similar subclonal proportions between pre-resistant and post-treatment populations (Fig. 2c). These results suggested that subclonal selection is already evident in the pre-resistant cells. Specifically, we found that subclone A was significantly enriched in the pre-resistant cells compared to pre-sensitive cells at all the treatment conditions, especially for the carboplatin and olaparib treatments ($P = 8.882 \times 10^{-11}$ and $P < 2.2 \times 10^{-16}$, respectively, two-proportion Z-test) (Fig. 2c).

With the identity of pre-resistant and pre-sensitive cells, as well as the identity of subclones, we determined rank-based transcriptomic signatures for pre-resistance and subclones (Table S2). We found that the signature of subclone A was positively correlated with the pre-resistance signatures (Fig. 2d), while signatures of subclones B, C or D correlated negatively. Subclone E showed a treatment-specific pattern, consistent with the enrichment patterns of the subclones in the pre-resistant populations (Fig. 2c and d). Gene ontology (GO) enrichment analysis revealed that changes in proteostasis were associated with primed resistance against olaparib and carboplatin and enriched in subclone A (Fig. 2e, Table S3). Response to topologically incorrect protein was associated with primed sensitivity towards olaparib or carboplatin, in line with DNA damaging therapies potentiating endoplasmic reticulum (ER) stress to lethal levels[23]. We also found that mitochondrial expression patterns were associated with primed sensitivity towards both olaparib and NK, consistent with recent studies[24,25]. Interestingly, although sensitivity to NK killing was also associated with mitosis[26], the cell cycle associations of the two DNA-damaging therapies diverged. Both mitosis and DNA replication were enriched in primed sensitivity against carboplatin that crosslinks DNA regardless of cell cycle phase[27]. In contrast, olaparib damages DNA by stalling replication forks, where the immediate damage itself needs to occur in the S phase[28]. Therefore, only concurrent DNA replication, not mitosis, was associated with primed olaparib sensitivity. Together, our results show that subtle

pre-resistant features can reflect the subclonal phenotypes that are subsequently enriched during each treatment.

HRD sensitises ovarian cancer cells to NK killing in a context-dependent manner

HRD is the key tumour phenotype that predicts clinical response to PARP inhibitors and carboplatin in HGSOC. As the capability to repair DNA can vary even within genetically homogeneous cell populations[29], we hypothesised that cells with HRD-like states could be primed for sensitivity to olaparib and carboplatin. To test the hypothesis, we first determined an HRD-associated transcriptional signature from an isogenic cell line pair model, constructed by comparing the gene expression of parental *BRCA1* mutant COV362 cells to the *BRCA1*-restored cells (Fig. 3a). We confirmed that the restored cells displayed a homologous recombination proficient (HRP) phenotype with increased RAD51 foci after DNA damage, resulting in reduced olaparib sensitivity (Fig. 3b and Fig. S5a). Kaplan-Meier survival analysis confirmed that HRD signature was associated with improved overall survival (OS) in ovarian serous cystadenocarcinoma patients in The Cancer Genome Atlas (TCGA) cohort (Fig. S5b).

We then compared the HRD signature to the pre-sensitivity signatures (i.e. the inverse of pre-resistance signatures) captured by ReSisTrace. We confirmed that the pre-sensitivity signatures in olaparib and carboplatin conditions were positively associated to the HRD signature, as expected based on their mechanisms of action that target HRD (Fig. 3c; association estimated by the connectivity scores defined in[30]). More interestingly, the pre-sensitivity signature for the NK condition showed even higher connectivity scores to the signature (Fig. 3c). This unexpected finding prompted us to explore a causal connection between HRD and sensitivity to NK treatment in ovarian cancer cells. Indeed, we found that the *BRCA2* restored Kuramochi cells (Fig. S5c) were consistently more resistant to NK when compared to cells transduced with control plasmid (Fig. 3d). Furthermore, the other two isogenic BRCA1 restored cell lines (COV362 and UWB1.289) also showed increased resistance against the NK cells from the same donor used in the ReSisTrace experiment (Fig. S5c). However, the responses to NK cells from other donors did not recapitulate this, suggesting donor-specificity of the NK pre-resistance signature (Fig. S5d). To investigate the potential of the NK treatment at a population level, we stratified TCGA ovarian cancer cohort[31] into HRD and HRP groups by a genomic classifier called ovaHRDscar[32], and compared their deconvoluted[33] gene expression profiles. We found that the NK pre-sensitivity signature was positively associated with the clinical HRD signature (connectivity score: 0.38 +/- 0.03). Furthermore, we used known NK cell markers to construct an NK signature that indicates presence of NK cells in clinical specimens. Indeed, the NK cell presence was associated with improved survival only in the HRD patients (P = 0.018, Cox proportional hazard model), but not in the HRP patients (Fig. 3e and Table S5). These results suggest that HRD tumours might be more responsive to innate immune cytotoxicity in the autologous context.

Targeting treatment-specific primed resistance with small molecules

We hypothesised that small molecules that shift cancer cells toward pre-sensitive states could sensitise them to the corresponding treatments. We searched the L1000 database[34] for such drugs that induce gene expression changes opposite to the pre-resistance signatures (Fig. 4a,

upper panel). Top drugs were ranked by the negative connectivity scores, which were further prioritised by a consensus *P*-value threshold derived from multiple variants of the pre-resistance signatures (Fig. 4b and Table S6). Drugs with non- or lowly expressed targets in Kuramochi cells were excluded, resulting in ten drugs for experimental validation (Table S7).

We performed a two-step functional validation for the predicted drugs. As transcriptional responses to many drug classes are cell-type specific[35], we first assessed whether the drugs indeed induced pre-sensitive signatures also in the context of interest, i.e. Kuramochi cells, and then measured their ability to sensitise these cells to either carboplatin, olaparib or NK cells, when used prior to the main treatment (Fig 4a, lower panels). For the first step, we generated drug-induced transcriptome profiles in Kuramochi cells through scRNA-seq. The UMAP projection shows that three proteasome inhibitors (bortezomib, delanzomib, and ixazomib) induced transcriptome profiles that were clearly clustered together and distinguished from the DMSO control (Fig. 4c, Fig. S6). In addition, pevonedistat, a neddylation inhibitor[36], when applied in two distinct concentrations, clearly shifted the transcriptional profile of Kuramochi cells to a unique state. We found that all the predicted drugs indeed induced changes opposite to the corresponding pre-resistance signatures, with significant negative connectivity scores (Fig. 4c, Fig. S6). For example, the top predicted pre-sensitisers for NK cells include pevonedistat, GW-843682X and nocodazole, all inducing transcriptomic changes with highly negative connectivity to NK pre-resistance, while pyrrolidine-dithiocarbamate and sitagliptin perturbed profiles had the weakest association with olaparib pre-sensitivity profiles .

As the second validation step, we assessed whether the predicted drugs indeed sensitised cells to the treatment in a synergistic manner. We pre-treated the Kuramochi cells with the predicted drugs, changed the media to remove the drugs, and then perturbed the cells with carboplatin, olaparib, or NK cells. The predictions were highly accurate: of the 12 predicted drugs, 11 showed synergy according to multiple reference models (Bliss, HSA, Loewe, or ZIP)[37,38], and eight showed synergy consistently by all of the reference models (Fig. 4c and Fig. S7). Importantly, these drugs were confirmed to induce pre-sensitive-like transcriptomics changes in Kuramochi cells, thus validating that the pre-resistance signatures identified by ReSisTrace indeed have functional significance. In particular, pevonedistat, ixazomib citrate, and GW-843682X were validated as the top synergistic drugs for carboplatin, olaparib, and NK cells, respectively (Fig. 4d). Furthermore, we confirmed that pevonedistat - that induced an expression profile highly similar to all three pre-sensitive states - strongly pre-sensitized Kuramochi cells to all the three treatment modalities already at a low concentration (0.33 μM), despite having a minor inhibition on cell viability as a monotherapy (Fig. S7a). Pevonedistat inhibits neddylation, a ubiquitin-like post-translational modification affecting a wide range of cellular functions, such as mRNA splicing and surveillance, DNA replication, and proteostasis[36]. Pevonedistat has shown synergy with multiple treatment modalities in preclinical models, and is currently undergoing several clinical trials as a combination therapy in both haematological and solid cancers[39,40]. Ixazomib citrate, together with other three proteasome inhibitors (delanzomib, carfilzomib, and bortezomib) synergized with olaparib, in line with the increased response to topologically incorrect protein in olaparib pre-sensitive cells (Fig. 2e). In contrast, the NF-κB inhibitor pyrrolidine-dithiocarbamate was less synergistic with

olaparib, whereas sitagliptin used to treat type II diabetes had no effect, in line with its failure to induce expression changes that resemble those of olaparib pre-sensitive cells. For the NK treatment, two drugs that cause mitotic arrest were identified, including the PLK1/3 inhibitor GW-843682X and the tubulin polymerisation inhibitor nocodazole, corroborating both the NK pre-sensitive-like expression changes they induced (Fig. 4c, S6) and gene ontology results (Fig. 2e). In addition, the purine nucleoside antimetabolite clofarabine was synergistic, suggesting that the cell cycle phase of cancer cells modulates their vulnerability to NK cells. Taken together, the drug signature analysis based on ReSisTrace enables a systematic and accurate prediction of sensitising small molecules that reverse the pre-resistant states to induce synergistic interactions with the target treatment.

Discussion

Fates of cancer cells with recently shared ancestry are closely intertwined, showing aligned drug responses[41] and preferential site of metastasis[42,43]. Cells of the same progeny, at least until a couple of generations apart, also present similar gene expression profiles[44,9]. However, transcriptional patterns of pre-treatment cells do not fully capture their future fate[9], and cells with extreme fates, such as high-dose drug persistence, may not display evident gene expression features prior treatment[11]. Compared to previous lineage tracing methods that allow more than one cell division after labelling, ReSisTrace has the advantage in identifying subtle and transient fate-coupled differences in gene expression patterns that are increasingly obscured by stochastic noise upon each cell division. Live cell tracking has revealed that divergence in protein expression significantly increases already from sister cells, i.e. two cell progenies, to cousins or four cell progenies[6]. However, a larger progeny size[11], and even clonal isolation[10] of pre-resistant populations are beneficial when characterising primed states of extremely rare persisters in order to maximise the odds of detecting each pre-resistant lineage. Thus, the selection of optimal lineage tracing method depends on the context, and is a trade-off between sensitivity to detect and analyse the rare progenitors of extreme perisister lineages, and sensitivity to identify more transient expression patterns.

Non-genetic heterogeneity should ideally be assayed in the context of genetic heterogeneity, as clinical tumours are never completely isogenic due their inherent genomic instability. To improve clinical relevance of ReSisTrace in HGSOC that is an extremely copy number unstable cancer[45], we analysed primed resistance within the subclonal context via scRNA-seq inferred copy number variations. We found that subclonal structures drove global transcriptional clustering, and discovered treatment-specific enrichment of subclones. These enrichment patterns explained a considerable part of the transcriptomic features of the pre-resistant populations, such as those representing changes in proteostasis, mitotic nuclear division, and chromatin conformation. It is important to consider that subclonal enrichment contributes to differential expression of not only CNV affected drivers but also the co-amplified passenger genes[46], thus providing an additional level of noise to the pre-resistance signal and supporting the analysis of pre-resistant states in their subclonal context. Thus, ReSisTrace revealed the interplay between genetic and phenotypic heterogeneity in primed resistance, an aspect that becomes increasingly relevant for patient-derived models such as organoids or xenografts.

DNA damage is known to induce NK-mediated cell killing[47] and aneuploidy activates NK clearance of non-transformed senescent cells via NF-κB signaling[48]. However, the effect of HRD on NK cytotoxicity has not been reported. In Kuramochi cells, we identified a connection between subtle intra-population differences in HRD state and sensitivity to NK killing, which was validated by showing increased resistance upon BRCA2 restoration against NK cells from four donors. However, other NK donor combinations did not show aligned responses in additional BRCA1 restored isogenic cell line pairs, and such discrepancy was not explained by differences in HLA-C haplotype matches. This suggests that primed resistance against innate immune cells can be very context-specific. Interestingly though, the presence of NK derived transcripts associated positively with ovarian cancer patient survival only in an HRD dependent manner, suggesting that in the native, autologous context, HRD patients could be more susceptible to NK based anti-tumor therapies than patients with HRP tumours. Taken together, the putative connection between innate immunity and HRD provides an intriguing extension to the body of evidence showing that HRD tumours are more susceptible to adaptive immunity, with improved T cell-mediated immunosurveillance[49].

Drug resistance driven by non-genetic heterogeneity is mechanistically complex. Functional genetic screens can reveal causal factors but are largely limited to mechanisms that rely on individual genes. Longitudinal transcriptomic comparisons can capture complex drug-induced patterns but do not allow detecting the evolutionary selection and enrichment of pre-existing resistant states during the treatment. On the contrary, ReSisTrace allows characterization of resistant populations already prior to the treatment, thus revealing also the contrasting vulnerable states that could be induced via pre-treatments. We used drug-perturbed gene expression signatures[34] to predict pre-sensitising drugs, and validated that context-specific induction of pre-sensitive states was a prerequisite for effective pre-treatments. We thus leveraged the plasticity of cell states by identifying sequential treatments, providing an alternative paradigm to simultaneous combinatorial therapy predictions that aim to kill the maximal proportion of malignant cells. Importantly, sequential treatments can mitigate many problems of simultaneous combinatorial therapies - such as elevated toxicities and adverse drug-drug interactions[50,51] - but have thus far been largely neglected in drug prediction approaches . In summary, sister-resolution lineage tracing can reveal complex cell states that prime treatment resistance in cancer, and allows opposing these states to overcome drug resistance already before it emerges.

Methods

Cell culture

COV362 cells (Merck, 07071910) were grown in DMEM (Corning, 15-013-CV) with 10% fetal bovine serum (FBS) (Gibco, 10270106), 1% penicillin–streptomycin (PS) (Gibco, 15140122), and 1% GlutaMAX (Gibco, 35050038). Kuramochi (JCRB Cell Bank, JCRB0098) cells were grown in RPMI1640 (Corning, 10-040-CV) with 10% FBS, and 1% PS. UWB1.289 (ATCC, CRL-2945) and UWB1.289+BRCA1 (ATCC, CRL-2946) cells were cultured according to the handling procedure from ATCC.

Small molecules

Pevonedistat hydrochloride (HY-10484), Ixazomib citrate (HY-10452), Delanzomib (HY-10454), Bortezomib (HY-10227), Carfilzomib (HY-10455), Pyrrolidinedithiocarbamate ammonium (HY-18738), Sitagliptin (HY-13749), GW843682X (HY-11003), Clofarabine (HY-A0005), Nocodazole (HY-13520) were all purchased from MedChemExpress.

Expansion of NK cells

NK cells were expanded by using K562-mbIL21-41BBL feeder cells as described previously[52]. Briefly, peripheral blood mononuclear cells (PBMCs) were isolated via gradient centrifugation with Ficoll-Paque from buffy coats of healthy donors. Five million PBMCs were suspended together with ten million feeder cells irradiated with 100 Gy, into 40 ml of R10 media supplemented with 10 ng/ml recombinant human IL-2 (R&D Systems, 202-IL-050). Cells were passaged twice per week and feeder cells were added in a 1:1 ratio after 7 days to maintain stimulation. After two weeks of culture, NK cells were purified with the NK Cell Isolation Kit (Miltenyi) and live-frozen. For the experiments, NK cells from four donors were used: donor 1 NK cells for ReSisTrace experiment, all donors' NK cells for BRCA1/2 restored cell line pairs, and NK cells from donor 2 to test predicted pre-sensitising drugs. All NK cells were cultured for 3 days between thawing and the experiments.

Synthesis of lineage barcodes library

The pBA439_UMI20 plasmid was constructed by ssDNA synthesis and NEBuilder® assembly tool. The lineage barcodes part ssDNA_UMI20 (5'-CTGGGGCACAAGCTTAATTAAGAATTCANNNNTGNNNNACNNNNGANNNNGTNNNNCTAGGGCCTAGAGGGCCCGTTTAAAC-3') was ordered from Integrated DNA Technologies. pBA439 was a gift from Jonathan Weissman (Addgene, plasmid # 85967)[19]. pBA439 vector was digested with EcoRI-HF (NEB, R3101S) and AvrII (NEB, R0174S), then the larger linear fragment pBA439_linear was recycled and purified by columns (Macherey-Nagel, 740609.50). The ssDNA_UMI20 was assembled with pBA439_linear to construct a pBA439_UMI20 library by the NEBuilder HiFi DNA assembly (NEB, E2621L). For amplification, the pBA439_UMI20 library was electroporated into 25 µl of Lucigen Endura ElectroCompetent Cells (Lucigen, 60242-2) using prechilled 1 mm electroporation cuvettes (BioRad, 1652089) in a BioRad GenePulser I machine set to 10 µF, 600 Ω, and 1800 V. Within seconds after the pulse, 1 ml of 37 °C Recovery Medium (Lucigen, 80026-1) was added and bacteria were grown in round-bottom tubes for 1 h at 37 °C while shaking at 180 r.p.m. Then, 1 ml of the bacterial culture was plated on a 25 × 25 cm bioassay plate containing LB medium with 100 µg/ml ampicillin. Plates were incubated at 32 °C for 22 h, then LB medium was added and cells were scraped off the plate. Bacterial cells were pelleted by 15 min of centrifugation at 5,000 RCF at 4 °C, and plasmid DNA was extracted with a Maxiprep endotoxin-free kit (Macherey-Nagel, 740424.10).

Lentivirus production and cell transduction

Lentiviral particles were produced by transfecting 293FT cells (Thermo Fisher, R70007) with the pBA439_UMI20 library, along with packaging constructs psPAX2 (Addgene, plasmid # 12260) and pMD2.G (Addgene, plasmid # 12259), gifts from Didier Trono. Constructs were

titred by serial dilution on the 293FT cells. The lentivirus was collected 48 h and 72 h after transfection and aliquoted to store at -80 °C, following filtering through a low-protein binding 0.45-µm filter.

Experimental BRCA +/- models

To clone the BRCA1/2 plasmids, lentiMPH v2 (Addgene, plasmid #89308) was modified by replacing Hygro selection marker with GFP to generate lentiMPH-GFP. *BRCA1* was amplified from pDEST-mCherry-LacR-BRCA1 (Addgene, plasmid #71115), and *BRCA2* was amplified from pCIN BRCA2 WT (Addgene, plasmid #16245). PCR amplified products (templates and primers presented in Table S8) were assembled by the NEBuilder HiFi DNA assembly (NEB, E2621L) to generate lenti-EMPTY-GFP, lenti-BRCA1-GFP, and lenti-BRCA2-GFP. Lentivirus were made as described earlier to transduce either COV362 or Kuramochi cells to restore the *BRCA1* and *BRCA2* functions. lentiMPH v2 was a gift from Feng Zhang, pDEST-mCherry-LacR-BRCA1 was a gift from Daniel Durocher, pCIN BRCA2 WT was a gift from Mien-Chie Hung ([53,54,55]).

The COV362 model was validated with a colony formation assay, where COV362 control and *BRCA1* restored cells (10000 cells per well) were exposed to 0, 2.5 or 5 µM olaparib. Colonies were stained with crystal violet after 12 days of incubation.

ReSisTrace experimental workflow

Six million Kuramochi cells were transduced with the pBA439_UMI20 lentivirus library at MOI between 0.2 and 0.25, and 2 days after transduction 2 µg/ml puromycin were added for a seven day selection to remove unbarcoded cells. All barcoded cells were equally divided into 6 vials (6 million cells per vial), 4 of which were cryopreserved and 2 were used for the subsequent experiment. Cells were synchronised with 2mM thymidine (Merck, T1895), then after releasing and washing out of thymidine, 16,000 cells/well were seeded on 6-well plates. After 48 hours, 4 wells were used for cell counting, and 8 wells were used for the ReSisTrace experiment. For each of the 8 samples, half of the cells were loaded for scRNA-seq as pre-treatment samples, and the remaining half were seeded back for either normal culture (13 days) or treated with carboplatin (Selleckchem, S1215, 1.2 µM, 3 days), olaparib (Selleckchem, S1060, 1.2 µM, 7 days), and NK cells (26,000 cells, 1 day). The drug concentrations or the number of NK cells were chosen to kill 70-80% of cancer cells compared to untreated cultures (Fig. S1g). After the treatment, cells were recovered for 10, 6 and 7 days respectively, after which about 16,000 cells from each well were loaded for scRNA-seq as post-treatment samples.

Mathematical modelling and simulation of lineage tracing

For the actual ReSisTrace experiments, the protocol involves multiple steps starting from labelling the initial pool of cells to harvesting viable cells for scRNA-seq. Additional uncertainties may be introduced from e.g. unequal doubling, or cell drop-out due to quality control (Fig. S1e). Therefore, the lineage tracing process was simulated by considering the following parameters that may affect the experiment outcome, including 1) SampleSize - total number of cells at the start of experiment; 2) KillRateTotal - killing rate of the treatment; 3) LineagePropAfterDoubling - proportions of singletons, twins, and quadruplets after doubling;

4) NCellBT - number of cells with lineage labels in the pre-treatment sample which passed the quality control; 5) NCellAT - number of cells with lineage labels in the post-treatment sample which passed the quality control; 6) NCellPreSeq - number of cells after recovering from treatment; 7) NCellSeq - number of cells loaded to scRNA-seq.

To reflect the actual experimental setting up, prior knowledge was utilised to constrain these parameters, including 1) SampleSize to be estimated as 16000; 2) KillRateTotal to be estimated between 70% and 80% of cell growth inhibition; 3) LineagePropAfterDoubling to be estimated as [42.46%, 53.04%, 4.5%] for the proportions of singletons, twins, and quadruplets separately, based on the observations shown in Fig. S1c. 4) NCellBT, NCellAT, NCellPreSeq, and NCellSeq were estimated based on the results of scRNA-seq data as shown in Table S1.

The simulation was initiated with these estimated parameters, creating a scenario where lineage labels were randomly assigned to cells and then traced into the final scRNA-seq data in both pre-treatment and post-treatment samples. Pre-sensitive predictive rate was determined, defined as the proportion of true pre-sensitive cells in the predicted pre-sensitive cells. In contrast, the pre-resistant predictive rate, defined as the proportion of true pre-resistant cells in the predicted pre-resistant cells, is always 100%. Therefore, A higher pre-sensitive predictive rate suggested a higher quality of the experiment data for determining the gene expression signature of primed resistance. To obtain a theoretical upper bound of the pre-sensitive predictive rate, an ideal scenario with perfect doubling and zero drop-out during the scRNA-seq was also simulated. Standard deviations were determined using 100 iterations (Fig. S1f).

Drugs-induced transcriptome profiles by scRNA-seq

Kuramochi cells were seeded in 6-well plates at around 50% confluence overnight. Predicted small molecules were added for 24h in duplicates, with DMSO as control. Then cells were washed and detached from plates. Each sample was individually multiplexed with 3′ CellPlex Kit Set A (10x Genomics, PN-1000261). After labelling and washing, every 12 samples were pooled and then loaded for scRNA-seq experiments according to 10x Genomics instructions.

Single-cell RNA-sequencing library preparation

On the 10x Genomics platform, for samples from ReSisTrace experiments, we used the Chromium Next GEM Single Cell 3′ Kit v3.1 (PN-1000268), Chromium Next GEM Chip G Single Cell kit (PN-1000120), and Dual Index Kit TT Set A (PN-1000215) were used. For samples from isogenic BRCAness +/- COV362 cells, cells were multiplexed with 3’ CellPlex Kit Set A (PN-1000261), then 7000 cells were loaded for scRNA-seq. For library preparation, the 3’ Feature Barcode Kit (PN-1000262) and the Dual Index Kit NN Set A (PN-1000243) were additionally used. The resulting cDNA libraries were quantified by both Agilent BioAnalyzer (Agilent, 5067-4626) and q-PCR (Roche, KK4835), and sequenced by Illumina NovaSeq 6000.

NK cells mediated killing for BRCA+/- models

On day 1, 8000 cancer cells per well were seeded on 96-well plates, and NK cells were cultured with 10 ng/ml IL2. On day 2, NK cells were added into cancer cells in indicated ratios to cancer cell numbers. On day 3, after removing the medium and the NK cells that were in suspension,

the viability of adherent cancer cells was measured by the CellTiter-Glo® Luminescent Assay (Promega, G9242).

Cell viability assay

1500 cells/well were seeded in 96-well plates and then allowed to adhere to the plates. Indicated concentrations of olaparib were added for 4 to 6 days. Cell viability was determined using CellTiter-Glo® Luminescent Assay (Promega, G9242).

Homologous recombination repair functional assay

HGSOC cells were seeded on 96-well plates in 60% confluence/10000 cells per well in triplicates and allowed to attach to the bottom for ~12-24 h. The cells were irradiated with 5 Gy of ionising radiation to induce DNA damage, followed by 8 h recovery time before fixation in 2% PFA. Fixed cells were immunostained according to standard procedures with primary antibodies against RAD51 (ab133534, Abcam, diluted 1:1000) and cyclin A2 (GTX634420, GeneTex, diluted 1:500), followed by incubation with secondary antibodies (goat anti-mouse IgG-Alexa Fluor 488, A11029; or goat anti-mouse IgG-Alexa Fluor 568, A11004; and goat anti-rabbit IgG-Alexa Fluor 647, A21245; LifeTechnologies, diluted 1:1000). Nuclei were counterstained with Hoechst. Images were acquired with Opera Phenix High Content Screening System (PerkinElmer) at 40x magnification. Images were analysed with ImageJ software using custom macros. At least 100 cyclin A2 positive nuclei per each well were analysed. Nuclei with ≥5 RAD51 foci were considered as RAD51 positive. Percentage of cyclin A2 and RAD51 positive nuclei out of all cyclin A2 positive nuclei was calculated.

Drug combination screening and synergy scoring

1500 (for olaparib or carboplatin combination) or 4000 (for NK cells combination) Kuramochi cells/well were seeded in 96-well plates and then allowed to adhere to the plates. Indicated concentrations of small molecules were added for 24 h, then cells were refreshed with normal medium with indicated concentrations of either olaparib or carboplatin added for 5 days or with different ratios of NK cells for 24 h. At the end point, cell viability was determined using CellTiter-Glo® Luminescent Assay (Promega, G9242).

Dose-response matrices were analysed with SynergyFinder Plus (www.synergyfinder.org) (R package v3.4.5)[56]. Four reference models were utilised to access the synergy, including HSA (Highest single agency), Bliss (Bliss independence), Loewe (Loewe additivity), and ZIP (Zero interaction potency). As the reference models rely on different mathematical assumptions of synergy, we reported the results for all of them. The degree of interactions between two drugs derived from the HSA model was visualised in a synergy landscape over the dose matrices (Fig. 4d and Fig. S7). Significance of average synergy score was evaluated by bootstrapping test.

Lineage label mapping and processing

The lineage barcodes sequence (5'-CTGGGGCACAAGCTTAATTAAGAATTCANNNNTGNNNNACNNNNGANNNNGTNNNNCTAGGGCCTAGAGGGCCCGTTTAAAC-3') was added to the GDCh38.d1.vd1 reference genome with GENCODE v25 annotation, as an individual gene named

pBA439_UMI_20 according to the parental lentiviral library. The Cell Ranger software (v5.0.1)[57] was used to perform read alignment and UMI quantification using the modified reference.

Lineage label sequences for each cell were extracted from the "possorted_genome_bam.bam" file output by cellranger count command, by considering the sequences mapped to gene *pBA439_UMI_20* and the sub-pattern "CANNNNTGNNNNACNNNNGANNNNGTNNNNCT". To mitigate the effect of sequencing errors, the directional network-based method from UMI-tools (v1.0.1)[58] was applied on the lineage label sequences. The sequences that differ at a single base from the representative sequence were corrected. In addition, we removed the non-unique lineage label sequences expressed by more than four cells in the pre-treatment samples. We assigned the cells from each pre-treatment sample into two groups: 1) pre-resistant cells with lineage labels matching those of the corresponding post-treatment samples; 2) pre-sensitive cells assigned to lineages that were not detected after the treatment.

Preprocessing of the scRNA-seq data

We used the Seurat (v4.0.4)[59] to perform the data quality control, normalization, top variable gene selection, scaling, dimensionality reduction, and differentially expressed gene (DEG) analysis. For each ReSisTrace sample, the genes expressed in less than three cells were removed. Based on the distribution of the UMI counts, number of genes, and percentage of mitochondrial transcripts, we filtered out the cells using the thresholds indicated in Table S9.

We used the LogNormalize method to normalize gene counts, and the FindVariableFeatures function to find the top 2,000 variable genes. The FindMarkers function was used to determine the differentially expressed genes by Wilcoxon rank sum test between different groups. To minimize experimental noise from cases where both sisters were sampled in the scRNA-seq analysis before the treatment, we performed the differential analysis only on the cells lacking sisters in the same pre-treatment samples.

Sister cell similarity

We merged the pre-treatment data from all the experiments. Cells with matching labels were defined as sister cells, gaining 2939 lineages with sister cells. Lowly expressed genes whose total normalised expressions are less than 1.25 (10% quantile) were removed. The UMAP coordinates for visualisation of randomly selected sister cells were determined with the first 20 principal components of the top 2,000 variable genes. For extracting the sister concordant genes, we considered only the lineages with two sister cells, and removed the sister cells with the Euclidean distance of their transcriptomes larger than 14.9 (90% quantile). The genes with $\log_2$ fold changes ($\log_2$FCs) significantly lower in the sister cell pairs compared to random cell pairs were defined as the sister-concordant genes (two-tailed t-test with an FDR-adjusted *P* value threshold of 0.05).

RNA velocity analysis

We selected expression-matched sets of sister-concordant and sister-discordant genes. We generated a binned histogram for the sum of expression in the pooled samples, and selected

genes that belong to the bin with the most balanced sister-concordant and sister-discordant genes (the second-lowest bin).

We used the velocyto (v0.17.17)[60] to align and quantify the spliced and unspliced RNA for each gene. Genes with RNA counts (both unspliced and spliced) less than 20 were removed. The dynamical model from scVelo (v0.2.3)[61] was used to fit the gene-specific rates of RNA transcription, splicing, and degradation.

Shannon diversity

Shannon diversity of lineage labels for each sample was determined by the vegan R package (v 2.6.2)[62].

Pathway analysis

KEGG pathways for the expression matched sister-concordant and discordant genes were determined by Enrichr (https://maayanlab.cloud/Enrichr/)[63,64]. Gene set enrichment analysis (GSEA) function from the clusterProfiler package (3.18.1)[65] was used for the ordered fold changes (logarithmic scaled) extracted from the fitness corrected pre-resistance differential expression analysis to inspect the enriched gene ontology biological process terms.

Subclone analysis

CNV profiles for the individual cancer cells for each sample were obtained with InferCNV (1.7.1)[66] using a common set of stromal cells from previously published data[67] as references. The Leiden algorithm was used for determining the underlying subclonal populations based on the CNV profiles. Exceptionally for the Olaparib1 sample, the algorithm was not able to recognise the smallest subclone (subclone E), due to the small sample size. Therefore, the inferCNV analysis was jointly run with the other Olaparib sample for the subclonal label determination.

Pre-resistant gene expression signatures

After identifying the pre-resistant and pre-sensitive cells in the pre-treatment samples, fold changes of gene expressions were determined. The $\log_2$FC between pre-resistant and pre-sensitive cells for each gene was averaged across two replicates and then subtracted by the $\log_2$FC of pre-surviving versus pre-extinct cells from the non-treatment control samples. Only the sister-concordant genes were included in the rank-based signature.

HRD signatures and NK signature

Two independent HRD signatures were derived. The clinically-derived HRD signature was extracted from The Cancer Genome Atlas Cohort (TCGA) cohort where HGSOC patients were stratified into HRD/HRP groups based on a genomic scar analysis using ovaHRDscar[32]. For transcriptomic comparison between these groups, we used the *limma* R package (3.46.0)[68] on the bulk expression data[31] that had been deconvoluted to retain only cancer cell derived expression using PRISM (0.9)[33]. The experimental HRD signature was extracted from the *BRCA1*+/- cell line pair by DEG analysis from the scRNA-seq data as described earlier. For the HRD signatures, we determined $\log_2$FCs of HRD to HRP using sister-concordant genes only, thresholded with indicated *P* values, We used PharmacoGx R package (v3.0.2)[30] to

determine the connectivity scores with method 'gwc' between the HRD signatures and control-corrected pre-resistance signatures.

The NK transcriptional signature was obtained from the coexpressed NK cell markers in our previous analysis of 22 clinical HGSOC tumours[67]. To estimate NK cell enrichment within the immune compartment of the TCGA cohort, we determined Mann-Whitney U statistic-like scores from the deconvoluted gene expressions using the *UCell* R package (1.3.1)[69]. A Cox proportional hazard regression model on the overall survival was built using the *survival* R package (3.3.1)[70], with the formula: *coxph (Surv (time, status) ~ nk_signature:HRD_status),* wherein *'nk_signature'* refers to the UCell score described above.

Predicting drugs to target primed resistance

Six variants of pre-resistant gene expression signatures were generated for each treatment with thresholds of *P* values based on different adjustment methods, including: 1) Original $\log_2$FC, 2) $\log_2$FC corrected by non-treatment control, 3) $\log_2$FC corrected by non-treatment control filtered by bootstrapping test *P* value < 0.05, 4) log2FC corrected by non-treatment control filtered by permutation test *P* value < 0.05, 5) Corrected log2FC filtered by both bootstrapping test *P* value < 0.05 and permutation test *P* value < 0.05, and 6) Original $\log_2$FC filtered by the Wilcoxon rank sum test *P* value < 0.05. The similarity of pre-resistant signatures and drug-induced consensus gene expression signatures were evaluated by the connectivity scores defined in the Connectivity Map (CAMP) project[71]. The drugs with negative connectivity scores consistently with all of the six variants of pre-resistant signatures (*P* value < 0.05) were selected as the top candidates (Table S6). We filtered out drugs whose targets were not expressed or expressed at below 10% of Kuramochi cells, including pifithrin-mu, AM-404, and methenamine for carboplatin; PSI-7976 and cilengitide for olaparib; oxyphenbutazone and WAY-100635 for NK cells (Table S7).

References

1. Sharma, S. V. *et al.* A chromatin-mediated reversible drug-tolerant state in cancer cell subpopulations. *Cell* **141**, 69–80 (2010).
2. Ramirez, M. *et al.* Diverse drug-resistance mechanisms can emerge from drug-tolerant cancer persister cells. *Nat. Commun.* **7**, 10690 (2016).
3. Hangauer, M. J. *et al.* Drug-tolerant persister cancer cells are vulnerable to GPX4 inhibition. *Nature* **551**, 247–250 (2017).
4. Shaffer, S. M. *et al.* Rare cell variability and drug-induced reprogramming as a mode of cancer drug resistance. *Nature* **546**, 431–435 (2017).
5. Thomas, P., Terradot, G., Danos, V. & Weiße, A. Y. Sources, propagation and consequences

of stochasticity in cellular growth. *Nat. Commun.* **9**, 4528 (2018).

6. Shaffer, S. M. *et al.* Memory Sequencing Reveals Heritable Single-Cell Gene Expression Programs Associated with Distinct Cellular Behaviors. *Cell* **182**, 947–959.e17 (2020).

7. Meyer, M. *et al.* Profiling the Non-genetic Origins of Cancer Drug Resistance with a Single-Cell Functional Genomics Approach Using Predictive Cell Dynamics. *Cell Syst* **11**, 367–374.e5 (2020).

8. Biddy, B. A. *et al.* Single-cell mapping of lineage and identity in direct reprogramming. *Nature* **564**, 219–224 (2018).

9. Weinreb, C., Rodriguez-Fraticelli, A., Camargo, F. D. & Klein, A. M. Lineage tracing on transcriptional landscapes links state to fate during differentiation. *Science* **367**, (2020).

10. Gutierrez, C. *et al.* Multifunctional barcoding with ClonMapper enables high-resolution study of clonal dynamics during tumor evolution and treatment. *Nat Cancer* **2**, 758–772 (2021).

11. Oren, Y. *et al.* Cycling cancer persister cells arise from lineages with distinct programs. *Nature* (2021) doi:10.1038/s41586-021-03796-6.

12. Domcke, S., Sinha, R., Levine, D. A., Sander, C. & Schultz, N. Evaluating cell lines as tumour models by comparison of genomic profiles. *Nat. Commun.* **4**, 2126 (2013).

13. Bryant, H. E. *et al.* Specific killing of BRCA2-deficient tumours with inhibitors of poly(ADP-ribose) polymerase. *Nature* **434**, 913–917 (2005).

14. Farmer, H. *et al.* Targeting the DNA repair defect in BRCA mutant cells as a therapeutic strategy. *Nature* **434**, 917–921 (2005).

15. Tumiati, M. *et al.* A Functional Homologous Recombination Assay Predicts Primary Chemotherapy Response and Long-Term Survival in Ovarian Cancer Patients. *Clin. Cancer Res.* **24**, 4482–4493 (2018).

16. Kim, H. *et al.* Combining PARP with ATR inhibition overcomes PARP inhibitor and platinum resistance in ovarian cancer models. *Nat. Commun.* **11**, 3726 (2020).

17. Sathe, P. *et al.* Innate immunodeficiency following genetic ablation of Mcl1 in natural killer cells. *Nat. Commun.* **5**, 4539 (2014).

18. Kivioja, T. *et al.* Counting absolute numbers of molecules using unique molecular identifiers.

Nat. Methods **9**, 72–74 (2011).

19.Adamson, B. *et al.* A Multiplexed Single-Cell CRISPR Screening Platform Enables Systematic Dissection of the Unfolded Protein Response. *Cell* **167**, 1867–1882.e21 (2016).

20.Lin, W.-H. W. *et al.* Asymmetric PI3K Signaling Driving Developmental and Regenerative Cell Fate Bifurcation. *Cell Rep.* **13**, 2203–2218 (2015).

21.Dey-Guha, I. *et al.* Asymmetric cancer cell division regulated by AKT. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 12845–12850 (2011).

22.Dainichi, T. *et al.* PDK1 Is a Regulator of Epidermal Differentiation that Activates and Organizes Asymmetric Cell Division. *Cell Rep.* **15**, 1615–1623 (2016).

23.Mandic, A., Hansson, J., Linder, S. & Shoshan, M. C. Cisplatin induces endoplasmic reticulum stress and nucleus-independent apoptotic signaling. *J. Biol. Chem.* **278**, 9100–9106 (2003).

24.Lahiguera, Á. *et al.* Tumors defective in homologous recombination rely on oxidative metabolism: relevance to treatments with PARP inhibitors. *EMBO Mol. Med.* **12**, e11217 (2020).

25.Belkahla, S. *et al.* The metabolism of cells regulates their sensitivity to NK cells depending on p53 status. *Sci. Rep.* **12**, 3234 (2022).

26.Nolte-'t Hoen, E. N. M. *et al.* Increased surveillance of cells in mitosis by human NK cells suggests a novel strategy for limiting tumor growth and viral replication. *Blood* **109**, 670–673 (2007).

27.Terasima, T. & Tolmach, L. J. X-ray sensitivity and DNA synthesis in synchronous populations of HeLa cells. *Science* **140**, 490–492 (1963).

28.Schoonen, P. M. *et al.* Progression through mitosis promotes PARP inhibitor-induced cytotoxicity in homologous recombination-deficient cancer cells. *Nat. Commun.* **8**, 15981 (2017).

29.Liu, J., François, J.-M. & Capp, J.-P. Gene Expression Noise Produces Cell-to-Cell Heterogeneity in Eukaryotic Homologous Recombination Rate. *Front. Genet.* **10**, 475 (2019).

30.Smirnov, P. *et al.* PharmacoGx: an R package for analysis of large pharmacogenomic datasets. *Bioinformatics* **32**, 1244–1246 (2016).

31.Cancer Genome Atlas Research Network. Integrated genomic analyses of ovarian carcinoma.

Nature **474**, 609–615 (2011).

32.Perez-Villatoro, F. *et al.* Optimized detection of homologous recombination deficiency improves the prediction of clinical outcomes in cancer. *NPJ Precis Oncol* **6**, 96 (2022).

33.Häkkinen, A. *et al.* PRISM: Recovering cell type specific expression profiles from individual composite RNA-seq samples. *Bioinformatics* (2021) doi:10.1093/bioinformatics/btab178.

34.Subramanian, A. *et al.* A Next Generation Connectivity Map: L1000 Platform and the First 1,000,000 Profiles. *Cell* **171**, 1437–1452.e17 (2017).

35.Niepel, M. *et al.* Common and cell-type specific responses to anti-cancer drugs revealed by high throughput transcript profiling. *Nat. Commun.* **8**, 1186 (2017).

36.Lobato-Gil, S. *et al.* Proteome-wide identification of NEDD8 modification sites reveals distinct proteomes for canonical and atypical NEDDylation. *Cell Rep.* **34**, 108635 (2021).

37.Yadav, B., Wennerberg, K., Aittokallio, T. & Tang, J. Corrigendum to 'Searching for drug synergy in complex dose-response landscapes using an interaction potency model' [Comput. Struct. Biotechnol. J. 13 (2015) 504-513]. *Comput. Struct. Biotechnol. J.* **15**, 387 (2017).

38.Tang, J., Wennerberg, K. & Aittokallio, T. What is synergy? The Saariselkä agreement revisited. *Front. Pharmacol.* **6**, 181 (2015).

39.Ferris, J. *et al.* Pevonedistat (MLN4924): mechanism of cell death induction and therapeutic potential in colorectal cancer. *Cell Death Discov* **6**, 61 (2020).

40.Zhou, X. *et al.* Phase I study assessing the mass balance, pharmacokinetics, and excretion of [C]-pevonedistat, a NEDD8-activating enzyme inhibitor in patients with advanced solid tumors. *Invest. New Drugs* **39**, 488–498 (2021).

41.Bhang, H.-E. C. *et al.* Studying clonal dynamics in response to cancer therapy using high-complexity barcoding. *Nat. Med.* **21**, 440–448 (2015).

42.Contreras-Trujillo, H. *et al.* Deciphering intratumoral heterogeneity using integrated clonal tracking and single-cell transcriptome analyses. *Nat. Commun.* **12**, 6522 (2021).

43.Belderbos, M. E. *et al.* Clonal selection and asymmetric distribution of human leukemia in murine xenografts revealed by cellular barcoding. *Blood* **129**, 3210–3220 (2017).

44.Kimmerling, R. J. *et al.* A microfluidic platform enabling single-cell RNA-seq of

multigenerational lineages. *Nat. Commun.* **7**, 10220 (2016).

45.Ciriello, G. *et al.* Emerging landscape of oncogenic signatures across human cancers. *Nat. Genet.* **45**, 1127–1133 (2013).

46.Leary, R. J. *et al.* Integrated analysis of homozygous deletions, focal amplifications, and sequence alterations in breast and colorectal cancers. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 16224–16229 (2008).

47.Gasser, S., Orsulic, S., Brown, E. J. & Raulet, D. H. The DNA damage pathway regulates innate immune system ligands of the NKG2D receptor. *Nature* **436**, 1186–1190 (2005).

48.Wang, R. W., Viganò, S., Ben-David, U., Amon, A. & Santaguida, S. Aneuploid senescent cells activate NF-κB to promote their immune clearance by NK cells. *EMBO Rep.* **22**, e52032 (2021).

49.Launonen, I.-M. *et al.* Single-cell tumor-immune microenvironment of BRCA1/2 mutated high-grade serous ovarian cancer. *Nat. Commun.* **13**, 835 (2022).

50.Lee, M. J. *et al.* Sequential application of anticancer drugs enhances cell death by rewiring apoptotic signaling networks. *Cell* **149**, 780–794 (2012).

51.Ducreux, M. *et al.* Sequential versus combination chemotherapy for the treatment of advanced colorectal cancer (FFCD 2000-05): an open-label, randomised, phase 3 trial. *Lancet Oncol.* **12**, 1032–1044 (2011).

52.Denman, C. J. *et al.* Membrane-bound IL-21 promotes sustained ex vivo proliferation of human natural killer cells. *PLoS One* **7**, e30264 (2012).

53.Joung, J. *et al.* Genome-scale CRISPR-Cas9 knockout and transcriptional activation screening. *Nat. Protoc.* **12**, 828–863 (2017).

54.Orthwein, A. *et al.* A mechanism for the suppression of homologous recombination in G1 cells. *Nature* **528**, 422–426 (2015).

55.Wang, S.-C. *et al.* Inhibition of cancer cell growth by BRCA2. *Cancer Res.* **62**, 1311–1314 (2002).

56.Zheng, S. *et al.* SynergyFinder Plus: Toward Better Interpretation and Annotation of Drug Combination Screening Datasets. *Genomics Proteomics Bioinformatics* (2022)

doi:10.1016/j.gpb.2022.01.004.

57.Zheng, G. X. Y. *et al.* Massively parallel digital transcriptional profiling of single cells. *Nat. Commun.* **8**, 14049 (2017).

58.Smith, T., Heger, A. & Sudbery, I. UMI-tools: modeling sequencing errors in Unique Molecular Identifiers to improve quantification accuracy. *Genome Res.* **27**, 491–499 (2017).

59.Hao, Y. *et al.* Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573–3587.e29 (2021).

60.La Manno, G. *et al.* RNA velocity of single cells. *Nature* **560**, 494–498 (2018).

61.Bergen, V., Lange, M., Peidli, S., Wolf, F. A. & Theis, F. J. Generalizing RNA velocity to transient cell states through dynamical modeling. *Nat. Biotechnol.* **38**, 1408–1414 (2020).

62.J. Oksanen, F. Guillaume Blanchet, M. Friendly, R. Kindt, P. Legendre, D. McGlinn, P. R. Minchin, R. B. O'Hara, G. L. Simpson, P. Solymos, M. Henry H. Stevens, E. Szoecs, H. Wagner. vegan: Community Ecology Package. R package version 2.6-4. *Comprehensive R Archive Network (CRAN)* (2022) doi:https://CRAN.R-project.org/package=vegan.

63.Kuleshov, M. V. *et al.* Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res.* **44**, W90–7 (2016).

64.Xie, Z. *et al.* Gene Set Knowledge Discovery with Enrichr. *Curr Protoc* **1**, e90 (2021).

65.Yu, G., Wang, L.-G., Han, Y. & He, Q.-Y. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* **16**, 284–287 (2012).

66.T. Tickle, G. C. Ti, M. Brown, B. Haas. inferCNV of the Trinity CTAT Project (Klarman Cell Observatory, Broad Institute of MIT and Harvard). (2019).

67.Zhang, K. *et al.* Longitudinal single-cell RNA-seq analysis reveals stress-promoted chemoresistance in metastatic ovarian cancer. *Sci Adv* **8**, eabm1831 (2022).

68.Ritchie, M. E. *et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47 (2015).

69.Andreatta, M. & Carmona, S. J. UCell: Robust and scalable single-cell gene signature scoring. *Comput. Struct. Biotechnol. J.* **19**, 3796–3798 (2021).

70.Therneau, T. M. & Grambsch, P. M. Modeling Survival Data: Extending the Cox Model.

Statistics for Biology and Health Preprint at https://doi.org/10.1007/978-1-4757-3294-8 (2000).

71.Lamb, J. *et al.* The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease. *Science* **313**, 1929–1935 (2006).

Acknowledgements

We thank FIMM High Content Imaging and Analysis unit for imaging services. The results are in part based upon data generated by the TCGA Research Network: https://www.cancer.gov/tcga. Cartoons in Figs. 3a and 4a were created with BioRender.com.

Funding: This work was supported by The Sigrid Jusélius Foundation (AV and JT), The Cancer Foundation Finland (AV), The Academy of Finland projects 289059 (AV), 319243 (AV), 317680 (JT), 320131 (JT), 351165 and 351196 (JT, AV), ERA PerMed JTC2020 PARIS/Academy of Finland project 344697 (AV), European Research Council No. 716063 (JT).

Author contributions: Conceptualization and supervision: J.T. and A.V. Study design, data interpretation, and preparation of the manuscript: A.V., J.T., J.D., S.Z., M.F., and J.E. Execution of experiments: J.D., J.E., J.J., L.G., S.F., S.P. and J.B.; NK cell preparation: O.D., and K.S. Computational and statistical analysis: S.Z., J.T., M.F., A.A., Y.W., W.W. and F.P. Resources: A.V., J.T., S.M., L.K. and A.F. All authors wrote the paper.

Competing interests: Authors declare that they have no competing interests.

Data and materials availability: All the data needed to evaluate the conclusions in the paper are available in the manuscript or the supplementary materials. All of the ReSisTrace scRNA-seq raw data have been deposited in the Gene Expression Omnibus (GEO) under accession code GSE223003. The source code for data analysis will be made available at GitHub (https://github.com/TangSoftwareLab/ReSisTrace) and Zenodo (inserted upon publication).

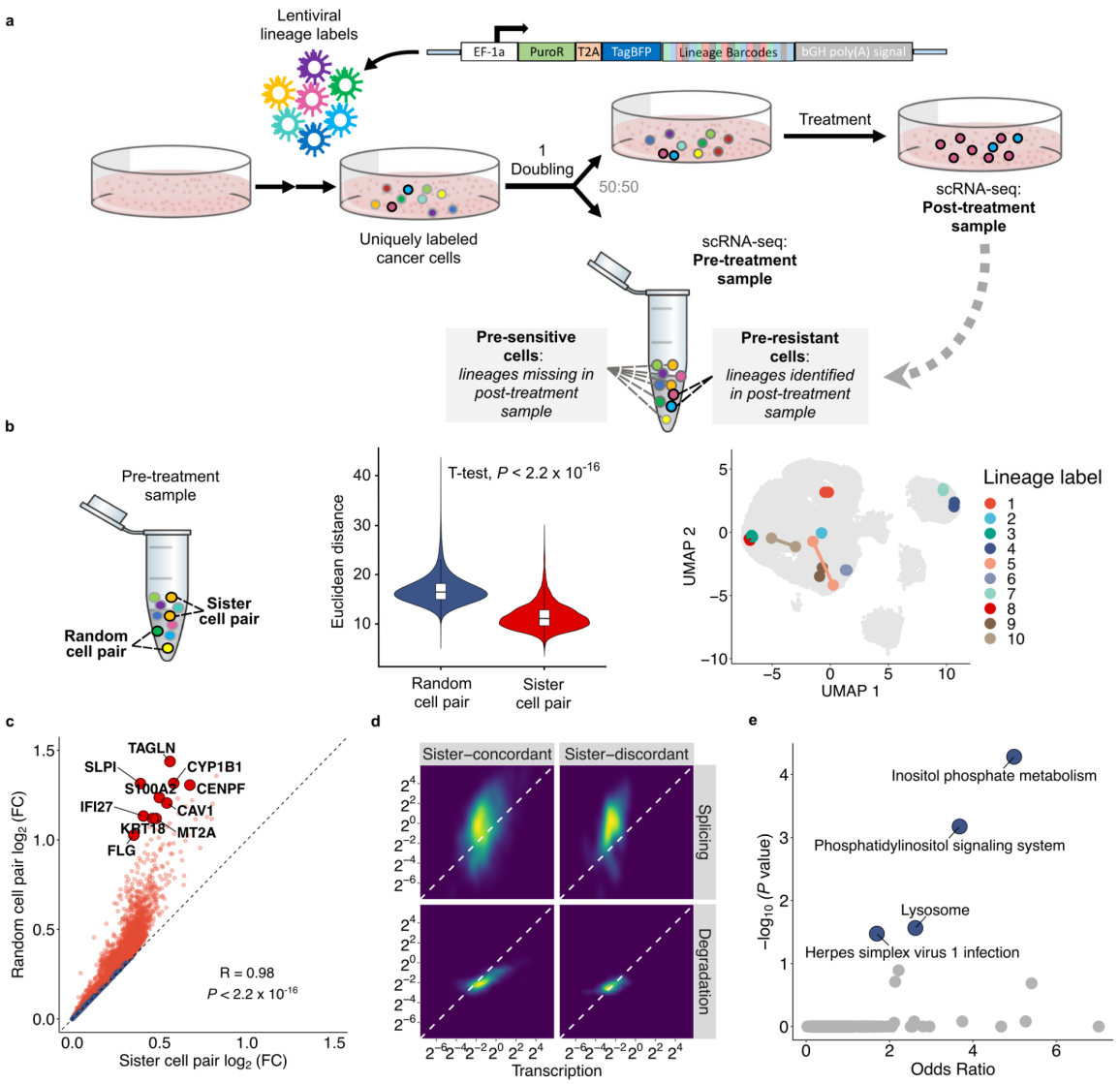

Fig. 1. ReSisTrace leverages sister cell similarity to discover cellular states primed for resistance. (**a**) Schematic representation of the ReSisTrace approach. Uniquely labelled cells are synchronised and allowed to divide once. Subsequently, the sample is divided in two for scRNA-seq and treatment. Next, barcodes of recovered cells in the post-treatment sample are determined by scRNA-seq and used to identify pre-resistant cells in the pre-treatment sample. (**b**) Sister cell similarity. Left, schematic of sister cell pairs and random cell pairs in a pre-treatment sample. Middle, Euclidean distance of sister cell pairs in comparison to random cell pairs ($P < 2.2 \times 10^{-16}$, two-tailed t-test, random cell pairs n = 6.85×10^{8}; sister cell pairs n = 3,477). Right, a subset of randomly selected 10 sister cell pairs shown in the Uniform manifold approximation and projection (UMAP). (**c**) Gene expression fold changes in sister cell pairs (x axis) versus random cell pairs (y axis). Sister-concordant genes are displayed as red, and sister-discordant genes as blue. (**d**) Inferred splicing and degradation rates of 919 sister-concordant and 107 sister-discordant genes. (**e**) Pathway enrichment of 1,332 sister-discordant genes.

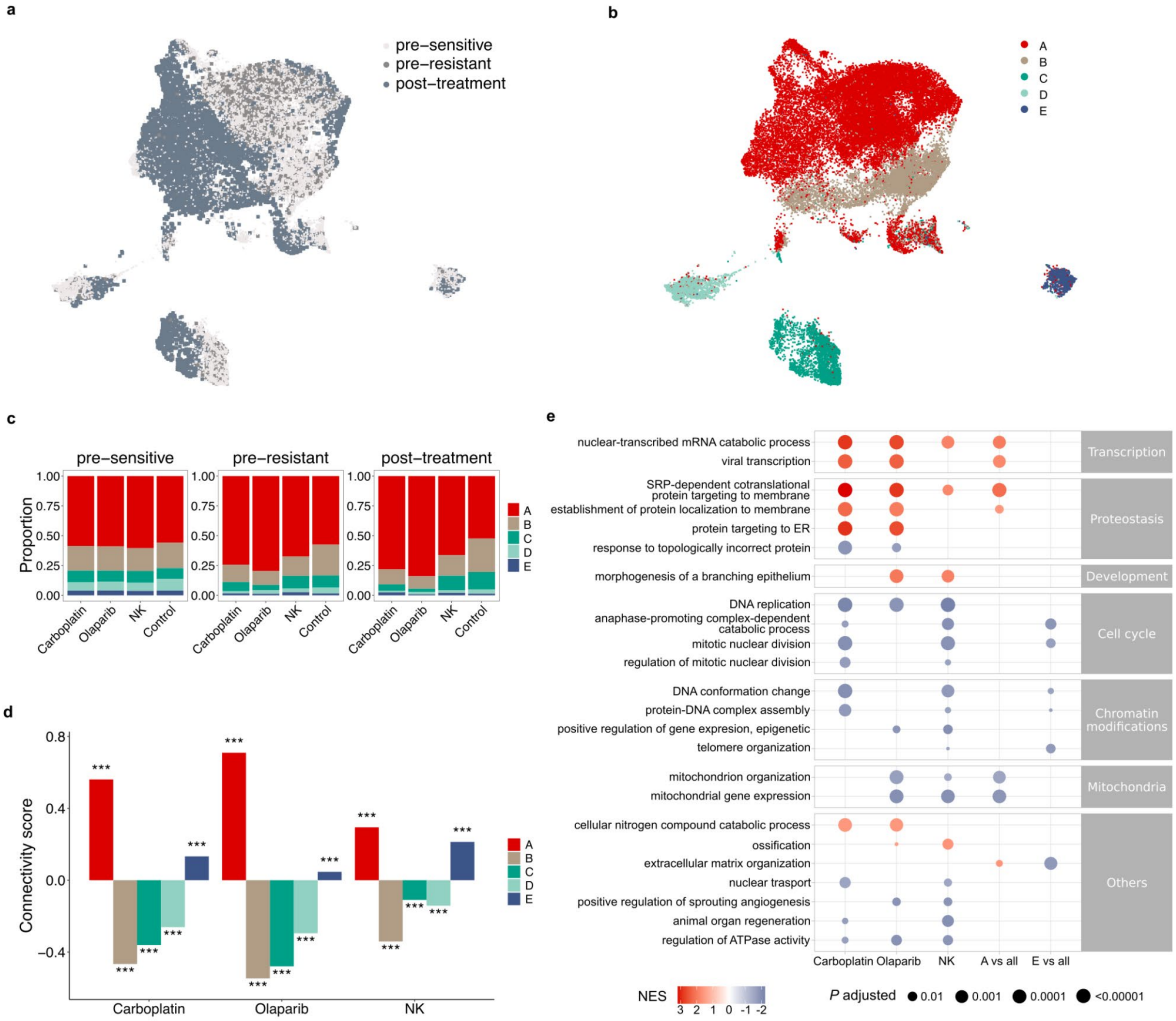

Fig. 2. Interplay between genetic and transcriptomic features in primed resistance. UMAP projection of single cells coloured by (**a**) the treatment groups (pre-sensitive n = 19,540, pre-resistant n = 3,102, and post-treatment n = 7,802), or (**b**) the inferred CNV subclones. (A, n = 24,903; B, n = 7,904; C n = 3,824; D, n = 2,482; E, n = 1,331) (**c**) Subclonal proportions in pre-sensitive, pre-resistant, and post-treatment populations for different treatments. (**d**) connectivity scores between the pre-resistance signatures and the subclone specific signatures. ***$P < 0.001$; two-tailed t-test; n = 8,419. (**e**) Representative gene ontologies enriched in the pre-resistance signatures as well as the signatures of subclones A and E.

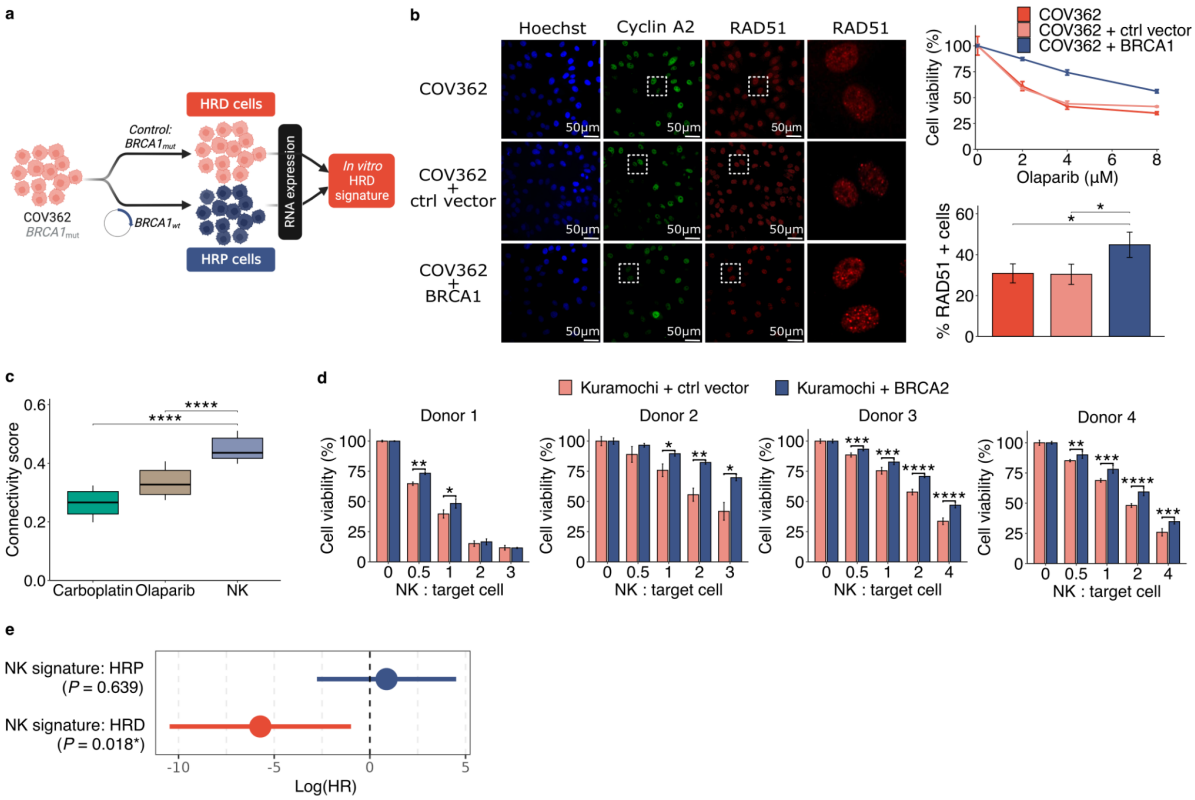

Fig. 3. Context-dependent vulnerability of HRD ovarian cancer to innate immunity. (**a**) Schematic representation of the method to determine an HRD signature, derived from scRNA-seq analysis of parental (*BRCA1* mutant) COV362 cells against *BRCA1*-restored cells. (**b**) Functional HR test of RAD51 foci within cyclin A2 positive cells (left: representative images, bottom right: quantification from three replicates) and olaparib sensitivity of the parental, control plasmid transduced or BRCA1 restored COV362 cells.* $0.01 < P < 0.05$, two-tailed t-test. Error bars are mean +/- s.d. of three replicates. (**c**) Connectivity scores between the pre-sensitive gene signatures and the HRD signature. Quantiles were determined by varying the adjusted *P* value threshold for filtering the genes in the HRD signature ($5\times10^{-16} < P < 0.05$, n = [211, 946]). **** $P < 1\times10^{-4}$, two-tailed t-test. (**d**) Measurements of NK cell-mediated killing of Kuramochi control and *BRCA2*-restored cells using NK cells from four donors. Error bars are mean +/- s.d. of at least three replicates. * $0.01 < P < 0.05$; ** $0.001 < P < 0.01$; *** $1\times10^{-4} < P < 0.001$; **** $P < 1\times10^{-4}$, two-tailed t-test. (**e**) Hazard ratios of overall survival in logarithmic scale for the interaction between NK transcriptional signature and patients' HRD/HRP status. *P* values are determined by the Cox regression models. Error bars are 95% confidence intervals.

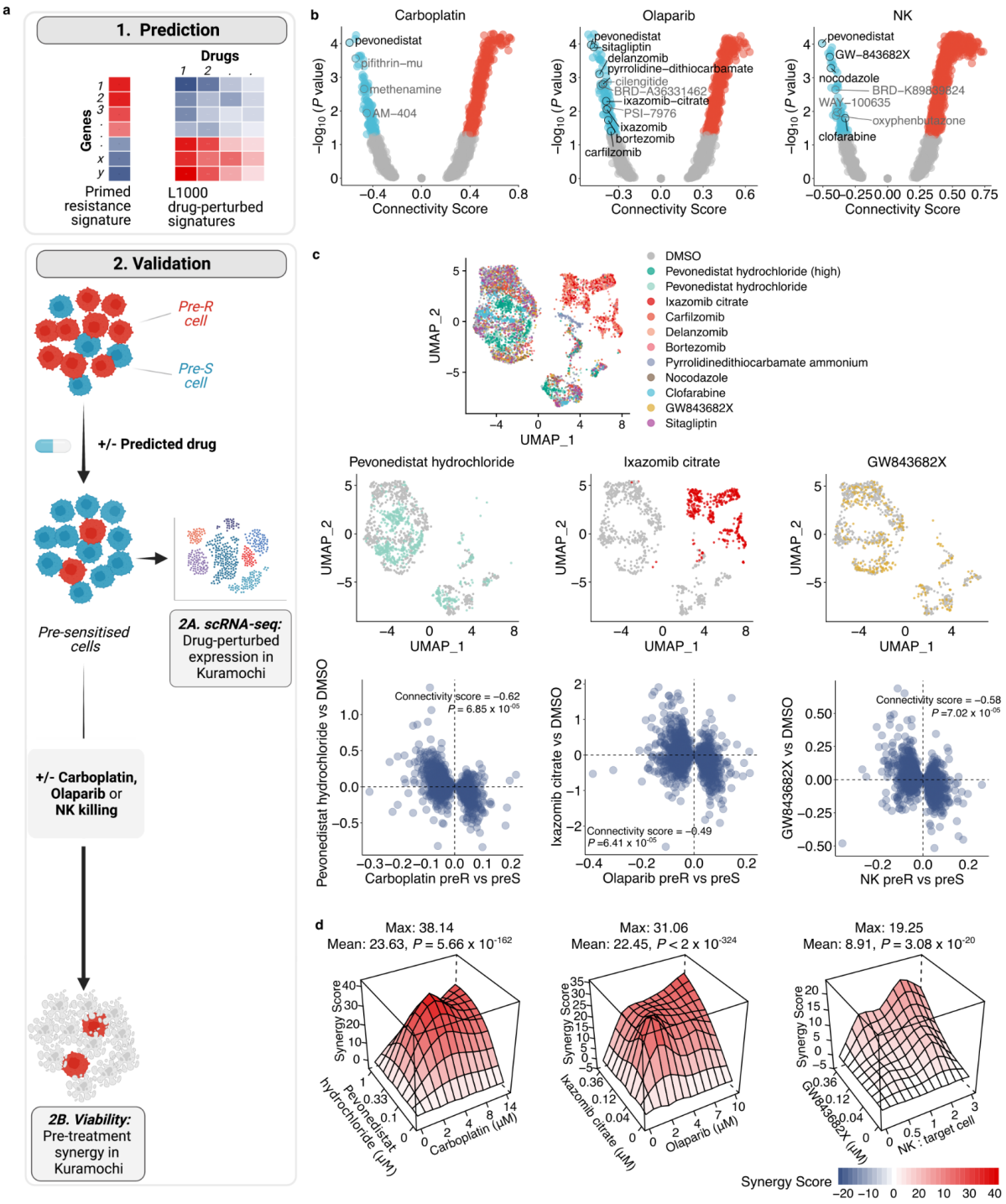

Fig. 4. Predicting drugs targeting of primed resistance. (**a**) Schematic representation of the approach to predict and validate drugs that leverage the acquired pre-resistance profiles by sequential treatment strategies. (**b**) Connectivity scores between the L1000 drug-induced signatures and acquired pre-resistance signatures, adjusted by non-treatment control. (**c**) Single-cell gene expression of drug perturbed Kuramochi cells projected on a UMAP (upper panel), with top predicted drugs shown as separate UMAP projections along with DMSO induced profiles (mid panel), and with induced fold changes compared to those of respective

pre-resistance profiles (lower panel). (**d**) Synergy landscapes for the top drugs for each treatment. Mean and maximal synergy scores were shown, with their *P* values determined by bootstrapping test.

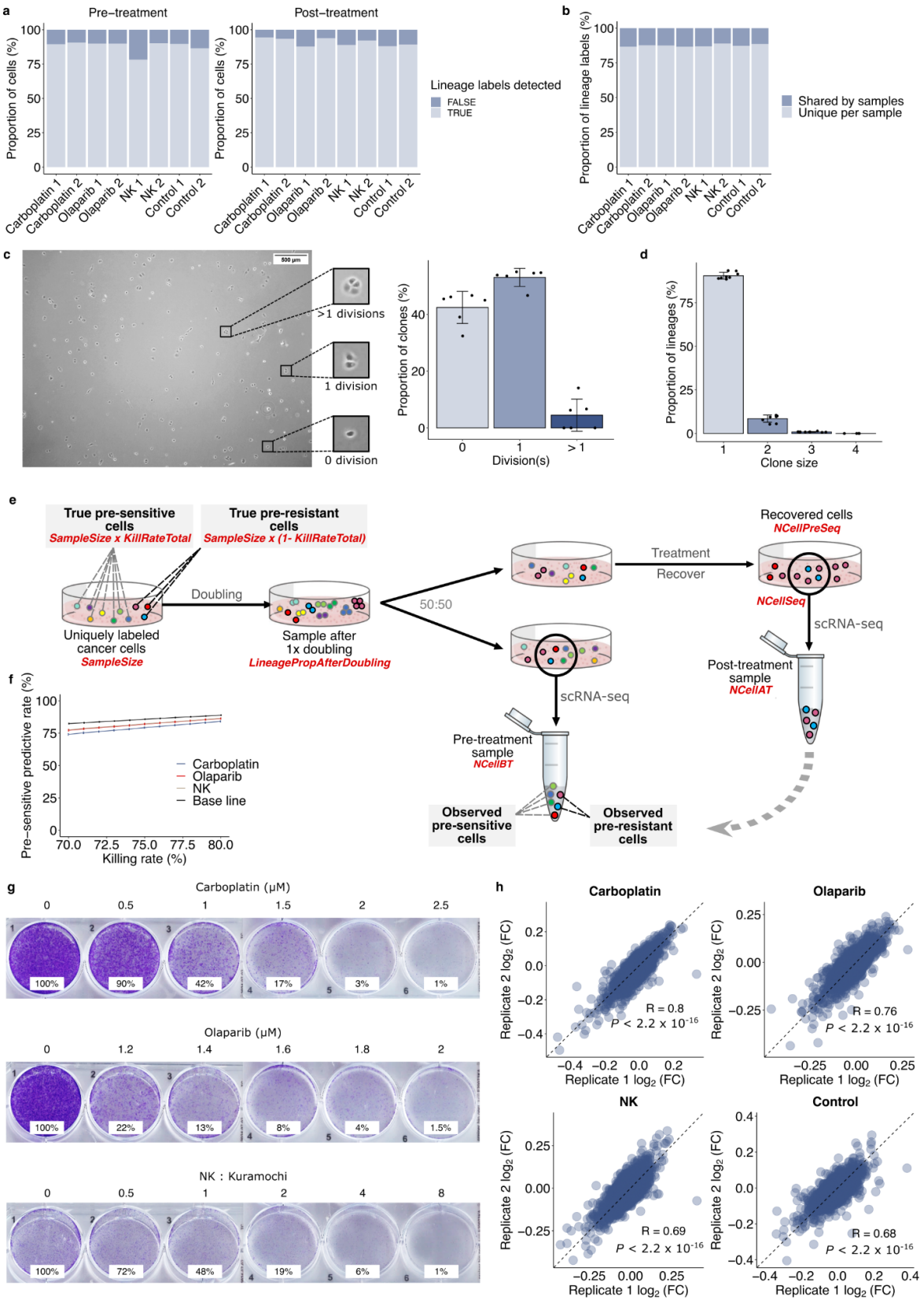
a
Pre−treatment
Post−treatment
Proportion of cells (%)
Carboplatin 1
Carboplatin 2
Olaparib 1
Olaparib 2
NK 1
NK 2
Control 1
Control 2
Lineage labels detected
FALSE
TRUE
b
Proportion of lineage labels (%)
Shared by samples
Unique per sample
c
>1 divisions
1 division
0 division
Proportion of clones (%)
Division(s)
d
Proportion of lineages (%)
Clone size
e
True pre-sensitive cells
SampleSize x KillRateTotal
True pre-resistant cells
SampleSize x (1- KillRateTotal)
Doubling
Uniquely labeled cancer cells
SampleSize
Sample after 1x doubling
LineagePropAfterDoubling
50:50
Treatment
Recover
Recovered cells
NCellPreSeq
NCellSeq
scRNA-seq
Post-treatment sample
NCellAT
Pre-treatment sample
NCellBT
Observed pre-sensitive cells
Observed pre-resistant cells
f
Pre−sensitive predictive rate (%)
Killing rate (%)
Carboplatin
Olaparib
NK
Base line
g
Carboplatin (μM)
Olaparib (μM)
NK : Kuramochi
h
Carboplatin
Olaparib
NK
Control
Replicate 2 log2 (FC)
Replicate 1 log2 (FC)
R = 0.8
R = 0.76
R = 0.69
R = 0.68
P < 2.2 x 10-16

Fig. S1. Cell doubling, lineage labelling efficiency, and quality assessment of ReSisTrace. (**a**) Percentage of cells with the lineage barcodes detected in each sample. (**b**) Proportion of sample-shared or sample-unique barcodes in each pre-treatment sample. **(c**) A representative phase-contrast image of cells after one doubling time of 48 hours (left), and proportion of clones for indicated numbers of divisions (right). Error bars are mean +/- s.d. of six images. (**d**) Proportion of lineages with indicated clone sizes detected in the pre-treatment scRNA-seq samples. Error bars are mean +/- s.d. of eight samples (4 treatment conditions each with 2 replicates). (**e**) Computer simulation of the ReSisTrace experiment. (**f**) Pre-sensitive predictive rate for each treatment condition, as compared to the upper bound of perfect doubling and no drop-out (baseline). (**g**) Colony formation of Kuramochi cell line treated by NK cells (1 day), olaparib (7 days), or carboplatin (3 days) with indicated ratio or concentration, followed by 7, 6, or 10 days recovery, respectively. The number under each well indicates the remaining cell number compared to the corresponding untreated control well. (**h**) Consistency of gene expression changes for pre-resistant against pre-sensitive cells between the replicates of each treatment condition. (Carboplatin n = 30,613; Olaparib n = 30,586 ; NK n = 30,613; Control n = 27,478)

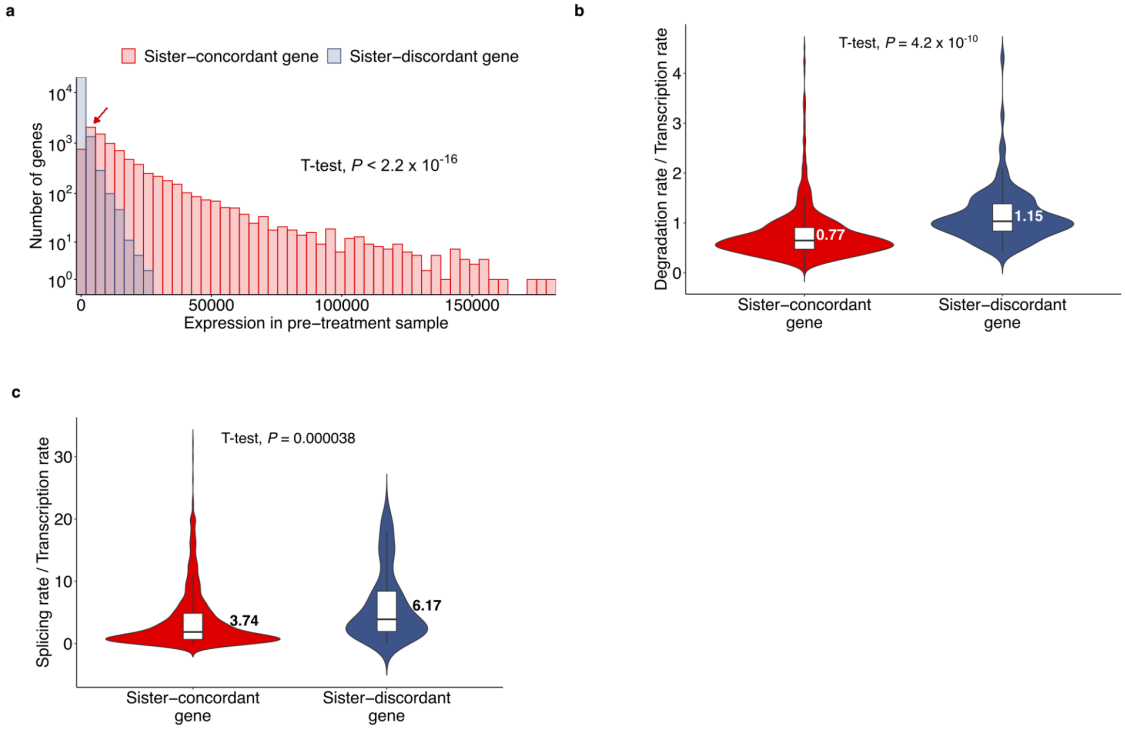

Fig. S2. Characteristics of sister-concordant and sister-discordant genes. (**a**) Histogram showing the number of genes at different expression levels of sister-concordant (red, n = 8,419) or sister-discordant (blue, n = 21,896) genes in the pre-treatment samples. Red arrow indicates the bin ([1839.89, 5519.11]) used to determine the subsets of sister-concordant and sister-discordant genes in the RNA velocity and pathway enrichment analyses. (**b**, **c**) Average RNA degradation to transcription ratio and average RNA splicing to transcription ratio of sister-concordant (n = 919) and discordant genes (n = 107) velocity genes used to fit the dynamical model.

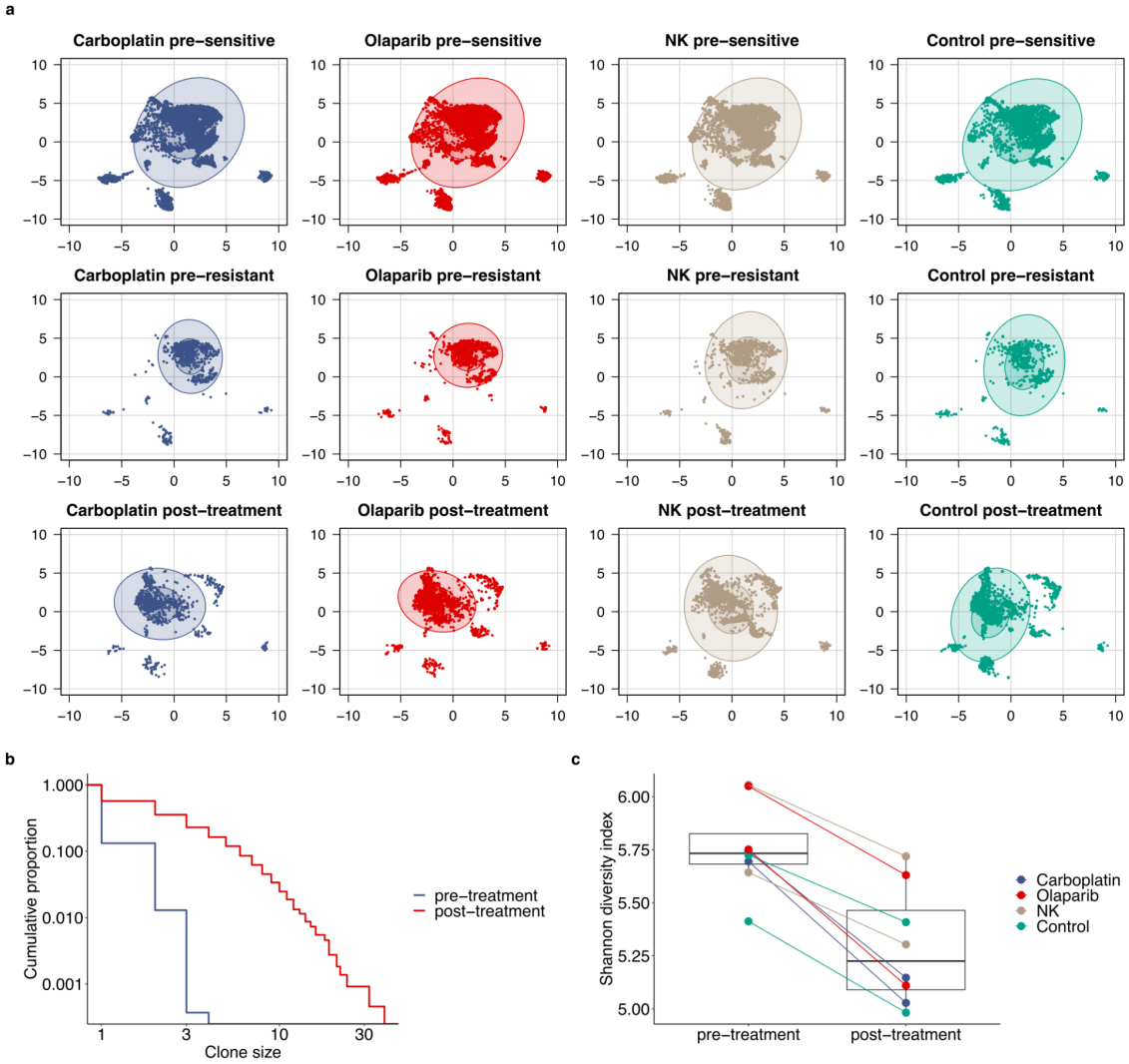

Fig. S3. Diversity of ReSisTrace populations. (**a**) UMAP projection of pre-sensitive, pre-resistant, and post-treatment cells for each treatment. (**b**) Cumulative proportion of lineages with different clone sizes in the pre-treatment or post-treatment samples. (**c**) Shannon diversity index in the pre-treatment or post-treatment samples for each replicate in different treatment conditions.

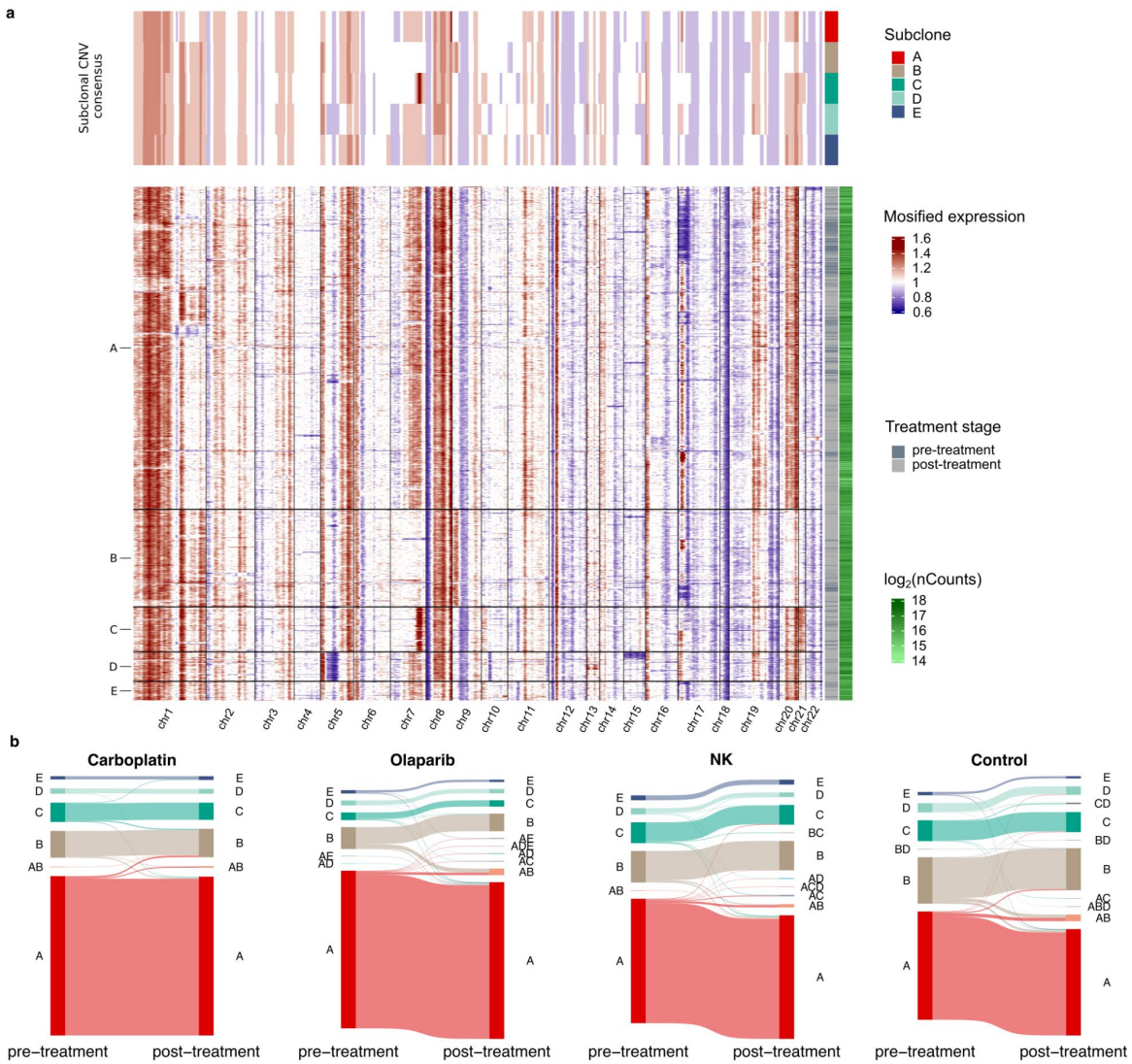

Fig. S4. Lineages maintain their subclonal identities during the treatment. (**a**) Heatmap showing the subclonal CNV profiles of the carboplatin samples inferred from their scRNA-seq data. The lower panel shows the CNV profiles for the individual cells. The upper panel shows the consensus CNV profiles for each identified subclone. (**b**) Sankey plot for the subclonal dynamics during each treatment. Each subclone is represented by a colored band, the height of which corresponds to the proportion of identified lineages.

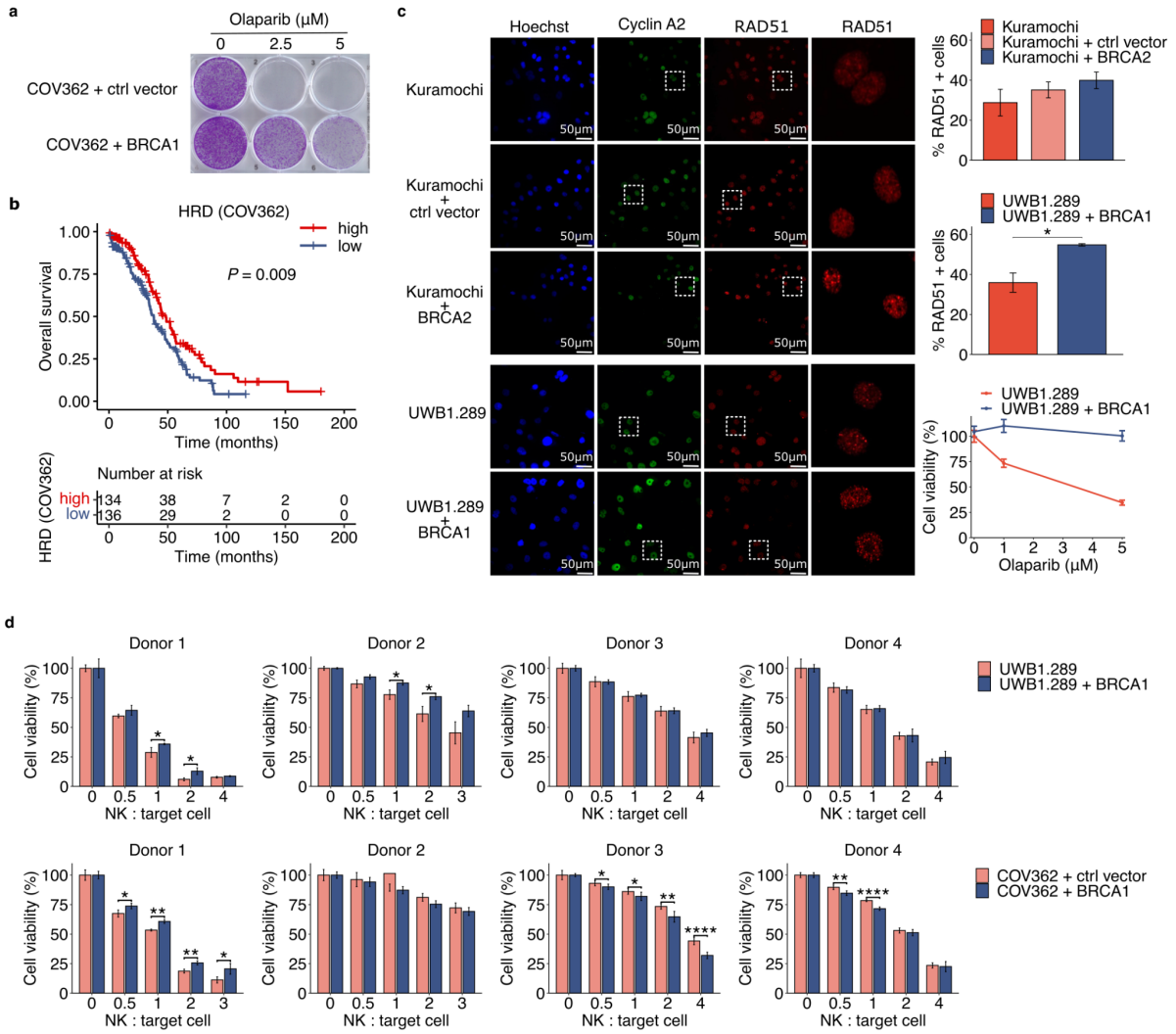

Fig. S5. BRCA restored cell lines display an HRP phenotype and are less sensitive to NK killing. (**a**) Colony formation assay of parental and BRCA1 restored COV362 cells treated with indicated concentrations of olaparib. (**b**) Kaplan-Mayer curves for the overall survival in the TCGA ovarian cancer cohort, stratified to high and low groups by the experimental HRD signature (COV362) (P = 0.009, log-rank test). The genes in the signatures were selected by adjusted P value threshold of 1×10^{-8}. (**c**) HRD test showing RAD51 foci for control plasmid transduced or BRCA2 restored Kuramochi cells, and UWB1.289 parental or BRCA1 restored cells. * $0.1 < P < 0.05$, two-tailed t-test. Error bars are mean +/- s.d. of three replicates. (**d**) NK killing of parental or BRCA2 restored UWB1.289 cells (top), and control plasmid transduced or BRCA1 restored COV362 cells (bottom). Killing rates of NK cells from four different donors are shown separately. * $0.01 < P < 0.05$; ** $0.001 < P < 0.01$; *** $1\times10^{-4} < P < 0.001$; **** $P < 1\times10^{-4}$, two-tailed t-test. Error bars are mean +/- s.d. of at least three replicates.

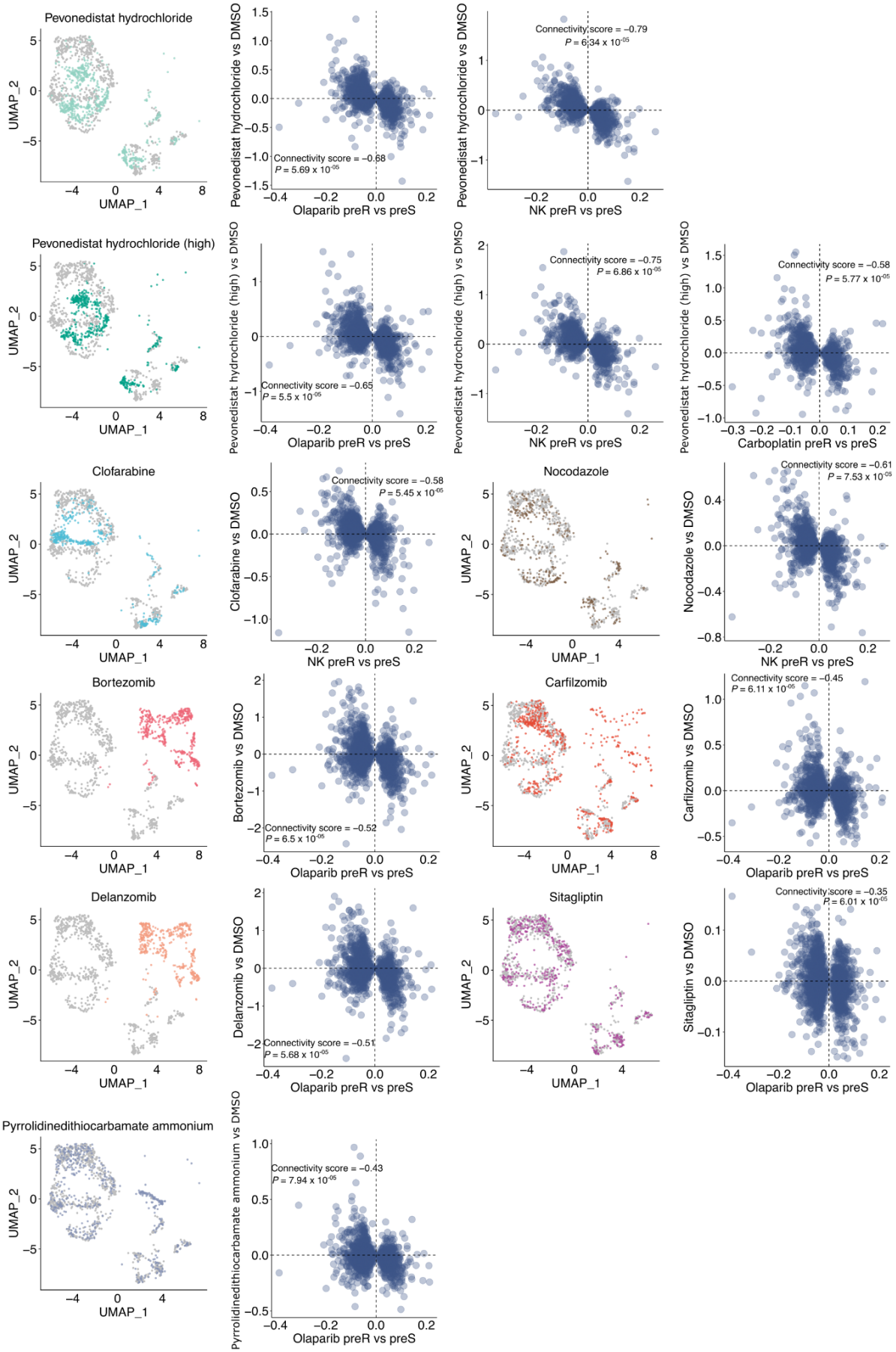
Pevonedistat hydrochloride
UMAP_2
UMAP_1
Pevonedistat hydrochloride vs DMSO
Olaparib preR vs preS
Connectivity score = −0.68
P = 5.69 x 10-05
NK preR vs preS
Connectivity score = −0.79
P = 6.34 x 10-05
Pevonedistat hydrochloride (high)
Pevonedistat hydrochloride (high) vs DMSO
Connectivity score = −0.65
P = 5.5 x 10-05
Connectivity score = −0.75
P = 6.86 x 10-05
Carboplatin preR vs preS
Connectivity score = −0.58
P = 5.77 x 10-05
Clofarabine
Clofarabine vs DMSO
Connectivity score = −0.58
P = 5.45 x 10-05
Nocodazole
Nocodazole vs DMSO
Connectivity score = −0.61
P = 7.53 x 10-05
Bortezomib
Bortezomib vs DMSO
Connectivity score = −0.52
P = 6.5 x 10-05
Carfilzomib
Carfilzomib vs DMSO
Connectivity score = −0.45
P = 6.11 x 10-05
Delanzomib
Delanzomib vs DMSO
Connectivity score = −0.51
P = 5.68 x 10-05
Sitagliptin
Sitagliptin vs DMSO
Connectivity score = −0.35
P = 6.01 x 10-05
Pyrrolidinedithiocarbamate ammonium
Pyrrolidinedithiocarbamate ammonium vs DMSO
Connectivity score = −0.43
P = 7.94 x 10-05

Fig. S6. Single-cell gene expression of drug perturbed Kuramochi cells as UMAP projections, with indicated drug-induced cells as coloured and DMSO control as grey, followed by mean log fold changes of drug-induced expression changes compared to respective pre-resistance profiles.

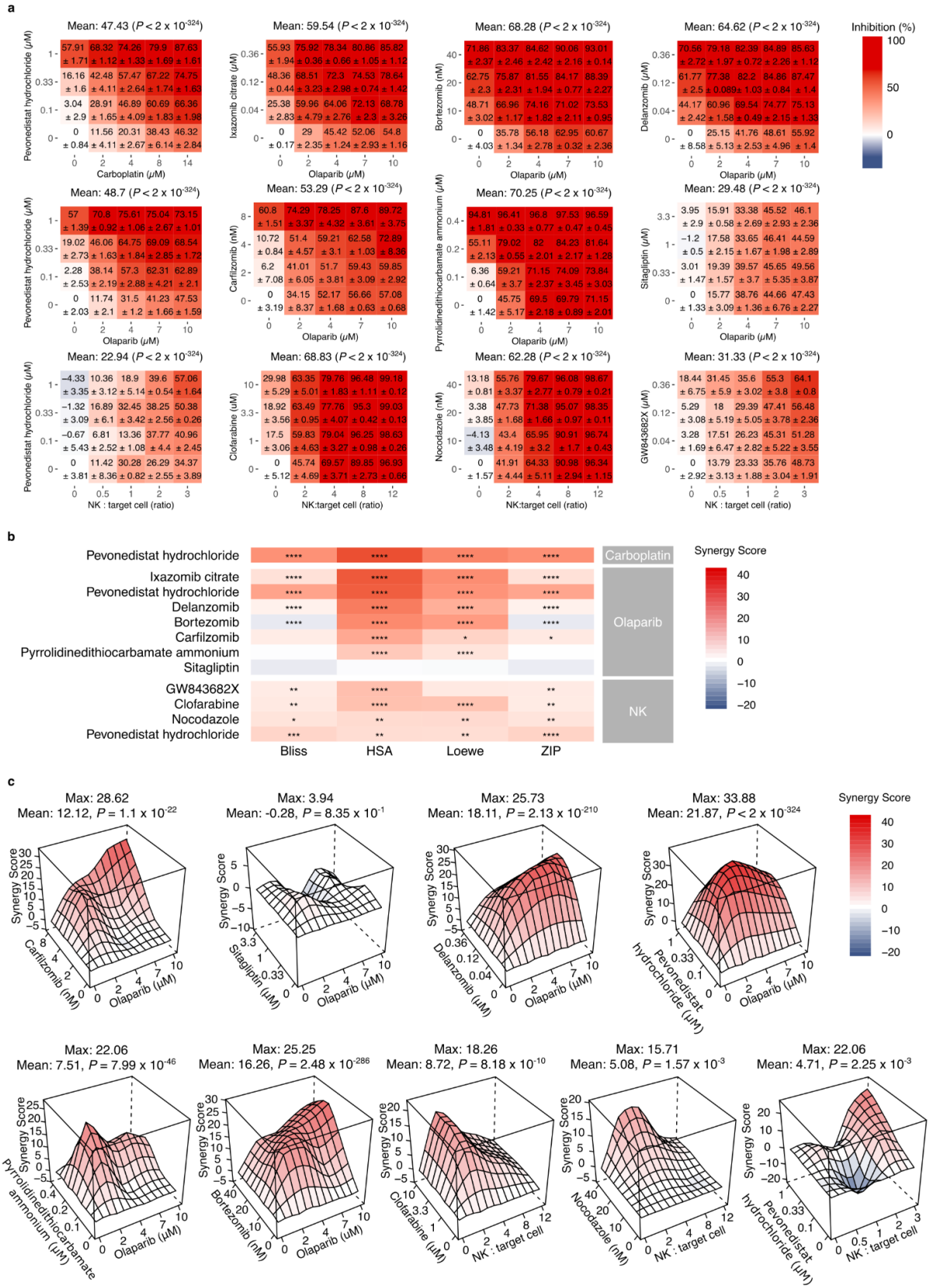
a
Mean: 47.43 (P < 2 x 10-324)
Pevonedistat hydrochloride (μM)
Carboplatin (μM)
Mean: 59.54 (P < 2 x 10-324)
Ixazomib citrate (μM)
Olaparib (μM)
Mean: 68.28 (P < 2 x 10-324)
Bortezomib (nM)
Olaparib (μM)
Mean: 64.62 (P < 2 x 10-324)
Delanzomib (μM)
Olaparib (μM)
Inhibition (%)
Mean: 48.7 (P < 2 x 10-324)
Olaparib (μM)
Mean: 53.29 (P < 2 x 10-324)
Carfilzomib (nM)
Olaparib (μM)
Mean: 70.25 (P < 2 x 10-324)
Pyrrolidinedithiocarbamate ammonium (μM)
Olaparib (μM)
Mean: 29.48 (P < 2 x 10-324)
Sitagliptin (μM)
Olaparib (μM)
Mean: 22.94 (P < 2 x 10-324)
NK : target cell (ratio)
Mean: 68.83 (P < 2 x 10-324)
Clofarabine (uM)
NK:target cell (ratio)
Mean: 62.28 (P < 2 x 10-324)
Nocodazole (nM)
NK:target cell (ratio)
Mean: 31.33 (P < 2 x 10-324)
GW843682X (μM)
NK : target cell (ratio)
b
Pevonedistat hydrochloride
Ixazomib citrate
Pevonedistat hydrochloride
Delanzomib
Bortezomib
Carfilzomib
Pyrrolidinedithiocarbamate ammonium
Sitagliptin
GW843682X
Clofarabine
Nocodazole
Pevonedistat hydrochloride
Carboplatin
Olaparib
NK
Bliss
HSA
Loewe
ZIP
Synergy Score
c
Max: 28.62
Mean: 12.12, P = 1.1 x 10-22
Synergy Score
Carfilzomib (nM)
Olaparib (μM)
Max: 3.94
Mean: -0.28, P = 8.35 x 10-1
Sitagliptin (μM)
Olaparib (μM)
Max: 25.73
Mean: 18.11, P = 2.13 x 10-210
Delanzomib (μM)
Olaparib (μM)
Max: 33.88
Mean: 21.87, P < 2 x 10-324
Pevonedistat hydrochloride (μM)
Olaparib (μM)
Synergy Score
Max: 22.06
Mean: 7.51, P = 7.99 x 10-46
Pyrrolidinedithiocarbamate ammonium (μM)
Olaparib (μM)
Max: 25.25
Mean: 16.26, P = 2.48 x 10-286
Bortezomib (nM)
Olaparib (μM)
Max: 18.26
Mean: 8.72, P = 8.18 x 10-10
Clofarabine (μM)
NK : target cell
Max: 15.71
Mean: 5.08, P = 1.57 x 10-3
Nocodazole (nM)
NK : target cell
Max: 22.06
Mean: 4.71, P = 2.25 x 10-3
Pevonedistat hydrochloride (μM)
NK : target cell

Fig. S7. Screening results for the predicted sequential drug combinations. (**a**) Dose-response matrices for the predicted sequential combinations, where cells were first treated with drugs with concentrations indicated on the vertical axis for 24 hours, followed by the treatment indicated on the horizontal axis (carboplatin or olaparib for 5 days, or by NK killing for 24 hours) (**b**) Synergy scores for the predicted drug combinations. * $0.01 < P < 0.05$; ** $0.001 < P < 0.01$; *** $1\times10^{-4} < P < 0.001$; **** $P < 1\times10^{-4}$, bootstrapping test. (**c**) HSA synergy landscapes for other predicted drug combinations. *P* values are determined by bootstrapping tests.

Supplementary Tables

Table S1. Cell counts from scRNA-seq experiments for determining the parameters in the computer simulations.

Table S2. Pre-resistance gene expression signatures for the treatments, and the CNV subclones.

Table S3. Full list of significant GO annotations for pre-resistance and subclonal signatures.

Table S4. HRD signatures from COV362 and TCGA clinical cohort.

Table S5. NK transcriptional signature.

Table S6. Predicted drugs for sensitising the intended treatments.

Table S7. Expression of the drug targets in Kuramochi cells.

Table S8. Primers for experimental BRCAness model constructs.

Table S9. Thresholds for the scRNA-seq data quality control.