Resistance patterns in drug-adapted cancer cell lines reflect the complex evolution in clinical tumours

Helen E. Grimsley[1,2], Magdalena Antczak[1,3], Ian G. Reddin[4], Katie-May McLaughlin[1], Andrea Nist[5], Marco Mernberger[6], Thorsten Stiewe[5,6], Tim R. Fenton[4], Daniel Speidel[7,8], Catherine Harper-Wynne[9], Karina Cox[10], Jindrich Cinatl jr.[11], Mark N. Wass[1]*, Michelle D. Garrett[1]*, Martin Michaelis[1,11]*

[1]School of Biosciences, Stacey Building, University of Kent, Canterbury, Kent, CT2 7NJ, UK

[2]Department of Radiation Oncology and the Molecular Biology Program, Memorial Sloan Kettering Cancer Center, New York, United States

[3]Current address: Queensland Cyber Infrastructure Foundation Ltd., Facility for Advanced Bioinformatics, Brisbane, QLD, 4072, Australia

[4]Cancer Sciences, Faculty of Medicine, University of Southampton, Southampton SO16 6YD, UK

[5]Genomics Core Facility, Philipps University, 35043 Marburg, Germany

[6]Institute of Molecular Oncology, Member of the German Center for Lung Research (DZL), Philipps University, 35032 Marburg, Germany

[7]Children's Medical Research Institute, The University of Sydney, 214 Hawkesbury Road, Westmead, New South Wales, Australia

[8]Current address: Breakpoint Therapeutics GmbH, Essener Bogen 7, 22419 Hamburg, Germany

[9]Kent Oncology Centre, Maidstone and Tunbridge Wells NHS Trust, Hermitage Lane, Maidstone, ME16 9QQ, UK

[10]Peggy Wood Breast Unit, Maidstone Hospital, Hermitage Lane, Maidstone, Kent, ME16 9QQ, UK

[11]Dr Petra Joh-Research Institute, 60528 Frankfurt am Main, Germany

*Corresponding authors' emails: M.Michaelis@kent.ac.uk; M.D.Garrett@kent.ac.uk; M.N.Wass@kent.ac.uk

Abstract

Background: Here, we introduce a novel set of triple-negative breast cancer (TNBC) cell lines consisting of MDA-MB-468, HCC38, and HCC1806 and their sublines adapted to cisplatin, doxorubicin, eribulin, paclitaxel, gemcitabine, or 5-fluorouracil.

Methods: The cell lines were characterized by whole exome sequencing and the determination of drug-response profiles. Moreover, genes harbouring resistance-associated mutations were investigated using TCGA data for potential clinical relevance.

Results: Sequencing combined with TCGA-derived patient data resulted in the identification of 682 biomarker candidates in the pan-cancer analysis. Thirty-five genes were considered the most promising candidates because they harboured resistance-associated variants in at least two resistant sublines, and their expression correlated with TNBC patient survival. Exome sequencing and response profiles to cytotoxic drugs and DNA damage response inhibitors identified revealed remarkably little overlap between the resistant sublines, suggesting that each resistance formation process follows a unique route. All of the drug-resistant TNBC sublines remained sensitive or even displayed collateral sensitivity to a range of tested compounds. Cross-resistance levels were lowest for the CHK2 inhibitor CCT241533, the PLK1 inhibitor SBE13, and the RAD51 recombinase inhibitor B02, suggesting that CHK2, PLK1, and RAD51 are potential drug targets for therapy-refractory TNBC.

Conclusions: We present novel preclinical models of acquired drug resistance in TNBC and many novel candidate biomarkers for further investigation. The finding that each cancer cell line adaptation process follows an unpredictable route reflects

recent findings on cancer cell evolution in patients, supporting the relevance of drug-adapted cancer cell lines as preclinical models of acquired resistance.

Key words

Triple Negative Breast Cancer, acquired drug resistance, exome sequencing DNA repair, *de novo* variants, TCGA

Introduction

Triple-negative breast cancer (TNBC) is characterized by the absence of estrogen, progesterone, and HER2 receptors [1]. TNBC is responsible for approximately 15% of breast cancer cases and is associated with a poorer prognosis than hormone receptor- or HER2-positive breast cancers [1,2]. Current TNBC therapies are largely based on cytotoxic anticancer drugs, including platinum drugs, anthracyclins, eribulin, gemcitabine, paclitaxel, and 5-fluorouracil [1]. TNBC often responds well initially to cytotoxic chemotherapy, but recurrence and resistance are common, eventually leading to therapy failure. This combination of an initial high response rate followed by rapid resistance is referred to as the 'TNBC paradox' [1,3]. To improve TNBC therapy outcomes, new treatment approaches are needed, particularly those that are effective against treatment-refractory disease characterized by acquired resistance to cytotoxic chemotherapy.

In contrast to intrinsic drug resistance (which occurs independently of therapy and is a consequence of pre-existing often stochastic events in cancer cells), acquired resistance is the direct consequence of selection and adaptation processes caused by cancer treatment (directed tumor evolution) [4–8]. Understanding acquired resistance mechanisms is essential for optimizing cancer treatment for patients with therapy-refractory tumors.

Drug-adapted cancer cell lines are preclinical models that have been shown to reflect clinically relevant acquired drug resistance mechanisms in numerous studies [4,9–17]. Furthermore, drug-adapted cell lines enable detailed functional and systems-level studies that are not possible using clinical samples [4].

Here, we introduce a novel set of three parental TNBC cell lines and their 15 sublines adapted to cisplatin, doxorubicin, eribulin, gemcitabine, paclitaxel, or 5-

fluorouracil. These cell lines were characterized by whole exome sequencing and the determination of response profiles to cytotoxic anti-cancer drugs and a panel of DNA damage repair inhibitors. The resulting data showed that each resistance formation process follows an individual and unpredictable route. The combined analysis of resistance-associated mutations in combination with patient data from The Cancer Genome Atlas (TCGA) [18] identified 35 novel candidate resistance biomarkers for further investigation.

Results

Project cell line panel

Here, we characterized a cell line panel consisting of the parental TNBC cell lines MDA-MB-468, HCC38, and HCC1806 and their sublines adapted to grow in the presence of cisplatin, doxorubicin, eribulin, paclitaxel, gemcitabine, or 5-fluorouracil, which are all drugs used for the treatment of TNBC (Fig. 1A, Suppl. File 1) [19–25]. The drug-resistant sublines were established by continuous exposure to stepwise increasing drug concentrations as previously described [16]. All parental cell lines were initially sensitive to therapeutic concentrations of the respective drugs, as indicated by IC_{50} (concentration that reduces cell viability by 50%) values within the range of clinical drug plasma concentrations (C_{max}) (Suppl. Fig. 1, Suppl. File 1) [26]. The relative resistance factors (IC_{50} drug-adapted subline/ IC_{50} respective parental cell line) ranged from 5.5-fold (HCC38^{r}PCL$^{2.5}$) to 5916.7-fold (HCC1806^{r}ERI^{50}) (Fig. 1B, Suppl. File 1).

Characterization of the cell line panel by whole exome sequencing

The cell line panel was investigated by whole exome sequencing. Among the identified variants, missense variants were most common, followed by synonymous variants (Suppl. Fig. 2A). Insertions/deletions (INDELs), frameshift mutations, stop-gain, stop-loss, and splice variants were identified at lower frequencies (Suppl. Fig. 2A). Between 217 (HCC38^{r}DOX^{40}) and 952 (HCC38^{r}GEM^{20}) variants differed in the drug-adapted sublines relative to the respective parental cell lines (Suppl. Fig. 2B).

We grouped the resistance-associated variants into five categories (Fig. 2A, see methods): 1. *Gained variants*, variants only called in the drug-adapted subline but detectable at low confidence in the respective parental cell line; 2. *De novo*

variants, variants called in the drug-adapted subline but undetectable in the respective parental cell line; 3. *Not-called variants*, variants only called in the parental cell line but detectable with low confidence in the resistant subline; 4. *Lost variants*; variants called in the parental cell line but undetectable in the drug-adapted subline; and 5. *Shared variants*; variants called in both the parental and the respective drug-adapted sublines (Fig. 2A).

The number of *gained* variants ranged from 44 (HCC38^{r}DOX^{40}) to 381 (HCC38^{r}GEM^{20}), the number of *de novo* variants ranged from 31 (HCC38^{r}DOX^{40}) to 225 (MDA-MB-468^{r}PCL^{20}), the number of *not-called* variants ranged from 88 (HCC38^{r}GEM^{20} and HCC1806^{r}DOX$^{12.5}$) to 345 (MDA-MB-468^{r}PCL^{20}), and the number of *lost* variants ranged from 129 (HCC38^{r}GEM^{20}) to 398 (MDA-MB-468^{r}PCL^{20}) (Fig. 2B, Fig. 2C, Suppl. File.2 and 3). The number of *shared* variants that were both called in the parental cell lines and their sublines ranged from 128 (MDA-MB-468^{r}PCL^{20}) to 368 (HCC38^{r}GEM^{20}) (Fig. 2D, Suppl. File 2 and 3). The number of *shared* variants that increased by at least two-fold in the resistant sub-lines vs. the respective parental ranged from four (HCC1806^{r}5-F^{1500}) to 21 (MDA-MB-468^{r}CDDP^{1000}), whilst the number of shared variants that decreased by at least two-fold ranged from two (MDA-MB-468^{r}PCL^{20}) to 24 (HCC38^{r}GEM^{20}) (Fig. 2E, Suppl. File 2).

Analysis of the distribution of *de novo* variants

To identify variants that may have a functional role in drug resistance, we initially considered the 81 genes that harbored *de novo* variants in at least two different sublines from more than one parental cell line (Fig. 3A, Suppl. File 4). This list included 46 genes that have already been reported to be involved in drug

resistance in cancer and 33 new candidate genes with a possible role in drug resistance (Fig. 3A, Suppl. File 4). Notably, 24 of the 33 new candidate genes are reported to be relevant in cancer (Fig. 3A, Suppl. File 4).

Four of the five genes with the greatest number of *de novo* variants in the drug-adapted sublines were mucin (MUC) genes. *MUC6* had *de novo* variants in 15 sublines, *MUC2* in 14 sublines, *MUC4* in 13 sublines, and MUC16 in nine sublines (Fig. 3A, Suppl. File 4). The MUC genes are large genes that are known to be commonly mutated in cancer and have been reported to be involved in cancer cell drug resistance [27–31]. *De novo* mutations in CDC27, which has also been linked to drug resistance in cancer, were also detected in nine drug-resistant sublines [32,33] (Fig. 3A, Suppl. File 4).

GXYLT1, *KRTAP4-11*, and *RGPD4* were amongst those genes, which had not previously been associated with drug resistance in cancer that displayed *de novo* mutations in a high number (7) of drug-resistant sublines (Fig. 3A, Suppl. File 4). A GXYLT1 mutation promoted metastasis in colorectal cancer through MAPK signalling, a pathway known to confer resistance to a range of anti-cancer drugs [34–37]. *RGPD4* mutations are correlated with vascular invasion in HBV-associated hepatocellular carcinoma, and it is known that there is an overlap between pro-angiogenic, pro-metastatic, and resistance-associated signalling in cancer [35,38]. There is no known link between KRTAP4-11 and cancer, but *KRTAP4-11* expression levels have been reported to predict the methotrexate response in rheumatoid arthritis patients [38]. Hence, it seems plausible that the products of these genes may be involved in cancer cell drug resistance.

Taken together, our analysis identified 48 genes known to be involved in cancer cell drug resistance alongside 33 novel candidates potentially contributing to

171 therapy failure. Further research will be required to characterize the roles of these
172 individual genes in detail.

173 When we compared the overlaps between *de novo* variants shared between
174 sublines adapted to the same drug, the numbers were too small to draw any
175 meaningful conclusions (Fig. 3B, Suppl. Fig. 3A).

176 Notably, *de novo* variants in drug-resistant sublines may not always represent
177 actual novel variants that are selected because they contribute to cancer cell
178 resistance. Many apparent *de novo* mutations may have already been present in a
179 small fraction of the cells of the parental cell line but may not have been detected
180 due to the sequencing depth. Hence, overlaps in *de novo* variants between sublines
181 of the same parental cell line can also be used to indicate the levels of relatedness
182 between the founding subpopulations of the different resistant sublines.

183 Analysis of the *de novo* variants shared between the sublines from the same
184 parental cell line indicated the largest overlap. On average, there was a 22.6%
185 overlap among the HCC1806 sublines, followed by a 15.0% overlap among the
186 HCC38 sublines and a 7.7% overlap among the MDA-MB-468 sublines (Fig. 3C).
187 However, there were also noticeable differences in the overlaps between *de novo*
188 variants identified in each of the sublines from the same parental cell line. For
189 example, only three *de novo* variants were shared between HCC38^{r}CDDP^{3000} (out of
190 98 in total, 3.1%) and HCC38^{r}PCL$^{2.5}$ (out of 92 in total, 3.3%), while 53 variants were
191 shared between HCC38^{r}ERI^{10} (out of 131 in total, 40.5%) and HCC38^{r}GEM^{20} (out of
192 203 in total, 26.1%) (Fig. 3C, Suppl. Fig. 3B). These numbers suggest that there are
193 no pre-existing cell line subpopulations that are consistently selected in response to
194 anti-cancer drug treatment.
195

196 **Protein functions related to variants that changed in drug-resistant sublines**

197 Next, we used the Gene Ontology (GO) annotation to perform an analysis of
198 the protein functions associated with genes present in the *de novo*, *gained*, *not*
199 *called*, and *lost* variant sets as well as *shared* variants with a two-fold increase or
200 decrease in allele frequency (Suppl. Fig 4A, B).

201 There was limited overlap between the GO terms for the variants detected in
202 the sublines adapted to the same drug (Suppl. Fig. 4C, E). The extracellular matrix-
203 related GO terms 'extracellular matrix constituent lubricant activity', 'extracellular
204 matrix', and 'maintenance of gastrointestinal epithelium' were most common, which
205 reflects the high number of variants observed in the mucin genes (Suppl. Fig. 4C, E).

206 GO term analysis of the sublines from the same parental cell line revealed
207 very similar results, again revealing an overrepresentation of extracellular matrix-
208 related GO terms (Suppl. Fig. 4D, F). Further research will be required to investigate
209 the potential role of mucins and the extracellular matrix in acquired drug resistance in
210 TNBC cells.

211

212 **Potential clinical relevance of selected variants**

213 The potential clinical relevance of genes harboring *de novo*, *gained*, and
214 *shared* variants with a two-fold increase in allele frequency in the resistant subline as
215 well as genes harboring truncating variants was analysed using patient data derived
216 from The Cancer Genome Atlas (TCGA) [39]. Notably, there were only data available
217 from patients treated with cisplatin, doxorubicin, gemcitabine, paclitaxel, and/or 5-
218 fluorouracil, but no data on eribulin treatment were available.

219 We performed two analyses, one pan-cancer analysis, in which we
220 considered all patient survival data available for the drugs, and a second analysis, in

which we considered TNBC patients and for which only doxorubicin and paclitaxel data were available (Fig. 4A). The pan-cancer analysis included data from 29 TCGA cancer types for which mutation status and gene expression data were available (Fig 4B).

Six cases with at least two mutations in a resistance-associated gene were associated with patient prognosis (Suppl. File 5). As the number of resistance-associated genes with mutations in patients was low, we also considered gene expression status associations with prognosis. For 1,018 cases there was a significant association between gene expression and patient prognosis. This included genes, whose products were known to play a role in cancer cell drug resistance, such as CHEK2 [40–42] and APC [43–45] (Fig. 4C, Suppl. File 5). Moreover, we also identified novel candidates, which had not previously been suggested to be involved in cancer cell drug resistance, including KIAA2018, EYS, NBPF10, and KIAA0586 (Fig. 4C, Suppl. File 5).

We further determined the association of the expression of genes harboring *de novo*, *gained*, and *shared* variants with a two-fold increase as well as genes harboring truncating variants with patient survival in response to treatment with cisplatin, doxorubicin, gemcitabine, paclitaxel, and 5-fluorouracil (Fig. 4 D-E, Suppl. File 5). In total, the expression of 682 genes was significantly correlated with patient survival in response to at least one drug in the pan-cancer data. For 513 of these 682 genes, gene expression was associated with tumor response to the drug of the respective resistant subline (Suppl. File 5). The expression of 91 genes was associated with patient response to two drugs, the expression of 51 genes associated with response to three drugs, the expression of 21 genes associated with

expression to four drugs, and the expression of 6 genes associated with response to all five drugs (Fig. 4D, Suppl. File 5).

Considering the TNBC data alone, the expression of 165 genes was significantly correlated with patient survival in response to either doxorubicin, paclitaxel, or both drugs (Fig. 4E, Suppl. File 5). The expression of 141 of these 165 genes was associated with tumor response to the drug of the respective drug-adapted subline. The expression of 22 genes was associated with patient response to both doxorubicin and paclitaxel (Fig. 4E, Suppl. File 5).

Comparison of the analysis of the 165 genes identified in the TCGA analysis with the 81 genes identified in the analysis of *de novo* variants (Suppl. File 4) revealed 35 overlapping genes present in both datasets. This included 23 genes that have already been associated with drug resistance and 12 genes (*ABCD1, AGAP6, CUBN, DNAJC13, FLG, GXYLT1, KIAA0586, PABPC3, RGPD3, RGPD4, SETX and USP6*) that are novel findings (Suppl. File 6).

Complex sensitivity patterns of drug-resistant sublines to cytotoxic drugs

Determining drug sensitivity profiles in the cell line panel against the drugs of adaptation, i.e., cisplatin, doxorubicin, eribulin, paclitaxel, gemcitabine, and 5-fluorouracil (Fig. 5A, Suppl. File 1), revealed complex resistance patterns that did not follow clear, predictable rules. For example, two of the three doxorubicin-adapted sublines (HCC38^{r}DOX^{40} and HCC1806^{r}DOX$^{12.5}$) displayed increased (collateral) sensitivity to cisplatin compared to the parental cell line, while MDA-MB-468^{r}DOX^{50} displayed cross-resistance to cisplatin (Fig. 5A, Suppl. File 1). Moreover, all resistant sublines remained sensitive to or showed collateral sensitivity to at least one of the other chemotherapeutic agents (Fig. 5A, Suppl. File 1). The 5-fluorouracil-resistant

HCC1806^{r}5-F^{1500} subline was the only resistant subline that remained sensitive to all other investigated cytotoxic drugs (Fig. 5A, Suppl. File 1).

The ATP-binding cassette (ABC) transporter ABCB1 (also known as P-glycoprotein and MDR1) is an efflux transporter that mediates resistance to many anti-cancer drugs, including doxorubicin, eribulin, and paclitaxel [46]. Only five of the nine sublines adapted to the ABCB1 substrates doxorubicin, eribulin, and paclitaxel (including all three eribulin-resistant sublines) displayed cross-resistance to all other ABCB1 substrates. Among the ABCB1 substrate-adapted sublines, all the eribulin-adapted sublines displayed cross-resistance to paclitaxel, and all the paclitaxel-adapted sublines displayed cross-resistance to eribulin (Fig. 5A, Suppl. File 1). Notably, eribulin and paclitaxel are both tubulin-binding agents but differ in their mechanisms of interaction with tubulin. Eribulin is a destabilizing agent that binds to the vinca binding site of tubulin and inhibits microtubule formation, while paclitaxel is a stabilizing agent that binds to the taxane binding site that impairs microtubule degradation [47–51]. Further research will be required to determine to what extent the tubulin-binding agent cross-resistance profile of the tubulin-binding agent-adapted sublines is the consequence of the expression of ABCB1 (and/or other transporters), tubulin-related resistance mechanisms, or both.

Taken together, it is not possible to predict how resistance to a certain drug will affect the sensitivity patterns of the resulting sublines to other cytotoxic agents. However, all of the drug-resistant TNBC sublines remained sensitive and/or displayed collateral sensitivity to at least one of the tested anti-cancer drugs. Future research will be needed to elucidate the underlying mechanisms to identify biomarkers for personalized therapy approaches that can guide effective drugs to the right patients [4].

295

296 **Complex sensitivity patterns of drug-resistant sublines to DNA damage**
297 **response (DDR) inhibitors**

298 Triple-negative breast cancer cells have been shown to harbor defects in DNA
299 damage repair signalling, which can result in a dependence on the remaining intact
300 DNA damage repair (DDR) pathways and, in turn, in sensitivity to certain DDR
301 inhibitors [52]. Hence, we tested a panel of inhibitors targeting critical nodes of DDR
302 signalling in our novel resistant TNBC cell line panel (Fig. 5B).

303 All parental cell lines displayed sensitivity to the tested DDR inhibitors at
304 therapeutic concentrations, i.e., within the C_{max} values reported for these agents (for
305 which this information was available) (Suppl. Fig. 7). Similar to the results obtained
306 for the cytotoxic anti-cancer drugs, the DDR sensitivity profiles were complex and
307 unpredictable in the resistant sublines (Fig. 5C, Suppl. File 1). Relative to the
308 respective parental cell lines, the sensitivity remained unchanged for 128 DDR
309 inhibitor/ resistant subline combinations. Increased resistance (cross-resistance) was
310 detected in 96 DDR inhibitor/resistant subline combinations, and increased
311 sensitivity (collateral vulnerability) was recorded in 16 DDR inhibitor/resistant subline
312 combinations. Neither sublines of the same parental cell line nor sublines adapted to
313 the same drugs displayed substantial overlap in their DDR inhibitor sensitivity
314 profiles. Generally, cross-resistance levels were lowest for the CHK2 inhibitor
315 CCT241533, the PLK1 inhibitor SBE13, and the RAD51 recombinase inhibitor B02
316 among the investigated DDR inhibitors (Fig. 5C, Suppl. File 1).

317 Cross-resistance patterns were even inconsistent between DDR inhibitors
318 with the same targets. For example, different sensitivity patterns were observed
319 between the ATR inhibitors ceralasertib and berzosertib as well as the CHK1

inhibitors rabusertib, MK-8776, SRA737, and prexasertib (Fig. 5C, Suppl. File 1). The reasons for these differences are unclear. Notably, the activity of the DDR inhibitors may be modified by interactions with additional targets, and off-target resistance mechanisms (e.g., processes associated with drug uptake or efflux) may contribute to these differences [53].

In summary, and in line with the findings from the investigation of cytotoxic anti-cancer drugs, the drug-adapted TNBC sublines displayed complex, unpredictable sensitivity patterns against DDR inhibitors. This further demonstrates that improved future therapies will depend on an advanced understanding of the underlying molecular processes that enable the identification of biomarkers that can guide effective therapies for individual patients after treatment failure[4]. Notably, CHK2, PLK1, and RAD51 may have potential as new drug targets for the discovery and development of next-line therapies for TNBC patients whose tumors have stopped responding to chemotherapy.

Investigation of patterns in cell line drug response profiles

Finally, we used the delta (Δ) method to identify potential patterns in the response of the cell lines to all investigated cytotoxic anti-cancer drugs and DDR inhibitors [54]. The IC_{50} values were transformed to ΔIC_{50} values for each compound (see methods) and correlated across the drug panel using linear regression analysis and testing for statistical significance (Suppl. Table 1). Positive correlations indicate that increased drug resistance is seen with both agents, whilst negative correlations indicate that whilst increasing drug resistance is observed for one agent, collateral sensitivity is observed for the other agent. In the MDA-MB-468, HCC38, and

344 HCC1806 sublines, we observed 19, 20, and 60 positive correlations and 2, 8, and 1
345 negative correlation, respectively (Suppl. Table 1).

346 We were most interested in the agents that demonstrated negative
347 correlations, as they may identify potential next-line treatments. However, among the
348 11 negative correlations, there were no consistent results across the cell line panel
349 (Fig. 6). This further confirms that acquired resistance mechanisms are complex,
350 individual, and unpredictable and that the identification of potential next-line
351 therapies after treatment failure will depend on an improved understanding of cancer
352 cell evolution enabling therapy monitoring and biomarker-guided treatment
353 adaptation.

Discussion

In this study, we introduced and characterized a novel set of 15 sublines derived from the TNBC cell lines HCC38, HCC1806, and MDA-MB-468 that had been adapted to cisplatin, doxorubicin, eribulin, paclitaxel, gemcitabine, or 5-fluorouracil.

We applied whole exome sequencing to identify biomarker candidates to guide the use of anti-cancer therapies. In the first step, we focused on *de novo* mutations, i.e., mutations found in a resistant subline but undetectable in the respective parental cell line. Considering genes that displayed *de novo* mutations in at least two sublines of two different parental cell lines resulted in 81 resistance-associated variants, 48 of which were already known to be involved in cancer cell drug resistance, while 33 variants were novel.

In a second approach, we used TCGA data to investigate the potential clinical relevance of genes that harbored resistance-associated variants in the resistant sublines. In the pan-cancer dataset, the expression of 682 of these genes was correlated with patient survival in response to at least one of the investigated drugs. Considering only TNBC, the expression of 165 genes was significantly correlated with patient survival.

Comparison of the *de novo* variant analysis with the TNBC TCGA analysis identified 35 overlapping genes. Twenty-three of these genes are known to be associated with drug resistance. Twelve genes (*ABCD1, AGAP6, CUBN, DNAJC13, FLG, GXYLT1, KIAA0586, PABPC3, RGPD3, RGPD4, SETX and USP6*) are novel findings that may represent novel resistance biomarkers that have not been previously associated with drug resistance in cancer. Further research will be needed to investigate and define in more detail the role of these gene variants in

379 cancer therapy response and the expression of these genes as biomarkers for the
380 tailoring of personalized cancer therapies. Notably, numerous studies have shown
381 that drug-adapted cancer cell lines exhibit clinically relevant resistance mechanisms
382 [4,9–17].

383 Interestingly, the analysis of exome sequencing data revealed remarkably few
384 overlapping mutations between the investigated resistant sublines, including sublines
385 derived from the same parental cell line and sublines adapted to the same drug. This
386 suggests that resistance formation is the consequence of a complex, individual, and
387 unpredictable evolutionary process.

388 This complexity was confirmed by the determination of drug sensitivity profiles
389 to both cytotoxic anti-cancer drugs and DNA damage repair (DDR) inhibitors. Drug-
390 adapted sublines of the same parental cell line and sublines adapted to the same
391 drug displayed substantially different drug response patterns.

392 Notably, all the drug-adapted sublines remained sensitive and/or displayed
393 increased sensitivity (collateral vulnerability) to a range of tested compounds. This
394 suggests that it will be possible in the future to establish an improved understanding
395 of the processes underlying acquired resistance formation that result in the
396 identification of biomarkers that indicate effective next-line treatments for patients for
397 whom currently no effective treatment is available.

398 Among the investigated DDR inhibitors, the CHK2 inhibitor CCT241533, the
399 PLK1 inhibitor SBE13, and the RAD51 recombinase inhibitor B02 had the lowest
400 cross-resistance levels. Thus, CHK2, PLK1, and RAD51 are potential drug targets in
401 TNBC patients after failure of established therapies, particularly if reliable biomarkers
402 are found that identify cancer patients who are likely to benefit from such treatments.

Overall, the results from the characterization of the project cell line panel indicated that cancer cell resistance is a complex, individual, and unpredictable process. This finding is in agreement with data from studies in which cancer cell lines were repeatedly adapted to the same drug in independent experiments [8,16,55–57] and with recent findings from a comprehensive analysis of cancer cell evolution in lung cancer patients [58–62].

In conclusion, we present a novel set of drug-adapted TNBC cell lines as preclinical models of acquired drug resistance. Overlapping genes detected through the characterization of *de novo* variants and patient-derived TCGA data identified 35 biomarker candidates for the guidance of personalized TNBC therapies for further investigation, including 12 novel genes that have not been previously associated with drug resistance in cancer. Finally, our results show that each cancer cell line adaptation process follows an individual, unpredictable route, which reflects recent clinical findings from the monitoring of cancer cell evolution in patients [58–62]. This further supports the relevance of drug-adapted cancer cell lines as preclinical models of acquired resistance that can be analysed and manipulated at a level of detail that is impossible in the clinical setting.

Materials and Methods

Cell culture

MDA-MB-468, HCC38, and HCC1806 cells were obtained from the American Type Culture Collection (ATCC). The drug-adapted sublines (Fig. 1A, Suppl. File.1) were established by continuous exposure to stepwise increasing drug concentrations as previously described and derived from the Resistant Cancer Cell Line (RCCL) collection (https://research.kent.ac.uk/industrial-biotechnology-centre/the-resistant-cancer-cell-line-rccl-collection)[4,63]. All cell lines were cultured in Iscove's Modified Dulbecco's medium (IMDM) supplemented with 10% fetal bovine serum (Sigma□Aldrich, Germany), 2 mM L-glutamine, 25 mM HEPES (Fisher Scientific, UK), 100 IU/mL penicillin, and 100 µg/mL streptomycin (Life Technologies, UK) at 37 °C in a humidified atmosphere with 5% CO_2. Each drug-adapted subline was continuously cultured in the presence of the specific adaptation drug at a defined concentration, as indicated by the cell line name (ng/mL), e.g., MDA-MB-468^{r}DOX^{50}, where r = the resistant subline, Dox = doxorubicin and 50 = 50 ng/ml.

Compounds

The following compounds were purchased from the indicated suppliers: Adavosertib, Alisertib, Berzosertib, Ceralasertib, MK-8776, Olaparib, Prexasertib, Rabusertib, Rucaparib, SBE13, Tozasertib (Adooq Bioscience), AZD0156, BI2536, Doxorubicin, Gemcitabine (Selleckchem), B02, Cisplatin, 5-Fluorouracil (Sigma□Aldrich), CCT241533, SRA737 (a gift from the Institute of Cancer Research), Eribulin (Eisia), and Paclitaxel (Cayman Chemicals). All drug stocks were prepared in DMSO and stored at -20 °C, except for cisplatin, which was prepared in 0.9% NaCl solution and stored in the dark at room temperature.

446 **Cell growth and viability assays**

447 Cell viability was tested using the 3-(4,5-dimethylthiazol-2-yl)-2,5-
448 diphenyltetrazolium bromide (MTT) dye reduction assay after 120 hours of
449 incubation with each compound, modified as previously described [64,65].
450 Concentrations that reduced cell viability by 50% relative to an untreated control
451 (IC_{50}) were determined and used to calculate the resistance factor (RF; IC_{50} of drug-
452 adapted cell line/IC_{50} of respective parental cell line).

453

454 **Whole exome sequencing**

455 Whole exome sequencing (WES) libraries were prepared using the Nextera
456 Rapid Capture Exome Kit (Illumina). Sequencing was performed on a HiSeq 1500
457 platform in Rapid Run mode with 2 x 100 nucleotide paired-end reads. The two lanes
458 of the Rapid Run flow cell provided two sets of FASTQ data per cell line.

459

460 **Variant calling**

461 FASTQC was used to control the quality of the raw sequence data [66] prior to
462 the removal of sequencing adaptors. Trimmomatic (settings: NexteraPE-
463 PE.fa:2:30:10 LEADING:3 TRAILING:3 SLIDING WINDOW: 4:15 MILEN:36) [67]. Raw
464 FASTQ files were aligned to the human reference genome (GRCH37) using
465 Burrows□Wheeler Alignment (v.0.7.17) with an output in sequence alignment map
466 (SAM) format applying the default settings -M -R [68–70]. Only paired reads were used,
467 and Samtools flagstat was used to print statistics throughout each of the subsequent
468 steps [68]. SAM files were input into Picard tools SortSam (v.2.17.10), where the read
469 alignments were sorted by coordinate and converted to a binary alignment map
470 (BAM) format (Picard Toolkit.2019. Broad Institute, GitHub Repository.

http://broadinstitute.github.io/picard/; Broad Institute). Picard Tools MarkDuplicates (v2.17.10) was used for the removal of PCR duplicates (Picard Toolkit. 2019. Broad Institute, GitHub Repository. http://broadinstitute.github.io/picard/; Broad Institute). GenomeAnalysisTK-3.7.0 RealignerTargetCreator was used to perform base score recalibration, and GenomeAnalysisTK-3.7.0 IndelRealigner was used for INDEL realignment MAX_READS = 20000 [71]. SAMtools mpileup was used to generate binary variant call format (BCF) files from the BAM files, which were then input into BCFtools to call the SNVs and INDELS to generate a variant calling format (VCF) [72]. Variants were annotated with VEP [73].

Variant filtering

Only variants in coding regions of the genome were considered. To identify high-confidence variants, variants with a Phred score < 30, variants with fewer than 10 reads supporting the base call, or variants with < 3 reads supporting the variant were removed. Moreover, common variants with a frequency of ≥ 0.001% in the genome aggregation database (gnomAD) were removed [74]; if not, ≥ 3 samples were annotated in The Cancer Genome Atlas (TCGA), or ≥ 10 samples were annotated in the Catalogue Of Somatic Mutations In Cancer (COSMIC) [39,75,76].

Definition of variants

Gained variants: variants that are called in the drug-resistant subline and are called with low confidence in the parental cell line. *De novo* variants: variants that are called in the drug-resistant subline but not called in the parental cell line. *Not called* variants: variants that are called in the parental cell line but not called in the drug-resistant subline, even at low confidence. *Lost* variants: variants that are called in the parental cell line and are called in low confidence in the drug-resistant subline.

Shared variants: variants that are called in both the parental and drug-resistant sublines.

Gene Ontology

Gene Ontology (GO) functional enrichment analysis was conducted using G:profiler [77]. Gene lists were submitted as queries to the g:GOSt functional profiling tool and run at a significance threshold of g:SCS and a user threshold of 0.05.

The Cancer Genome Atlas (TCGA) analysis

TCGA data retrieval

The data were collected from the UCSC Xena functional genomics browser [https://xenabrowser.net]. Batch-corrected gene expression data (RNAseq, log2(normalized value + 1)) for 11,060 patients (version 2016-12-29), clinical data for 12,591 patients (version 2018-09-13), and somatic mutation data (HG19) for 9,104 patients (version 2016-12-29) were downloaded for the TCGA pancancer (PANCAN) cohort. Curated drug data were obtained from Moiso 2021 for 4,321 patients [78].

Final datasets

The 4 downloaded datasets were filtered to a final dataset for each drug for which every data type was available (gene expression, somatic mutation, clinical, and drug data). If a patient did not have at least 1 somatic mutation recorded, they were excluded from further somatic mutation analyses. This resulted in final datasets of 683 patients (23 cancer types) treated with cisplatin, 385 (17) with doxorubicin, 367 (11) with fluorouracil, 349 (20) with gemcitabine, and 544 (16) with paclitaxel for which somatic mutation and clinical data were available (table – “mutations/treatment_by_cancer_type_mutation_patients.tsv”). The gene expression

and treatment data included 765 patients (24 cancer types) treated with cisplatin, 571 (18) with doxorubicin, 452 (11) with fluorouracil, 438 (21) with gemcitabine, and 828 (15) with paclitaxel (table – "expression/treatment_by_cancer_type_expression_patients.tsv). Datasets including only TNBC patients were also created for further analysis[79]. This was only completed for those patients treated with doxorubicin (96 patients) and paclitaxel (63), as the number of patients treated with cisplatin (2), gemcitabine (4), and fluorouracil (21) was too low for meaningful analysis. For doxorubicin and paclitaxel treatments, gene expression and clinical data were available for 93 and 62 TNBC patients, respectively, while somatic mutation and clinical data were available for 74 doxorubicin- and 49 paclitaxel-treated patients. One TNBC patient (TCGA-AR-A256) whose disease-specific survival (DSS) data were incomplete was excluded.

Survival analysis

Analysis was performed in R version 4.3.0. Kaplan-Meier (KM) plots were generated for mutation status (mutated – MUT or wild type – WT) and for gene expression status (high or low) using the survival (v3.5-5) and survminer (0.4.9) packages. Somatic nonsynonymous mutations were considered in the genes of interest. The cut-off for high/low gene expression was calculated using the surv_cutpoint function in survminer, which makes use of the R package maxstat (v0.7-25). This gave a threshold for high/low expression based on the most significant relation with outcome, in this case, disease-specific survival. Any sample with gene expression > the calculated threshold was considered to have "high expression", and any sample with gene expression < the threshold was considered to have "low expression". The p value displayed on the KM plots was calculated using the log-rank test.

Statistical analysis and data manipulation

GraphPad Prism 6 (GraphPad Software, Inc., USA) was used to generate dose□response curves and determine IC_{50} values via nonlinear regression (with variable slopes). Statistical significance was calculated using a two-tailed t-test, assuming unequal variance, in GraphPad Prism 6 (GraphPad Software, Inc., USA).

The delta method was used as described by Bracht *et al.*, 2006 [54]. IC_{50} values were transformed to Δ IC_{50} values: Δ IC_{50} = *log (average* IC_{50} *in drug over all cell lines) – log (individual* IC_{50} *in drug for each cell line)*. Linear regression analysis of ΔIC_{50X} versus $\Delta IC_{50Y,}$ where X and Y represent two different compounds from the panel, was performed. The Pearson correlation coefficient (*r*) was used to establish the level of significance in a two-tailed test with (n-2) degrees of freedom, where $p \leq 0.05$.

Availability of data and materials

Data generated or analyzed during this study are included in this published article and its supplementary information files. The exome sequencing datasets generated and analyzed during the current study are available at Gene Expression Omnibus (GEO) repository (#PRJNA1155201).

Competing interests

Nothing to declare.

Funding

This work was supported by grants from the Frankfurter Stiftung für krebskranke Kinder, Kent Health, Kent Cancer Trust, and the Rosetrees Trust.

Authors contribution

HEG, MMi, MDG, and MNW conceived and designed the study. HG, MA, IGR, KMM, AN, MMe, and JC acquired data. All authors analyzed and interpreted data. HEG, MM, MDG, and MNW drafted the work. All authors substantively revised the work and approved the submitted version.

Acknowledgements

We would like to thank the Institute of Cancer Research for their kind gift of CCT241533 and SRA737. Figures were created using BioRender.com.

References

1. Gupta, G.K., Collier, A.L., Lee, D., Hoefer, R.A., Zheleva, V., van Reesema, L.L.S., Tang-Tan, A.M., Guye, M.L., Chang, D.Z., Winston, J.S., et al. (2020). Perspectives on triple-negative breast cancer: Current treatment strategies, unmet needs, and potential targets for future therapies. Cancers (Basel). *12*, 1–33. 10.3390/cancers12092392.

2. Hampton, J.M., Song, J., Jayasekera, J., Schechter, C., Alagoz, O., Stout, N.K., Trentham-dietz, A., Lee, S.J., Huang, X., Mandelblatt, J.S., et al. (2024). Analysis of Breast Cancer Mortality in the US—1975 to 2019. Jama *5405*, 233–241. 10.1001/jama.2023.25881.

3. Fornier, M., and Fumoleau, P. (2012). The paradox of triple negative breast cancer: Novel approaches to treatment. Breast J. *18*, 41–51. 10.1111/j.1524-4741.2011.01175.x.

4. Michaelis, M., Wass, M.N., and Cinatl, J. (2019). Drug-adapted cancer cell lines as preclinical models of acquired resistance. Cancer Drug Resist. *2*, 447–456. 10.20517/cdr.2019.005.

5. Oellerich, T., Schneider, C., Thomas, D., Knecht, K.M., Buzovetsky, O., Kaderali, L., Schliemann, C., Bohnenberger, H., Angenendt, L., Hartmann, W., et al. (2019). Selective inactivation of hypomethylating agents by SAMHD1 provides a rationale for therapeutic stratification in AML. Nat. Commun. *10*. 10.1038/s41467-019-11413-4.

6. Santoni-Rugiu, E., Melchior, L.C., Urbanska, E.M., Jakobsen, J.N., De Stricker, K., Grauslund, M., and Sørensen, J.B. (2019). Intrinsic resistance to EGFR-tyrosine kinase inhibitors in EGFR-mutant non-small cell lung cancer: Differences and similarities with acquired resistance. Cancers (Basel). *11*, 1–57. 10.3390/cancers11070923.

7. Touat, M., Li, Y.Y., Boynton, A.N., Spurr, L.F., Iorgulescu, J.B., Bohrson, C.L., Cortes-Ciriano, I., Birzu, C., Geduldig, J.E., Pelton, K., et al. (2020). Mechanisms and therapeutic implications of hypermutation in gliomas. Nature *580*, 517–523. 10.1038/s41586-020-2209-9.

8. Rothenburger, T., Thomas, D., Schreiber, Y., Wratil, P.R., Pflantz, T., Knecht, K., Digianantonio, K., Temple, J., Schneider, C., Baldauf, H.M., et al. (2021). Differences between intrinsic and acquired nucleoside analogue resistance in acute myeloid leukaemia cells. J. Exp. Clin. Cancer Res. *40*, 1–19. 10.1186/s13046-021-02093-4.

9. Juliano, R.L., and Ling, V. (1976). A surface glycoprotein modulating drug permeability in Chinese hamster ovary cell mutants. BBA - Biomembr. *455*, 152–162. 10.1016/0005-2736(76)90160-7.

10. Cole, S.P.C., Bhardwaj, G., Gerlach, J.H., Mackie, J.E., Grant, C.E., Almquist, K.C., Stewart, A.J., Kurz, E.U., Duncan, A.M.V., and Deeley, R.G. (1992). Overexpression of a transporter gene in a multidrug-resistant human lung cancer cell line. Science (80-.). *258*, 1650–1654. 10.1126/science.1360704.

11. Engelman, J., Zejnullahu, K., Mitsudomi, T., Song, Y., and Hyland (2007). MET amplification leads to gefitinib resistance in lung cancer by activating ERBB3 signaling. Science (80-.). *316*, 1039–1043.

12. Sharma, S. V., Haber, D.A., and Settleman, J. (2010). Cell line-based platforms to evaluate the therapeutic efficacy of candidate anticancer agents. Nat. Rev. Cancer *10*, 241–253. 10.1038/nrc2820.

13. Nazarian, R., Shi, H., Wang, Q., Kong, X., Koya, R.C., Lee, H., Chen, Z., Lee, M.K., Attar, N., Sazegar, H., et al. (2010). Melanomas acquire resistance to B-RAF(V600E) inhibition by RTK or N-RAS upregulation. Nature *468*, 973–977. 10.1038/nature09626.

14. Crystal, A.S., Shaw, A.T., Sequist, L. V., Friboulet, L., Niederst, M.J., Lockerman, E.L., Frias, R.L., Gainor, J.F., Amzallag, A., Greninger, P., et al. (2014). Patient-derived models of acquired resistance can identify effective drug combinations for cancer. Science (80-.). *346*, 1480–1486. 10.1126/science.1254721.

15. Schneider, C., Oellerich, T., Baldauf, H.M., Schwarz, S.M., Thomas, D., Flick, R., Bohnenberger, H., Kaderali, L., Stegmann, L., Cremer, A., et al. (2017). SAMHD1 is a biomarker for cytarabine response and a therapeutic target in acute myeloid leukemia. Nat. Med. *23*, 250–255. 10.1038/nm.4255.

16. Michaelis, M., Rothweiler, F., Barth, S., Cinat, J., Van Rikxoort, M., Löschmann, N., Voges, Y., Breitling, R., Von Deimling, A., Rödel, F., et al. (2011). Adaptation of cancer cells from different entities to the MDM2 inhibitor nutlin-3 results in the emergence of p53-mutated multi-drug-resistant cancer cells. Cell Death Dis. *2*. 10.1038/cddis.2011.129.

17. Berlak, M., Tucker, E., Dorel, M., Winkler, A., McGearey, A., Rodriguez-Fos, E., da Costa, B.M., Barker, K., Fyle, E., Calton, E., et al. (2022). Mutations in ALK signaling pathways conferring resistance to ALK inhibitor treatment lead to collateral vulnerabilities in neuroblastoma cells. Mol. Cancer *21*, 1–19. 10.1186/s12943-022-01583-z.

18. Weinstein, J.N., Collisson, E.A., Mills, G.B., Shaw, K.R.M., Ozenberger, B.A., Ellrott, K., Sander, C., Stuart, J.M., Chang, K., Creighton, C.J., et al. (2013). The cancer genome atlas pan-cancer analysis project. Nat. Genet. *45*, 1113–1120. 10.1038/ng.2764.

19. Klaas, E., Sung, E., Azizi, E., Martinez, M., Barpujari, A., Roberts, J., and

Lucke-Wold, B. (2023). Advanced breast cancer metastasized in the brain: treatment standards and innovations. J. Cancer Metastasis Treat. *9*. 10.20517/2394-4722.2022.125.

20. Wang, B., Sun, T., Zhao, Y., Wang, S., Zhang, J., Wang, Z., Teng, Y.E., Cai, L., Yan, M., Wang, X., et al. (2022). A randomized phase 3 trial of Gemcitabine or Nab-paclitaxel combined with cisPlatin as first-line treatment in patients with metastatic triple-negative breast cancer. Nat. Commun. *13*. 10.1038/s41467-022-31704-7.

21. Liu, Y., Fan, L., Wang, Z.H., and Shao, Z.M. (2023). Nab-paclitaxel Followed by Dose-dense Epirubicin/Cyclophosphamide in Neoadjuvant Chemotherapy for Triple-negative Breast Cancer: A Phase II Study. Oncologist *28*, 86-e76. 10.1093/oncolo/oyac223.

22. Gumusay, O., Huppert, L.A., Magbanua, M.J.M., Wabl, C.A., Assefa, M., Chien, A.J., Melisko, M.E., Majure, M.C., Moasser, M., Park, J., et al. (2023). A phase Ib/II study of eribulin in combination with cyclophosphamide in patients with advanced breast cancer. Breast Cancer Res. Treat. *203*, 197–204. 10.1007/s10549-023-07073-0.

23. Kim, S.H., Im, S.A., Suh, K.J., Lee, K.H., Kim, M.H., Sohn, J., Park, Y.H., Kim, J.Y., Jeong, J.H., Lee, K.E., et al. (2023). Clinical activity of nivolumab in combination with eribulin in HER2-negative metastatic breast cancer: A phase IB/II study (KCSG BR18-16). Eur. J. Cancer *195*, 113386. 10.1016/j.ejca.2023.113386.

24. Velikova, G., Morden, J.P., Haviland, J.S., Emery, C., Barrett-Lee, P., Earl, H., Bloomfield, D., Brunt, A.M., Canney, P., Coleman, R., et al. (2023). Accelerated versus standard epirubicin followed by cyclophosphamide, methotrexate, and fluorouracil or capecitabine as adjuvant therapy for breast cancer (UK TACT2; CRUK/05/19): quality of life results from a multicentre, phase 3, open-label, randomised, . Lancet Oncol. *24*, 1359–1374. 10.1016/S1470-2045(23)00460-6.

25. Takahashi, M., Cortés, J., Dent, R., Pusztai, L., McArthur, H., Kümmel, S., Denkert, C., Park, Y.H., Im, S.A., Ahn, J.H., et al. (2023). Pembrolizumab Plus Chemotherapy Followed by Pembrolizumab in Patients with Early Triple-Negative Breast Cancer: A Secondary Analysis of a Randomized Clinical Trial. JAMA Netw. Open *6*, E2342107. 10.1001/jamanetworkopen.2023.42107.

26. Liston, D.R., and Davis, M. (2017). Clinically Relevant Concentrations of Anticancer Drugs□: A Guide for Nonclinical Studies. Clin. Cancer Res., 3489–3498. 10.1158/1078-0432.CCR-16-3083.

27. Kim, N., Hong, Y., Kwon, D., and Yoon, S. (2013). Somatic Mutaome Profile in

693 Human Cancer Tissues. Genomics Inform. *11*, 239. 10.5808/gi.2013.11.4.239.

694 28. Reynolds, I.S., Fichtner, M., McNamara, D.A., Kay, E.W., Prehn, J.H.M., and
695 Burke, J.P. (2019). Mucin glycoproteins block apoptosis; promote invasion,
696 proliferation, and migration; and cause chemoresistance through diverse
697 pathways in epithelial cancers. Cancer Metastasis Rev. *38*, 237–257.
698 10.1007/s10555-019-09781-w.

699 29. Pandey, K., Lee, E., Park, N., Hur, J., Cho, Y. Bin, Katuwal, N.B., Kim, S.K.,
700 Lee, S.A., Kim, I., An, H.J., et al. (2021). Deregulated immune pathway
701 associated with palbociclib resistance in preclinical breast cancer models:
702 Integrative genomics and transcriptomics. Genes (Basel). *12*, 1–14.
703 10.3390/genes12020159.

704 30. Chang, Y., Wang, Y., Li, B., Lu, X., Wang, R., Li, H., Yan, B., Gu, A., Wang,
705 W., Huang, A., et al. (2021). Whole-Exome Sequencing on Circulating Tumor
706 Cells Explores Platinum-Drug Resistance Mutations in Advanced Non-small
707 Cell Lung Cancer. Front. Genet. *12*, 1–11. 10.3389/fgene.2021.722078.

708 31. Patel, N.M., Geropoulos, G., Patel, P.H., Bhogal, R.H., Harrington, K.J.,
709 Singanayagam, A., and Kumar, S. (2023). The Role of Mucin Expression in the
710 Diagnosis of Oesophago-Gastric Cancer: A Systematic Literature Review.
711 Cancers (Basel). *15*. 10.3390/cancers15215252.

712 32. Kim, S.H., Ho, J.N., Jin, H., Lee, S.C., Lee, S.E., Hong, S.K., Lee, J.W., Lee,
713 E.S., and Byun, S.S. (2016). Upregulated expression of BCL2, MCM7, and
714 CCNE1 indicate cisplatin-resistance in the set of two human bladder cancer
715 cell lines: T24 cisplatin sensitive and T24R2 cisplatin resistant bladder cancer
716 cell lines. Investig. Clin. Urol. *57*, 63–72. 10.4111/icu.2016.57.1.63.

717 33. Feng, Z., Zhang, L., Zhou, J., Zhou, S., Li, L., Guo, X., Feng, G., Ma, Z.,
718 Huang, W., and Huang, F. (2017). mir-218-2 promotes glioblastomas growth,
719 invasion and drug resistance by targeting CDC27. Oncotarget *8*, 6304–6318.
720 10.18632/oncotarget.13850.

721 34. Peng, L., Zhao, M., Liu, T., Chen, J., Gao, P., Chen, L., Xing, P., Wang, Z., Di,
722 J., Xu, Q., et al. (2022). A stop-gain mutation in GXYLT1 promotes metastasis
723 of colorectal cancer via the MAPK pathway. Cell Death Dis. *13*, 1–12.
724 10.1038/s41419-022-04844-3.

725 35. Michaelis, M., Klassert, D., Barth, S., Suhan, T., Breitling, R., Mayer, B.,
726 Hinsch, N., Doerr, H.W., Cinatl, J., and Cinatl, J. (2009). Chemoresistance
727 acquisition induces a global shift of expression of aniogenesis-associated
728 genes and increased pro-angogenic activity in neuroblastoma cells. Mol.
729 Cancer *8*, 80. 10.1186/1476-4598-8-80.

730 36. Bahar, M.E., Kim, H.J., and Kim, D.R. (2023). Targeting the RAS/RAF/MAPK

pathway for cancer therapy: from mechanism to clinical studies. Signal Transduct. Target. Ther. *8*. 10.1038/s41392-023-01705-z.

37. Wang, P., Laster, K., Jia, X., Dong, Z., and Liu, K. (2023). Targeting CRAF kinase in anti-cancer therapy: progress and opportunities. Mol. Cancer *22*, 1–34. 10.1186/s12943-023-01903-x.

38. Xu, J., Zhou, Y., Dong, K., Gong, J., Xiong, W., Wang, X., Gu, C., Lu, X. yu, Huang, D. pei, Shen, X. dong, et al. (2023). Gene variation profile and it's potential correlation with clinical characteristics in HBV-associated HCC patients of Sichuan Han nationality in China. Asian J. Surg. *46*, 4371–4377. 10.1016/j.asjsur.2023.02.056.

39. Weinstein, J.N. (2013). Cancer Genome Atlas Pan-cancer analysis project. Nat Genet *45*, 113–1120. 10.3779/j.issn.1009-3419.2015.04.02.

40. Xu, H., Cheung, I.Y., Wei, X.X., Tran, H., Gao, X., and Cheung, N.K. V. (2011). Checkpoint kinase inhibitor synergizes with DNA-damaging agents in G 1 checkpoint-defective neuroblastoma. Int. J. Cancer *129*, 1953–1962. 10.1002/ijc.25842.

41. Ling, V.Y., Straube, J., Godfrey, W., Haldar, R., Janardhanan, Y., Cooper, L., Bruedigam, C., Cooper, E., Tavakoli Shirazi, P., Jacquelin, S., et al. (2023). Targeting cell cycle and apoptosis to overcome chemotherapy resistance in acute myeloid leukemia. Leukemia *37*, 143–153. 10.1038/s41375-022-01755-2.

42. Zeng, L., Nikolaev, A., Xing, C., della Manna, D.L., and Yang, E.S. (2020). CHK1/2 inhibitor prexasertib suppresses notch signaling and enhances cytotoxicity of cisplatin and radiation in head and neck squamous cell carcinoma. Mol. Cancer Ther. *19*, 1279–1288. 10.1158/1535-7163.MCT-19-0946.

43. VanKlompenberg, M.K., Bedalov, C.O., Soto, K.F., and Prosperi, J.R. (2015). APC selectively mediates response to chemotherapeutic agents in breast cancer. BMC Cancer *15*, 1–14. 10.1186/s12885-015-1456-x.

44. Stefanski, C.D., Keffler, K., McClintock, S., Milac, L., and Prosperi, J.R. (2019). APC loss affects DNA damage repair causing doxorubicin resistance in breast cancer cells. Neoplasia (United States) *21*, 1143–1150. 10.1016/j.neo.2019.09.002.

45. Astarita, E.M., Maloney, S.M., Hoover, C.A., Berkeley, B.J., Van Klompenberg, M.K., Murlidharan Nair, T., and Prosperi, J.R. (2021). Adenomatous Polyposis Coli loss controls cell cycle regulators and response to paclitaxel in MDA-MB-157 metaplastic breast cancer cells. PLoS One *16*, 1–18. 10.1371/journal.pone.0255738.

769 46. Szakács, G., Paterson, J.K., Ludwig, J.A., Booth-Genthe, C., and Gottesman,
770 M.M. (2006). Targeting multidrug resistance in cancer. Nat. Rev. Drug Discov.
771 *5*, 219–234. 10.1038/nrd1984.

772 47. Lu, J.F., Pokharel, D., and Bebawy, M. (2015). MRP1 and its role in anticancer
773 drug resistance. Drug Metab. Rev. *47*, 406–419.
774 10.3109/03602532.2015.1105253.

775 48. Derry, W.B., Wilson, L., and Jordan, M.A. (1995). Substoichiometric Binding of
776 Taxol Suppresses Microtubule Dynamics. Biochemistry *34*, 2203–2211.
777 10.1021/bi00007a014.

778 49. Snyder, J.P., Nettles, J.H., Cornett, B., Downing, K.H., and Nogales, E. (2001).
779 The binding conformation of Taxol in β-tubulin: A model based on electron
780 crystallographic density. Proc. Natl. Acad. Sci. U. S. A. *98*, 5312–5316.
781 10.1073/pnas.051309398.

782 50. Jordan, M.A., and Wilson, L. (2004). Microtubules as a target for anticancer
783 drugs. Nat. Rev. Cancer *4*, 253–265. 10.1038/nrc1317.

784 51. Smith, J.A., Wilson, L., Azarenko, O., Zhu, X., Lewis, B.M., Littlefield, B.A., and
785 Jordan, M.A. (2010). Eribulin binds at microtubule ends to a single site on
786 tubulin to suppress dynamic instability. Biochemistry *49*, 1331–1337.
787 10.1021/bi901810u.

788 52. Jin, J., Tao, Z., Cao, J., Li, T., and Hu, X. (2021). DNA damage response
789 inhibitors: An avenue for TNBC treatment. Biochim. Biophys. Acta - Rev.
790 Cancer *1875*, 188521. 10.1016/j.bbcan.2021.188521.

791 53. Baxter, J.S., Zatreanu, D., Pettitt, S.J., and Lord, C.J. (2022). Resistance to
792 DNA repair inhibitors in cancer. Mol. Oncol. *16*, 3811–3827. 10.1002/1878-
793 0261.13224.

794 54. Bracht, K., Boubakari, Grünert, R., and Bednarski, P.J. (2006). Correlations
795 between the activities of 19 anti-tumor agents and the intracellular glutathione
796 concentrations in a panel of 14 human cancer cell lines: Comparisons with the
797 National Cancer Institute data. Anticancer. Drugs *17*, 41–51.
798 10.1097/01.cad.0000190280.60005.05.

799 55. Michaelis, M., Schneider, C., Rothweiler, F., Rothenburger, T., Mernberger,
800 M., Nist, A., von Deimling, A., Speidel, D., Stiewe, T., and Cinatl, J. (2018).
801 TP53 mutations and drug sensitivity in acute myeloid leukaemia cells with
802 acquired MDM2 inhibitor resistance. bioRxiv, 404475.

803 56. Michaelis, M., Wass, M.N., Reddin, I., Voges, Y., Rothweiler, F., Hehlgans, S.,
804 Cinatl, J., Mernberger, M., Nist, A., Stiewe, T., et al. (2020). YM155-adapted
805 cancer cell lines reveal drug- induced heterogeneity and enable the
806 identification of biomarker candidates for the acquired resistance setting.

Cancers (Basel). *12*, 1–17. 10.3390/cancers12051080.

57. Hata, A.N., Niederst, M.J., Archibald, H.L., Gomez-Caraballo, M., Siddiqui, F.M., Mulvey, H.E., Maruvka, Y.E., Ji, F., Bhang, H.E.C., Radhakrishna, V.K., et al. (2016). Tumor cells can follow distinct evolutionary paths to become resistant to epidermal growth factor receptor inhibition. Nat. Med. *22*, 262–269. 10.1038/nm.4040.

58. Karasaki, T., Moore, D.A., Veeriah, S., Naceur-Lombardelli, C., Toncheva, A., Magno, N., Ward, S., Bakir, M. Al, Watkins, T.B.K., Grigoriadis, K., et al. (2023). Evolutionary characterization of lung adenocarcinoma morphology in TRACERx. Nat. Med. *29*. 10.1038/s41591-023-02230-w.

59. Martínez-Ruiz, C., Black, J.R.M., Puttick, C., Hill, M.S., Demeulemeester, J., Larose Cadieux, E., Thol, K., Jones, T.P., Veeriah, S., Naceur-Lombardelli, C., et al. (2023). Genomic–transcriptomic evolution in lung cancer and metastasis. Nature *616*, 543–552. 10.1038/s41586-023-05706-4.

60. Al Bakir, M., Huebner, A., Martínez-Ruiz, C., Grigoriadis, K., Watkins, T.B.K., Pich, O., Moore, D.A., Veeriah, S., Ward, S., Laycock, J., et al. (2023). The evolution of non-small cell lung cancer metastases in TRACERx 10.1038/s41586-023-05729-x.

61. Frankell, A.M., Dietzen, M., Al Bakir, M., Lim, E.L., Karasaki, T., Ward, S., Veeriah, S., Colliver, E., Huebner, A., Bunkum, A., et al. (2023). The evolution of lung cancer and impact of subclonal selection in TRACERx. Nature *616*, 525–533. 10.1038/s41586-023-05783-5.

62. Abbosh, C., Frankell, A.M., Harrison, T., Kisistok, J., Garnett, A., Johnson, L., Veeriah, S., Moreau, M., Chesh, A., Chaunzwa, T.L., et al. (2023). Tracking early lung cancer metastatic dissemination in TRACERx using ctDNA. Nature *616*, 553–562. 10.1038/s41586-023-05776-4.

63. Michaelis, M., Rothweiler, F., Barth, S., Cinat, J., Van Rikxoort, M., Löschmann, N., Voges, Y., Breitling, R., Von Deimling, A., Rödel, F., et al. (2011). Adaptation of cancer cells from different entities to the MDM2 inhibitor nutlin-3 results in the emergence of p53-mutated multi-drug-resistant cancer cells. Cell Death Dis. *2*. 10.1038/cddis.2011.129.

64. Mosmann, T. (1983). Rapid colorimetric assay for cellular growth and survival: Application to proliferation and cytotoxicity assays. J. Immunol. Methods *65*, 55–63. 10.1016/0022-1759(83)90303-4.

65. Onafuye, H., Pieper, S., Mulac, D., Cinatl, J., Wass, M.N., Langer, K., and Michaelis, M. (2019). Doxorubicin-loaded human serum albumin nanoparticles overcome transporter-mediated drug resistance in drug-adapted cancer cells. Beilstein J. Nanotechnol. *10*, 1707–1715. 10.3762/bjnano.10.166.

66. Andrews S (2018). FastQC A Quality control tool for high throughput sequence data. Babraham Bioinfo, 3–5.

67. Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. Bioinformatics *30*, 2114–2120. 10.1093/bioinformatics/btu170.

68. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics *25*, 2078–2079. 10.1093/bioinformatics/btp352.

69. Burrows, M., and Wheeler, D. (1994). A block-sorting lossless data compression algorithm. Algorithm, Data Compression, 18. 10.1.1.37.6774.

70. Church, D.M., Schneider, V.A., Graves, T., Auger, K., Cunningham, F., Bouk, N., Chen, H.C., Agarwala, R., McLaren, W.M., Ritchie, G.R.S., et al. (2011). Modernizing reference genome assemblies. PLoS Biol. *9*. 10.1371/journal.pbio.1001091.

71. McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., et al. (2010). The genome analysis toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. *20*, 1297–1303. 10.1101/gr.107524.110.

72. Li, H. (2011). A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. Bioinformatics *27*, 2987–2993. 10.1093/bioinformatics/btr509.

73. McLaren, W., Gil, L., Hunt, S.E., Riat, H.S., Ritchie, G.R.S., Thormann, A., Flicek, P., and Cunningham, F. (2016). The Ensembl Variant Effect Predictor. Genome Biol. *17*, 1–14. 10.1186/s13059-016-0974-4.

74. Karczewski, K.J., Francioli, L.C., Tiao, G., Cummings, B.B., Alföldi, J., Wang, Q., Collins, R.L., Laricchia, K.M., Ganna, A., Birnbaum, D.P., et al. (2019). Variation across 141,456 human exomes and genomes reveals the spectrum of loss-of-function intolerance across human protein-coding genes. bioRxiv, 531210. 10.1101/531210.

75. Ghandi, M., Huang, F.W., Jané-Valbuena, J., Kryukov, G. V., Lo, C.C., McDonald, E.R., Barretina, J., Gelfand, E.T., Bielski, C.M., Li, H., et al. (2019). Next-generation characterization of the Cancer Cell Line Encyclopedia. Nature. 10.1038/s41586-019-1186-3.

76. Bamford, S., Dawson, E., Forbes, S., Clements, J., Pettett, R., Dogan, A., Flanagan, A., Teague, J., Futreal, P.A., Stratton, M.R., et al. (2004). The COSMIC (Catalogue of Somatic Mutations in Cancer) database and website. Br. J. Cancer *2*, 355–358. 10.1038/sj.bjc.6601894.

77. Raudvere, U., Kolberg, L., Kuzmin, I., Arak, T., Adler, P., Peterson, H., and Vilo, J. (2019). g:Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). Nucleic Acids Res. *47*, W191–W198. 10.1093/nar/gkz369.

78. Moiso, E. (2021). Manual curation of TCGA treatment data and identification of potential markers of therapy response. medRxiv, 2021.04.30.21251941.

79. Lehmann, B.D., Colaprico, A., Silva, T.C., Chen, J., An, H., Ban, Y., Huang, H., Wang, L., James, J.L., Balko, J.M., et al. (2021). Multi-omics analysis identifies therapeutic vulnerabilities in triple-negative breast cancer subtypes. Nat. Commun. *12*, 1–18. 10.1038/s41467-021-26502-6.

Figure Legends

Figure 1. Confirmation of the resistance status of the project cell lines. A) Panel of drug-naïve (MDA-MB-468, HCC38, HCC1806) and drug-adapted Triple Negative Breast Cancer cell lines. B) Left: dose–response curve; bottom: IC_{50} values; right: resistance factor ($IC_{50,}$ drug-adapted subline/$IC_{50,}$ respective parental cell line)); when drug-naïve and drug-adapted cell lines are treated with the respective agent: cisplatin, doxorubicin, eribulin, paclitaxel, gemcitabine, or 5-fluorouracil. Circles indicate drug-naïve cell lines, and crosses indicate drug-adapted cell lines. Green, MDA-MB-468-derived; blue, HCC38-derived; orange, HCC1806-derived. The data are from ≥ 3 independent experiments, and the statistics were calculated using Student's t-test and are plotted as the means ± SDs.

Figure 2. Genomic characterization of drug-adapted cell lines. A) Diagram illustrating the differences between *gained, de novo*, *not called*, *lost* and *shared* variants. B) Count of *Gained* (blue) and *De novo* (green) variants, C) count of *Lost* (orange) and *Not-called* (pink) variants, D) left panel; count of all *Shared* (purple) variants, right panel; two-fold increase or decrease of shared variants.

Figure 3. Identification of novel candidates associated with therapy failure. A) Flow chart of genes with *de novo* variants observed in two or more sublines from more than one parental cell line. B) Venn diagrams of *de novo* variants shared between sublines adapted to the same drug. C) Summary of relatedness between sublines drug-adapted from the same parental cell line (%).

Figure 4. Tumor patient data available for mutations in resistance-associated genes. (A) TCGA pan-cancer datasets for mutation status and gene expression. Only patients for which clinical, drug and mutation status/gene expression data was available for were considered in the TCGA pan-cancer analysis. (B) TCGA pan-cancer mutation status and expression data available for chemotherapy drugs for 29 TCGA cancer classifications. (C) Kaplan-Meier plots for gene expression with most significant association with prognosis in the pan-cancer dataset. Log-rank test was the statistical test used with multiple test correction performed using Benjamini-Hochberg method. (D-E) Genes for which expression is significantly associated with

patient prognosis. Upset plot showing the number of genes that are associated with patient prognosis for (D) pan-cancer and (E) TNBC.

Figure 5. Complex sensitivity patterns to cytotoxic and DDR-targeted agents. A) Heatmap of fold resistance and collateral sensitivity to cytotoxic agents. B) Summary of pathways targeted by DNA damage response and repair (DDRR) inhibitors used in screening. C) Heatmap of fold change resistance and collateral sensitivity to DDRR inhibitors.

Figure 6. Lack of trends in drug or inhibitor sensitivity patterns. Graphs demonstrating a negative correlation; collateral sensitivity to one agent but resistance to the other (blue); positive correlation; resistance to both agents (red); and no statistical correlation (black) for each set of sublines adapted from the MDA-MB-468, HCC38 or HCC1806 TNBC cell lines.

Supplementary Figure 1. Chemo-naïve cell lines are clinically sensitive to chemotherapy agents. IC_{50} values of drug-naïve parental cell lines treated with the respective chemotherapy agents: cisplatin, doxorubicin, eribulin, paclitaxel, gemcitabine or 5-fluorouracil. Green, MDA-MB-468 cells; blue, HCC38 cells; orange, HCC1806 cells. The black line indicates known C_{max} values for each chemotherapy agent. Data from $n \geq 3$, statistics were calculated using Student's t-test and are plotted as the mean ± SD.

Supplementary Figure 2: Variant counts. A) Total number of variants called for in the panel of drug-naïve and drug-resistant cell lines. B) Types of variants called for in the panel of drug-naïve and drug-resistant cell lines, including missense, synonymous, frameshift, inframe insertion, inframe deletion, stop loss, stop gain, splice acceptor and splice donor variants.

Supplementary Figure 3. *De novo* variant overlaps. The number of *de novo* variants that overlap in A) drug-resistant cell lines adapted to the same chemotherapy drug and B) drug-resistant cell lines adapted from the same parental cell line but to different chemotherapy drugs.

Supplementary Figure 4. Gene ontology terms related to variants in drug-resistant sublines. A) The number of variants increased in drug-resistant sublines (*de novo* variants, *gained* variants and *shared* variants that demonstrated a ≥2 increase in variant allele frequency). B) The number of variants decreased in drug-resistant sublines (*not-called* variants, *lost* variants and *shared* variants that demonstrated ≤2 decreases in variant allele frequency). The number and overlapping terms found in increased and decreased variants were compared between cell lines adapted to the same chemotherapy drug (C, E) and sublines derived from the same parental cell line but adapted to different chemotherapy drugs (D, F). Green bars indicate increased variants (A, C, D), and red bars indicate decreased variants (B, C, D).

Supplementary Figure 5. Chemo-naïve cell lines are clinically sensitive to DNA damage response and repair (DDRR) inhibitors. IC_{50} values of drug-naïve cell lines treated with the indicated drug. Green, MDA-MB-468-derived; blue, HCC38-derived; orange, HCC1806-derived. The black line indicates known C_{max} values for each DDRR agent. The data are from ≥ 3 independent experiments, and the statistics were calculated using Student's t-test and are plotted as the means ± SDs.

Supplementary Table 1. Drug correlation of delta (Δ) values. The IC_{50} values were transformed to ΔIC_{50} values for each drug (see methods) and correlated across the drug panel, with linear regression analysis and statistical significance. The values in the table indicate the r values of the correlations, where positive values indicate positive correlations and negative values indicate negative correlations. P values of the correlations are indicated in the blue color scheme, with light blue ($p \leq 0.05$) indicating the lowest statistical significance and dark blue ($p \leq 0.00001$) indicating the highest statistical significance.

Supplementary File 1. Mean IC_{50} values, SDs and resistance factors for the project panel treated with chemotherapy drugs and DNA damage response inhibitors.

Supplementary File 2. Basic variant characterization of the cell line panel.

Supplementary File 3. Variants found to be *de novo*, *gained*, *not called*, *lost* and shared in drug-resistant cell lines.

Supplementary File 4. List of genes with *de novo* variants in ≥2 drug-resistant cell lines. The values in the table indicate the variant allele frequencies of the *de novo* variants identified in the indicated genes. PMIDs for genes previously implicated in cancer and drug resistance.

Supplementary File 5. Step-by-step analysis of both TNBC and pan-cancer patient data extracted from the TCGA.

Supplementary File 6. Comparison of genes identified through *de novo* variant analysis and TCGA analysis.

1018

1019

1020

Figure 1

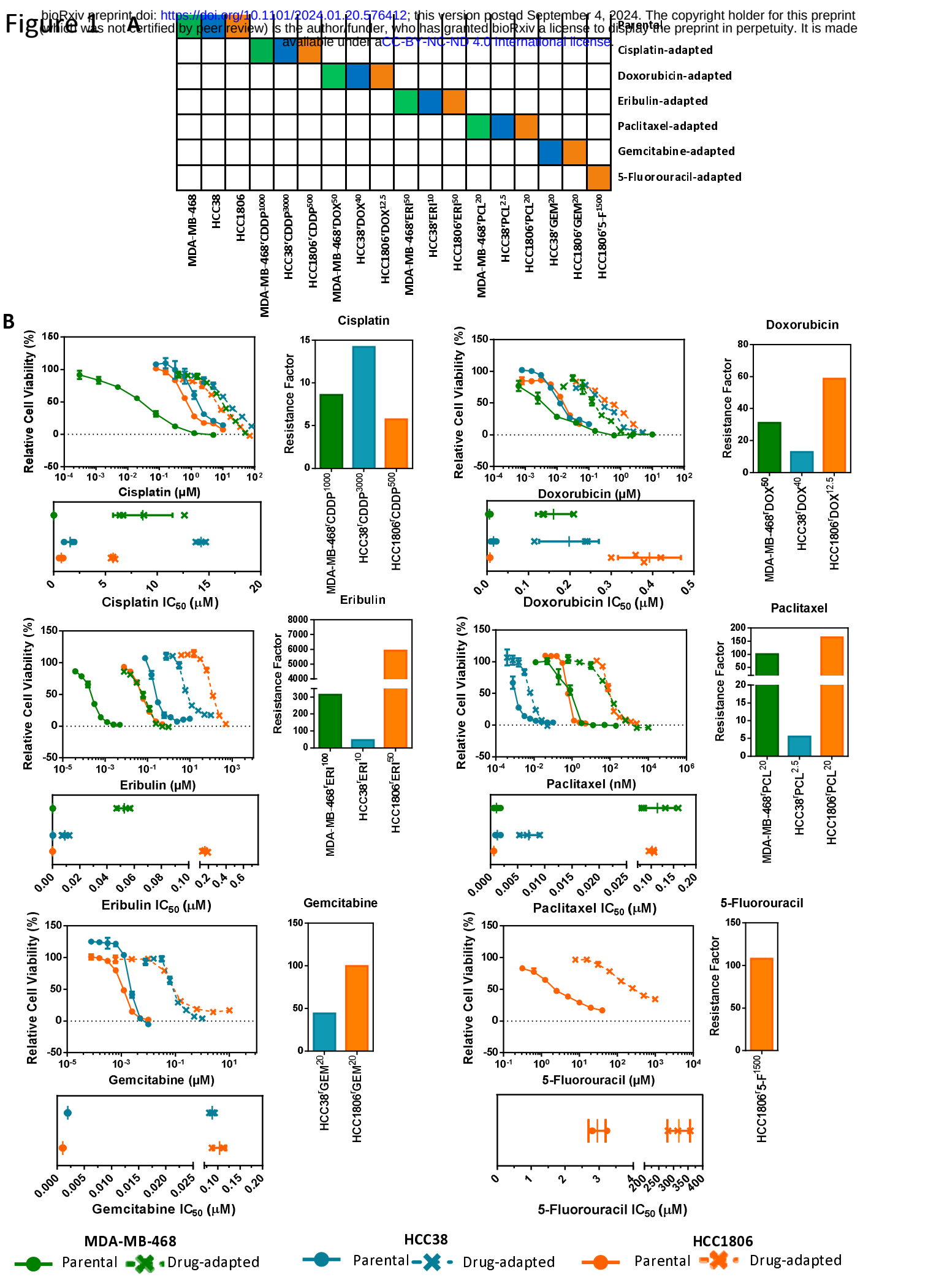

Figure 2

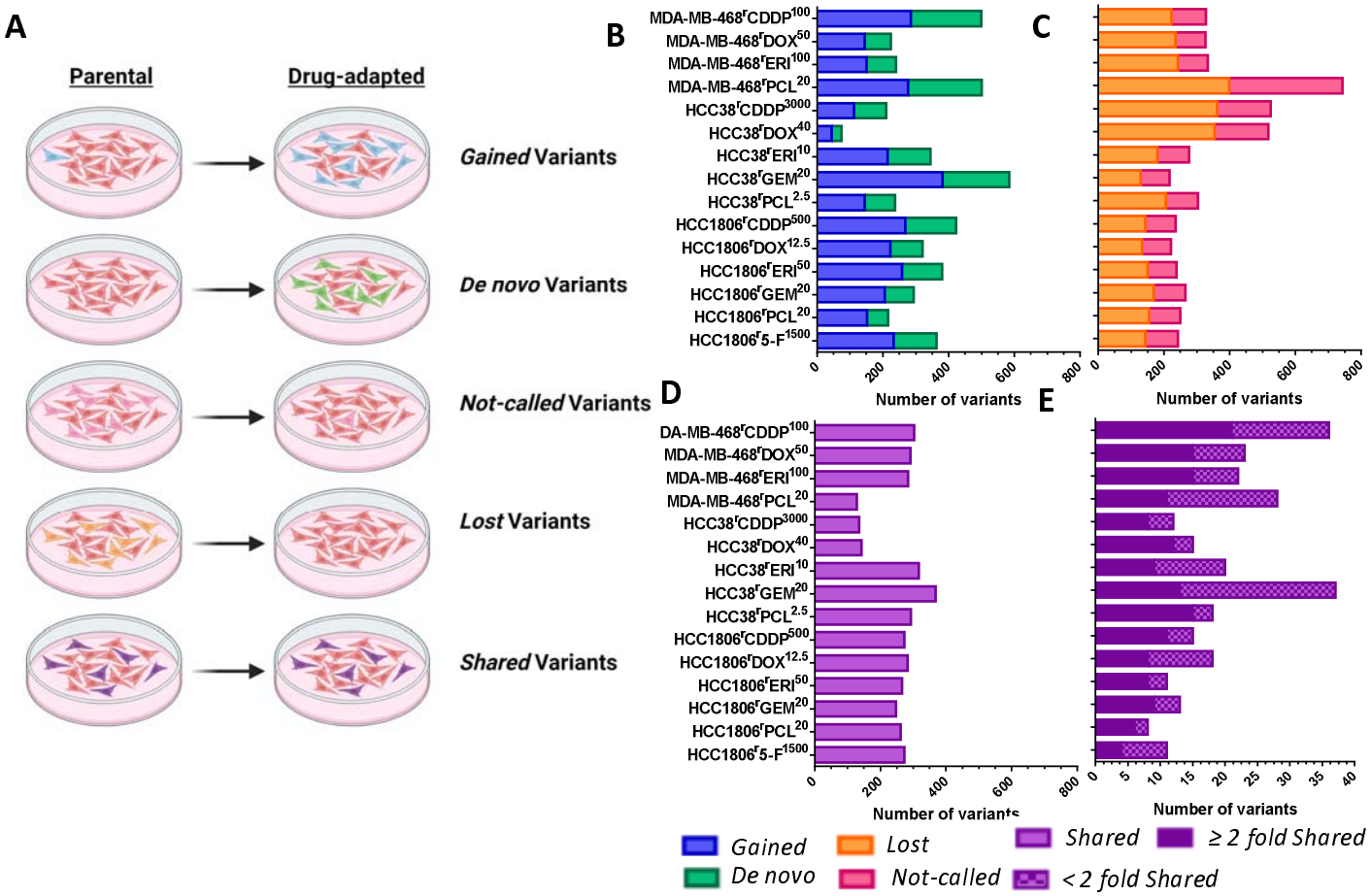

Figure 3

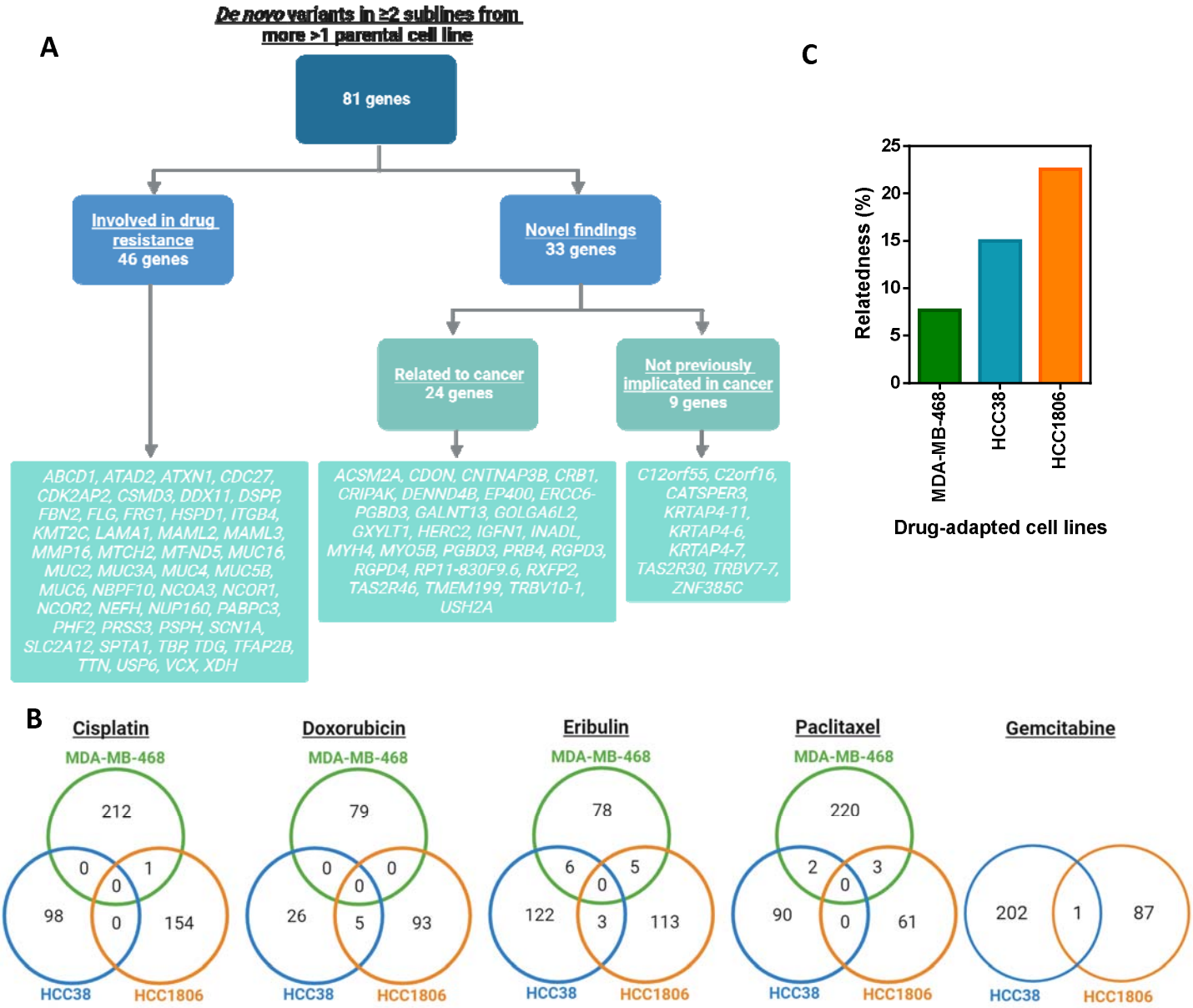

Figure 4

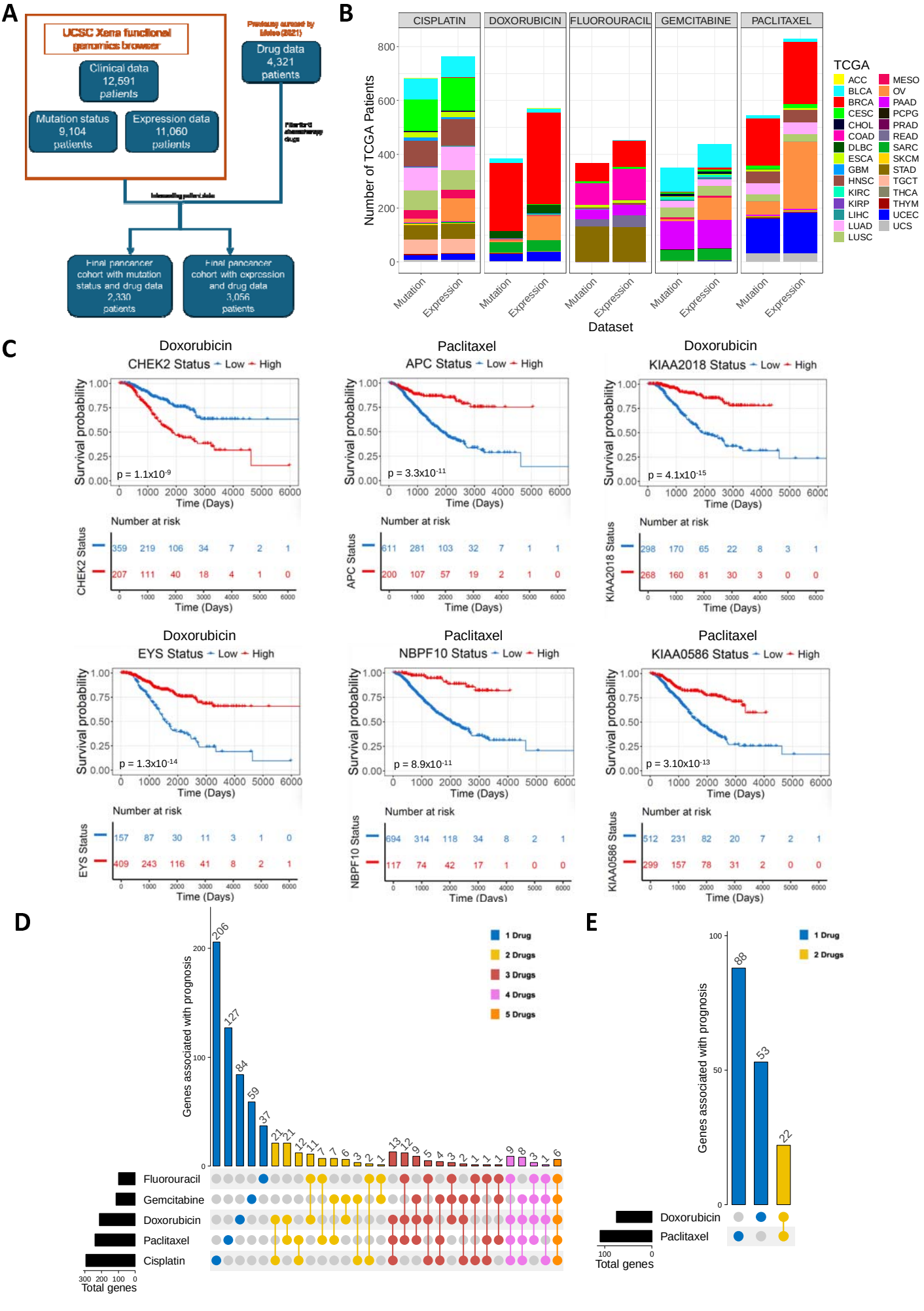

Figure 5