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

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

[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]Genomics Core Facility, Philipps-University, 35043 Marburg, Germany.

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

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

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

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

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

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

[10]Institute for Medical Virology, University Hospital, Goethe University Frankfurt am Main, 60596 Frankfurt, Germany.

26 *Corresponding authors emails: M.Michaelis@kent.ac.uk; M.D.Garrett@kent.ac.uk;

27 M.N.Wass@kent.ac.uk

Abstract

Here, we introduce a novel set of drug-adapted triple-negative breast cancer (TNBC) cell lines consisting of the parental cell lines MDA-MB-468, HCC38, and HCC1806 and their sublines adapted to cisplatin, doxorubicin, eribulin, paclitaxel, gemcitabine, or 5-fluorouracil. Whole exome sequencing in combination with the analysis of TCGA-derived patient data resulted in the identification of 135 biomarker candidates for the guidance of personalized TNBC therapies for further investigation, including 58 novel ones that had not been associated with drug resistance before. The analysis of exome sequencing data showed remarkably few overlaps among the resistant sublines, suggesting that each resistance formation process follows an individual, unpredictable route. This complexity was confirmed by cancer cell line drug sensitivity profiles to cytotoxic anti-cancer drugs and DNA damage repair inhibitors. Drug-adapted sublines of the same parental cell line and sublines adapted to the same drug substantially differed in their drug response patterns. Cross-resistance levels were lowest for the CHK2 inhibitor CCT241533, the PLK1 inhibitor SBE13, and the RAD51 recombinase inhibitor B02, making CHK2, PLK1, and RAD51 promising drug targets for therapy-refractory TNBC. In conclusion, we present novel preclinical models of acquired drug resistance in TNBC and 58 novel candidate biomarkers for further investigation. Whole exome data and drug sensitivity profiles showed that each cancer cell line adaptation process follows an unpredictable route, which 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]. It is responsible for about 15% of breast cancer cases and associated with a poorer prognosis than hormone receptor or HER2 positive breast cancers [1,2]. Current TNBC therapies are largely based on cytotoxic anti-cancer drugs with treatments including cisplatin, doxorubicin, eribulin, gemcitabine, paclitaxel, and 5-fluorouracil [1]. TNBC often responds initially well to cytotoxic chemotherapy, but recurrence and resistance formation are common, eventually leading to therapy failure. This combination of an initial high response rate followed by rapid resistance formation is referred to as the 'TNBC paradox' (Fornier and Fumoleau, 2012; Gupta *et al.*, 2020). To improve TNBC therapy outcomes, new treatment approaches are needed, in particular those that are effective against treatment-refractory disease characterized by acquired resistance to cytotoxic chemotherapy.

The processes underlying the formation of acquired drug resistance in cancer differ from those responsible for intrinsic resistance (Michaelis, Wass and Cinatl, 2019; Oellerich *et al.*, 2019; Santoni-Rugiu *et al.*, 2019; Touat *et al.*, 2020; Rothenburger *et al.*, 2021). In contrast to intrinsic drug resistance, that occurs independently of therapy, and is a consequence of pre-existing often stochastic events in the tumor, acquired resistance is the direct consequence of selection and adaptation processes caused by cancer treatment (directed tumor evolution). These discrepancies in origin can result in differences between the mechanisms underlying intrinsic and acquired drug resistance (Michaelis, Wass and Cinatl, 2019; Oellerich *et al.*, 2019; Santoni-Rugiu *et al.*, 2019; Touat *et al.*, 2020; Rothenburger *et al.*, 2021).

Identifying and understanding these mechanisms is therefore essential to optimize cancer treatment for patients with therapy-refractory tumors.

Cancer cell line adaptation to anti-cancer drugs provides a preclinical platform for the investigation of treatment-induced cancer cell evolution that has been shown in numerous studies to reflect clinically relevant acquired drug resistance mechanisms [5,10–18]. Furthermore, the resulting drug-resistant cell lines allow detailed functional and systems level studies that are not possible using clinical samples (Michaelis, Wass and Cinatl, 2019).

Here, we introduce a novel set of three parental TNBC sublines and their 15 sublines adapted to cisplatin, doxorubicin, eribulin, gemcitabine, paclitaxel, or 5-fluorouracil. The project cell lines were characterized by whole exome sequencing and the determination of response profiles to cytotoxic drugs and DNA damage repair inhibitors. The resulting data showed that the resistance formation processes are individual and unpredictable. The combined analysis of the resistance-associated mutations in combination with patient data from The Cancer Genome Atlas (TCGA) resulted in 58 candidate resistance biomarkers for further investigation [19].

Results

Project cell line panel

The project cell line panel consists 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, drugs from drug classes that are used for the treatment of this cancer type (Fig.1A, SupFile.1) [20–26]. Drug-resistant sublines were established by continuous exposure to stepwise increasing drug concentrations as previously described [17]. 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}) (SupFig.1A) [27]. 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, SupFile.1).

Characterization of the cell line panel by whole exome sequencing

We initially performed whole exome sequencing on the project cell line panel. Between 186 (HCC38^{r}DOX^{40}) and 739 (HCC38^{r}GEM^{20}) DNA sequence variants were detected in the drug-adapted sublines that differed from the respective parental cell lines (SupFig.2A, SupFile.2). Missense variants were most common, followed by synonymous variants (SupFig.2B). Insertions/ deletions (INDELs), frameshift mutations, stop-gain, stop-loss, and splice variants were identified at lower frequencies (SupFig.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 drug-adapted cell lines that increased or decreased by at least two-fold are presented (Fig.2A).

The number of *gained* variants ranged from 44 (HCC38^{r}DOX^{40}) to 381 (HCC38^{r}GEM^{20}), of *de novo* variants from 31 (HCC38^{r}DOX^{40}) to 225 (MDA-MB-468^{r}PCL^{20}), of *not-called* variants from 88 (HCC38^{r}GEM^{20} and HCC1806^{r}DOX$^{12.5}$) to 345 (MDA-MB-468^{r}PCL^{20}), of *lost* variants from 129 (HCC38^{r}GEM^{20}) to 398 (MDA-MB-468^{r}PCL^{20}), and of shared variants from 128 (MDA-MB-468^{r}PCL^{20}) to 368 (HCC38^{r}GEM^{20}) (Fig.2B-2D, SupFile.3).

Analysis of the distribution of *de novo* variants

To identify variants that may have a functional role in drug resistance, we initially looked at genes that harbored *de novo* variants in at least two different sublines from more than one parental cell line, resulting in a list of 81 genes (Fig.3A, SupFile.4). This list includes 48 genes that had already been described to be involved in drug resistance in cancer and 33 new candidate genes with a possible role in drug resistance (Fig.3A, SupFile.4). Notably, 24 of the 33 new candidate genes had already been reported to be of relevance in cancer (Fig.3A, SupFile.4).

Among the five genes with *de novo* variants in the most cell lines were the mucin genes *MUC6* (15 cell lines), *MUC2* (14 cell lines), *MUC4* (13 cell lines), and MUC16 (9 cell lines) (Fig.3A, SupFile.4). These are large genes that are known to be

148 commonly mutated and have been reported to be involved in cancer cell drug
149 resistance [28–32]. *De novo* mutations in CDC27, which has also been linked to drug
150 resistance in cancer, were also found in 9 resistant sublines [33,34] (Fig.3A, SupFile.4).
151 *GXYLT1*, *KRTAP4-11*, and *RGPD4* were the genes among those had not
152 previously been associated with drug resistance that displayed *de novo* mutations in
153 the most (7) resistant sublines (Fig.3A, SupFile.4). GXYLT1 promotes metastasis
154 formation in colorectal cancer through MAPK signaling, a pathway known to provide
155 resistance to a range of anti-cancer drugs [35–38]. RGPD4 was correlated with
156 vascular invasion in HBV-associated hepatocellular carcinoma, and it is known that
157 there is an overlap between pro-angiogenic, pro-metastatic, and resistance-
158 associated signaling in cancer [36,39]. There is no known link between KRTAP4-11 and
159 cancer, but KRTAP4-11 expression levels were reported to predict the methotrexate
160 response in rheumatoid arthritis patients [39]. Hence, it seems plausible that these
161 genes and their products may be involved in cancer cell drug resistance.
162 Taken together, our analysis identified many genes already known to be
163 involved in cancer cell drug resistance alongside a substantial number of novel
164 candidates potentially contributing to therapy failure. Further research will have to
165 characterize the roles of these individual genes in detail.
166 When we compared the overlaps between exactly the same *de novo* variants
167 in sublines adapted to the same drug the numbers were too small to draw any
168 meaningful conclusions (Fig.3B, Sup.Fig3A). Notably, *de novo* variants in drug-
169 resistant sublines may not always represent actual novel variants that are selected
170 because they contribute to cancer cell resistance. Many apparent *de novo* mutations
171 have probably already been present in a small fraction of the cells of the parental cell
172 line, but have not been detected due to limited sequencing depth. Hence, overlaps in

173 *de novo* variants between sublines of the same parental cell line can also be used to
174 indicate the levels of relatedness between the founding subpopulations of the
175 different resistant sublines.

176 Analysing *de novo* variants shared between the sublines of each resistant
177 sublines indicated the largest overlap (22.6% on average) and, thus, relatedness
178 among the HCC1806 sublines, followed by the HCC38 (15.0%), and the MDA-MB-
179 468 (7.7%) sublines (Fig.3C). There were noticeable differences in the overlaps
180 between *de novo* variants of the sublines of the individual parental cell lines. For
181 example, only three *de novo* variants were shared between HCC38^{r}CDDP^{3000} (out of
182 98 in total, 3.1%) and HCC38^{r}PCL$^{2.5}$ (out of 92 in total, 3.3%), while 53 variants were
183 shared between HCC38^{r}ERI^{10} (out of 131 in total, 40.5%) and HCC38^{r}GEM^{20} (out of
184 203 in total, 26.1%) (Fig.3C, SupFig.3B). However, there were no patterns
185 suggesting consistent overlaps between sublines adapted to certain drugs.
186 Therefore, there is no indication that certain drugs may select certain pre-existing
187 cell line subpopulations.

188

189 **Gene ontology (GO) terms related to gene variants that changed in drug-**
190 **resistant sublines**

191 Next, we performed a gene ontology (GO) term analysis of *de novo*, *gained*,
192 *not called*, and *lost* variants as well as *shared* variants with a two-fold increase or
193 decrease in allele frequency (Sup.Fig4A, B).

194 There was a limited overlap between GO terms among sublines adapted to
195 the same drug (SupFig.4C, E). The extracellular matrix-related GO terms
196 'extracellular matrix constituent lubricant activity', 'extracellular matrix', and

197 'maintenance of gastrointestinal epithelium' were most common which reflects the
198 high number of variants observed in the mucin genes (SupFig.4C, E).

199 A GO term analysis among the sublines of the same parental cell lines
200 resulted in very similar results, again revealing an overrepresentation of extracellular
201 matrix-related GO terms (SupFig.4D, F). Further research will have to investigate the
202 potential role of mucins and changes to the extracellular matrix in acquired drug
203 resistance in TNBC cells.

204

205 **Potential clinical relevance of selected variants**

206 The potential clinical relevance of *de novo*, *gained*, and *shared* variants with a
207 two-fold increase was analyzed using patient data derived from The Cancer Genome
208 Atlas (TCGA) data [40]. Initially, we investigated exact variants, i.e., variants that have
209 the same chromosomal position and base change in the TCGA data set. Next, we
210 analyzed variants in the same chromosomal position but with a different base
211 change that resulted in the same consequence i.e., a missense variant. Here, 27
212 Exact and 40 Same Consequence resistance-associated increased variants were
213 identified in TCGA-derived patient mutation data (Fig.4A, Fig.4B).

214 We also focused on all protein truncating variants in cells lines. Variants were
215 selected from the TCGA database if they were of a similar consequence (i.e., a
216 frameshift). Here, 65 resistance-associated variants were found to have protein
217 truncating variants of similar consequence identified in the sublines were also
218 identified in the TCGA-derived patient mutation data (Fig.4, Fig.4C).

219 The fraction of mutated tumors was too low for a meaningful analysis of the
220 potential role of the variants in clinical drug response. Hence, we used gene
221 expression data to evaluate further the role of the respective genes and their

products in response to the respective drugs. The following numbers of Exact and Same Consequence variants could be examined for the respective drugs: cisplatin: 22 (9 Exact, 13 Same consequence), doxorubicin: 9 (3 Exact, 6 Same Consequence), paclitaxel: 12 (5 Exact, 7 Same Consequence), gemcitabine: 16 (8 Exact and 8 Same Consequence), 5-fluorouacil: 8 (2 Exact, 6 Same Consequence) (Fig.4B, Sup.File.5).

For the protein truncating variants, the number of analyzed genes was; cisplatin: 16, doxorubicin: 14, paclitaxel: 7, gemcitabine: 22, 5-fluorouacil: 2. From these genes, Kaplan Meier curves were plotted for high and low gene expression and filtered for FDR and statistical significance (see methods). For the protein truncating variant analysis, we focused on Kaplan Meier curves in which low gene expression was associated with poor patient outcome, as truncations are most likely to result in a loss of function (Fig.4C, SupFile.6).

In total, we identified 62 genes whose expression was associated with the therapy response in cancer patients (Fig.4, Sup.File.5, SupFile.6). This included five genes with a known role in resistance to the drug that the cell line, in which we identified the variants, was adapted to (*COL22A1, FAT4, RGS9, SLC2A12, SLC4A8*) (Fig.4, SupFile.7). Expression levels of 18 further genes have been reported to mediate resistance to other drugs (*ABCB10, ADNP, C20orf27, CHST11, EXT1, FKBP7, NCOR1, PIK3C2B, ABCA8, ACIN1, DNAH5, MTCH2, INHBA, KLF11, SLC22A23, SLC24A1, SMC1B, TYK2*) (SupFile.7). For another 12 genes, there is evidence that they contribute to both (*HUWE1, ITGB4, MSK1, PHF2, TRPM7, BRD7, CES2, IDO1, MSH2, PRLR, RPL14, TOP2A* (Fig.4F) (SupFile.7).

In addition to these 35 genes and their products with a known role in drug resistance, we identified 27 genes that have not previously been linked to drug

resistance (*ANK2, C11orf80, C5orf42, DNAJC13, EPB41, FGF14, FLG, GBGT1, GXYLT1, KCND2, OGN, RNF213, TBC1D9, USH2A, AGAP6, CDON, CEBPZ, CNEP1R1, COG6, CUBN, EFCAB6, HSD17B3, KIAA0586, SETX, SYNGR1, ZKSCAN3, ZNF442*) (Fig.4E).

Eight of the genes and gene products (*CDON*, *FLG*, *GXYLT1*, *ITGB4*, *NCOR1*, *PHF2*, *SLC2A12*, *USHS2A*) had already been identified in our analysis of *de novo* variants (SupFile.4). Six of them have previously been associated with cancer drug resistance (*ITGB4*, *NCOR1*, *PHF2*, *SLC2A12*, *USHS2A*) (Fig.3, Fig.4, SupFile.4, SupFile.7. Further research will have to define in more detail the potential use of the expression levels of and variants in these genes as biomarkers for the direction of clinical therapies.

Complex sensitivity patterns of drug-resistant sublines against cytotoxic drugs

Determining drug sensitivity profiles in the project cell line panel against the drugs of adaptation, i.e., cisplatin, doxorubicin, eribulin, paclitaxel, gemcitabine, and 5-fluorouracil (Fig.5A, SupFile.1), revealed complex resistance patterns that did not follow clear, predictable rules. For example, two out of the three doxorubicin-adapted sublines (HCC38^{r}DOX^{40}, HCC1806^{r}DOX$^{12.5}$) displayed increased (collateral) sensitivity to cisplatin, while MDA-MB-468^{r}DOX^{50} displayed cross-resistance to cisplatin (Fig.5A, SupFile.1). Moreover, all resistant sublines remained sensitive to or showed collateral sensitivity against at least one of the other chemotherapeutic agents (Fig.5A, SupFile.1). The 5-fluorouracil-resistant HCC1806^{r}5-F^{1500} subline was the only resistant cell line that remained sensitive to all other investigated cytotoxic drugs (Fig.5A, SupFile.1).

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 eribulin and paclitaxel-adapted sublines displayed cross resistance to the respective other drug (Fig.5A, SupFile.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 [41–45]. Further research will have to show to which 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 formation to a certain drug will affect the sensitivity patterns of the resulting sublines to other cytotoxic agents. However, all of the drug-resistant triple-negative breast cancer sublines remained sensitive and/ or displayed collateral sensitivity to at least one of the tested chemotherapeutic drugs. Future research will have to elucidate the underlying mechanisms to identify biomarkers for personalized therapy approaches that can guide effective drugs to the right patients [5].

Complex sensitivity patterns of drug-resistant sublines against DNA damage response inhibitors

Triple-negative breast cancer cells have been shown to harbor defects in DNA damage repair signaling, which can result in a dependence on the remaining intact

297 DNA damage repair pathways and, in turn, in sensitivity to certain DNA damage
298 response inhibitors [4]. Hence, we tested a panel of inhibitors targeting critical nodes
299 of DNA damage repair signaling in the project cell lines (Fig.5B).

300 All parental cell lines displayed sensitivity to the tested DNA damage
301 response inhibitors in therapeutic concentrations, i.e., within the Cmax values
302 reported for these agents (if available) (SupFig.5). However and similarly to the
303 results obtained for cytotoxic anti-cancer drugs, the DNA damage response inhibitor
304 sensitivity profiles in the resistant sublines were complex and unpredictable (Fig.5C,
305 SupFile.1). Relative to the respective parental cell lines, the sensitivity remained
306 unchanged for 128 DNA damage response inhibitor/ resistant subline combinations.
307 Increased resistance (cross-resistance) was detected in 96 DNA damage response
308 inhibitor/ resistant subline combinations, and increased sensitivity (collateral
309 vulnerability) was recorded in 16 DNA damage response inhibitor/ resistant subline
310 combinations. Neither sublines of the same parental cell line nor sublines adapted to
311 the same drugs displayed substantial overlaps in their DNA damage response
312 inhibitor sensitivity profiles. Generally, cross-resistance levels were lowest for the
313 CHK2 inhibitor CCT241533, the PLK1 inhibitor SBE13, and the RAD51 recombinase
314 inhibitor B02 among the investigated DNA damage response inhibitors (Fig.5C,
315 SupFile.1).

316 Cross-resistance patterns were even inconsistent between DNA damage
317 repair inhibitors with the same targets. For example, different sensitivity patterns
318 were observed between the ATR inhibitors ceralasertib and berzosertib as well as
319 the CHK1 inhibitors rabusertib, MK-8776, SRA737, and prexasertib (Fig.5C,
320 SupFile.1). The reasons for these discrepancies are unclear. Notably, the activity of
321 the DNA damage repair inhibitors may be modified by interaction with additional

targets, and off-target resistance mechanisms (e.g., processes associated with drug uptake or efflux) may contribute to these differences [46].

In summary and in line with the findings from the investigation of cytotoxic anti-cancer drugs, the drug-adapted triple-negative breast cancer sublines displayed complex, unpredictable sensitivity patterns against DNA damage agents, further demonstrating that improved future therapies will depend on an improved understanding of the underlying molecular processes resulting in the identification of biomarkers that can guide effective therapies to individual patients after treatment failure [5]. Notably, CHK2, PLK1, and RAD51 may have potential as next-line therapies for triple-negative breast cancer patients, whose tumors have stopped responding to chemotherapy.

Investigation of drug sensitivity patterns by the Delta (Δ) method

Finally, we used the delta (Δ) method to identify potential patterns in the response of the project cell lines to all investigated cytotoxic drugs and DNA damage response inhibitors [47]. 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 (Sup. Table1). Positive correlations indicate that increased drug resistance is seen with both agents, whilst negative correlations indicate that whilst increasing drug resistance is observed to one agent, acquired vulnerability is observed in the other agent. In the MDA-MB-468, HCC38, and HCC1806 sublines, we observed 19, 20, and 60 positive correlations and 2, 8, and 1 negative correlations, respectively (Sup. Table1).

We were most interested in the agents that demonstrate negative correlations as they may identify potential next-line treatments. However, among the 11 negative

correlations, there were no consistent results across all three cell lines (Fig.6). This further confirms that resistance mechanisms are complex, individual, and unpredictable and that the identification of potential next-line therapies after treatment failure will depend on an improved understanding enabling therapy monitoring and biomarker-guided treatment adaptation.

Discussion

In this study, we introduce and characterize a novel set of 15 drug-adapted cell lines derived from three parental cell lines that were sensitive to clinical plasma concentrations of the respective drugs. Viability tests confirmed that all drug-adapted sublines had developed substantial resistance to the respective drugs.

Next, we applied whole exome sequencing to identify biomarker candidates for the guidance of anti-cancer therapies. In a 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. This resulted in the identification of 64 genes, whose expression was associated with drug response in cancer patients. This included 37 genes and gene products with a known role in drug resistance and 27 genes that had not been linked to drug resistance before.

Eight of the genes and gene products (*CDON*, *FLG*, *GXYLT1*, *ITGB4*, *NCOR1*, *PHF2*, *SLC2A12*, *USHS2A*) were detected by both approaches, including six ones with an already documented role in cancer drug resistance (*ITGB4*, *NCOR1*, *PHF2*, *SLC2A12*, *USHS2A*). Hence, our study identified in total 135 genes that may represent novel resistance biomarkers, 58 of which had not been associated with drug resistance in cancer before. Further research will have to investigate and define in more detail the role of variants in and the expression of these genes as biomarkers for the tailoring of personalized cancer therapies. Notably, drug-adapted cancer cell lines have already been shown to represent clinically relevant resistance mechanisms in numerous studies [5,10–18].

Interestingly, the analysis of exome sequencing data resulted in remarkably few overlaps between the investigated resistant sublines, including sublines derived from the same parental cell line and sublines adapted to the same drug. This suggests that resistance formation is each time the consequence of a complex, individual, and unpredictable evolutionary process.

This complexity was confirmed by the determination of drug sensitivity profiles, both to cytotoxic anti-cancer drugs and DNA damage repair inhibitors. Drug-adapted sublines of the same parental cell line and sublines adapted to the same drug displayed substantially different drug response patterns. Nevertheless and notably, cross-resistance levels were lowest for the CHK2 inhibitor CCT241533, the PLK1 inhibitor SBE13, and the RAD51 recombinase inhibitor B02 among the investigated DNA damage response inhibitors. Thus, CHK2, PLK1, and RAD51 may be promising drug targets in TNBC patients after failure of the established therapies, in particular if reliable biomarkers are found that identify cancer patients that are likely to benefit from the respective treatments.

Overall, the results from the characterization of the project cell line panel by whole exome sequencing and from the determination of drug sensitivity profiles both indicated that cancer cell resistance formation 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 [17,48,49] and with recent findings from the comprehensive analysis of cancer cell evolution in lung cancer patients [50–54].

In conclusion, we here present a novel set of drug-adapted TNBC cell lines as preclinical models of acquired drug resistance. An initial characterization by whole exome sequencing in combination with patient-derived TCGA data resulted in the identification of 135 biomarker candidates for the guidance of personalized TNBC therapies for further investigation, including 58 ones that are novel and had not been associated with drug resistance in cancer before. Finally, whole exome data and drug sensitivity profiles showed 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 [50–54]. This further supports the relevance of drug-adapted cancer cell lines as preclinical models of acquired resistance that can be analyzed 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 were obtained from ATCC. The drug-adapted sublines (Fig.1A, SupFile.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/)(Michaelis *et al.*, 2011; Michaelis, Wass and Cinatl, 2019). All cell lines were cultured in Iscove's Modified Dulbecco's medium (IMDM) supplemented with 10% fetal bovine serum (Sigma-Aldrich, Germany), 2mM L-glutamine, 25mM HEPES (Fisher Scientific, UK), 100IU/mL penicillin, and 100µg/mL streptomycin (Life Technologies, UK) at 37 °C in a humidified atmosphere at 5 % CO_2. The drug-adapted sublines were continuously cultured in the presence of the respective concentrations of the drugs.

Compounds

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 (Institute of Cancer Research), Eribulin (Eisia), Paclitaxel (Cayman Chemicals). All drug stocks were prepared in DMSO, and stored at -20 °C, except cisplatin which was prepared in 0.9% NaCl solution and stored in the dark at room temperature.

Cell growth and viability assays

Cell viability was tested using the 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) dye reduction assay after 120-hour incubation, modified as previously described [55,56]. Concentrations that reduce cell viability by 50% relative to an untreated control (IC_{50}) were determined and used to calculate the resistance factor (RF; IC_{50} of drug-adapted cell line / IC_{50} of drug-naive cell line).

Whole exome sequencing

Whole exome sequencing (WES) was performed using the Nextera Exome Enrichment Kit (Illumina). 2 x 100 nucleotide paired end sequences were input into Illumina HisSeq2000 with an output of 100 nucleotide paired end reads in FASTQ format. The sequencing was performed in two lanes providing two sets of FASTQ data per cell line.

Variant calling

FASTQC was used to control the quality of the raw sequence data [57], prior to the removal of sequencing adaptors with parameters. Trimmomatic (settings: NexteraPE-PE.fa:2:30:10 LEADING:3 TRAILING:3 SLIDING WINDOW: 4:15 MILEN:36) [58]. Raw FASTQ files were aligned to the human reference genome (GRCH37) using the Burrows-Wheeler Alignment (v.0.7.17) with an output as Sequence Alignment Map (SAM) format applying the default settings -M -R [59–61]. Only paired reads were used and Samtools flagstat used to print statistics throughout each of the subsequent steps. SAM files were input into Picard tools SortSam (v.2.17.10), where the read alignments were sorted by coordinate and converted to a Binary Alignment Map (BAM) format output (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 was used to perform base score recalibration [62]. 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) [63]. Variants were annotated with VEP [64].

Variant filtering

Only variants in the protein sequences were considered. To identify high confidence variants, variants with a Phred score < 30, variants with less than 10 reads supporting the base call, or 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 [65], if not ≥ 3 samples were annotated in The Cancer Genome Atlas (TCGA), or ≥ 10 samples in the Catalogue Of Somatic Mutations In Cancer (COSMIC) [40,66,67].

Definition of variants

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

493

494 **Gene Ontology**

495 Gene ontology (GO) functional enrichment analysis was conducted using
496 G:profiler [68].

497

498 **TCGA analysis**

499 Variant data was extracted via the GDSC Data portal and the Bioconductor R
500 package *TCGAbiolinks* was used to obtain clinical data [69,70]. Chromosomal locations
501 of patient variants were remapped from GRCh38 to GRCh37 using the NCBI
502 Genome Remapping service. Pan-cancer gene expression and survival data was
503 extracted for each chemotherapeutic agent. Survival analyses were conducted to
504 determine the response of the patient treated with the chemotherapeutic agent for
505 when the gene expression was high or low. Cox proportional hazards regression was
506 used to calculate the hazard ratio for cohorts expressing high vs low expression
507 levels of the given gene. The '*surv*_cutpoint' function of the package *survminer* in R
508 allowed for the identification of the optimal expression cut-ff point to give the lowest
509 p-value for high vs low expression. The cut-off selected was between the 20th and
510 80th percentiles of gene expression values as previously described by Uhlen *et al.*,
511 2017. The calculations used overall survival as the measure of clinical outcome.
512 Overall survival is defined as days to last medical follow up or death as was
513 previously described by Ng *et al.*, 2016. The calculations were performed using the
514 R *survminer* and *survival* packages. From this Kaplan-Meier survival curves were
515 generated using the R package *ggsurvplot*. Statistical analysis using the Wald test
516 (or log rank (Mantel-Cox)) test was performed to obtain the p-value of significance
517 for each Kaplan-Meier graph. It should be noted that eribulin variants were omitted

518 from this analysis as at the time of analysis, no eribulin patient data was available for
519 the further analysis steps.
520
521 **Statistical analysis and data manipulation**
522 GraphPad Prism 6 (GraphPad software Inc, USA) was used to generate
523 dose-response curves and determine GI_{50} values using non-linear regression (with
524 variable slope). Statistical significance was calculated using a two-tailed T-test,
525 assuming unequal variance in GraphPad Prism 6 (GraphPad software Inc, USA).
526 Delta method was used as described by Bracht *et al.*, 2006 . IC_{50} values were
527 transformed to Δ IC_{50} values: Δ IC_{50} = *log (average* IC_{50} *in drug over all cell lines) –*
528 *log (individual* IC_{50} *in drug for each cell line)*. Linear regression analysis of ΔIC_{50X}
529 versus ΔIC_{50Y} where X and Y represent two different drugs from the panel, were
530 performed. The Pearson correlation coefficient (*r*) was used to establish the level of
531 significance in a two tailed test with (n-2) degrees of freedom, where $p \leq 0.05$.
532
533

Acknowledgements

This work was supported by grants from the Frankfurter Stiftung für krebskranke Kinder, Kent Health, Kent Cancer Trust, and the Rosetrees Trust. We would like to thank Tim Fenton and lab members for experimental guidance and useful discussions and the Institute of Cancer Research for their kind gift of CCT241533 and SRA737. Figures were created using BioRender.com.

Conflict of Interests

Nothing to declare.

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

5. 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.

6. 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.

7. 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.

8. 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).

574 Mechanisms and therapeutic implications of hypermutation in gliomas. Nature
575 *580*, 517–523. 10.1038/s41586-020-2209-9.

576 9. Rothenburger, T., Thomas, D., Schreiber, Y., Wratil, P.R., Pflantz, T., Knecht,
577 K., Digianantonio, K., Temple, J., Schneider, C., Baldauf, H.M., et al. (2021).
578 Differences between intrinsic and acquired nucleoside analogue resistance in
579 acute myeloid leukaemia cells. J. Exp. Clin. Cancer Res. *40*, 1–19.
580 10.1186/s13046-021-02093-4.

581 10. Juliano, R.L., and Ling, V. (1976). A surface glycoprotein modulating drug
582 permeability in Chinese hamster ovary cell mutants. BBA - Biomembr. *455*,
583 152–162. 10.1016/0005-2736(76)90160-7.

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

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

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

594 14. Nazarian, R., Shi, H., Wang, Q., Kong, X., Koya, R.C., Lee, H., Chen, Z., Lee,
595 M.K., Attar, N., Sazegar, H., et al. (2010). Melanomas acquire resistance to B-
596 RAF(V600E) inhibition by RTK or N-RAS upregulation. Nature *468*, 973–977.
597 10.1038/nature09626.

598 15. Crystal, A.S., Shaw, A.T., Sequist, L. V., Friboulet, L., Niederst, M.J.,
599 Lockerman, E.L., Frias, R.L., Gainor, J.F., Amzallag, A., Greninger, P., et al.
600 (2014). Patient-derived models of acquired resistance can identify effective
601 drug combinations for cancer. Science (80-.). *346*, 1480–1486.
602 10.1126/science.1254721.

603 16. Schneider, C., Oellerich, T., Baldauf, H.M., Schwarz, S.M., Thomas, D., Flick,
604 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.

17. 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.

18. 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.

19. 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.

20. 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.

21. 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.

22. 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.

23. 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.

24. 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.

25. 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.

26. 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.

27. 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.

28. Kim, N., Hong, Y., Kwon, D., and Yoon, S. (2013). Somatic Mutaome Profile in Human Cancer Tissues. Genomics Inform. *11*, 239. 10.5808/gi.2013.11.4.239.

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

30. Pandey, K., Lee, E., Park, N., Hur, J., Cho, Y. Bin, Katuwal, N.B., Kim, S.K.,

667 Lee, S.A., Kim, I., An, H.J., et al. (2021). Deregulated immune pathway
668 associated with palbociclib resistance in preclinical breast cancer models:
669 Integrative genomics and transcriptomics. Genes (Basel). *12*, 1–14.
670 10.3390/genes12020159.

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

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

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

684 34. Feng, Z., Zhang, L., Zhou, J., Zhou, S., Li, L., Guo, X., Feng, G., Ma, Z.,
685 Huang, W., and Huang, F. (2017). mir-218-2 promotes glioblastomas growth,
686 invasion and drug resistance by targeting CDC27. Oncotarget *8*, 6304–6318.
687 10.18632/oncotarget.13850.

688 35. Peng, L., Zhao, M., Liu, T., Chen, J., Gao, P., Chen, L., Xing, P., Wang, Z., Di,
689 J., Xu, Q., et al. (2022). A stop-gain mutation in GXYLT1 promotes metastasis
690 of colorectal cancer via the MAPK pathway. Cell Death Dis. *13*, 1–12.
691 10.1038/s41419-022-04844-3.

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

697 37. 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.

38. 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.

39. 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.

40. 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.

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

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

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

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

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

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

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

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

49. Michaelis, M., Wass, M.N., Reddin, I., Voges, Y., Rothweiler, F., Hehlgans, S., Cinatl, J., Mernberger, M., Nist, A., Stiewe, T., et al. (2020). YM155-adapted cancer cell lines reveal drug- induced heterogeneity and enable the identification of biomarker candidates for the acquired resistance setting. Cancers (Basel). *12*, 1–17. 10.3390/cancers12051080.

50. 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.

51. 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.

52. 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.

53. 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.

54. Abbosh, C., Frankell, A.M., Harrison, T., Kisistok, J., Garnett, A., Johnson, L.,

760 Veeriah, S., Moreau, M., Chesh, A., Chaunzwa, T.L., et al. (2023). Tracking
761 early lung cancer metastatic dissemination in TRACERx using ctDNA. Nature
762 *616*, 553–562. 10.1038/s41586-023-05776-4.

763 55. Mosmann, T. (1983). Rapid colorimetric assay for cellular growth and survival:
764 Application to proliferation and cytotoxicity assays. J. Immunol. Methods *65*,
765 55–63. 10.1016/0022-1759(83)90303-4.

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

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

772 58. Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: A flexible
773 trimmer for Illumina sequence data. Bioinformatics *30*, 2114–2120.
774 10.1093/bioinformatics/btu170.

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

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

780 61. Church, D.M., Schneider, V.A., Graves, T., Auger, K., Cunningham, F., Bouk,
781 N., Chen, H.C., Agarwala, R., McLaren, W.M., Ritchie, G.R.S., et al. (2011).
782 Modernizing reference genome assemblies. PLoS Biol. *9*.
783 10.1371/journal.pbio.1001091.

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

788 63. Li, H. (2011). A statistical framework for SNP calling, mutation discovery,
789 association mapping and population genetical parameter estimation from
790 sequencing data. Bioinformatics *27*, 2987–2993.

10.1093/bioinformatics/btr509.

64. 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.

65. 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.

66. 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.

67. 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.

68. 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.

69. Colaprico, A., Silva, T.C., Olsen, C., Garofano, L., Cava, C., Garolini, D., Sabedot, T.S., Malta, T.M., Pagnotta, S.M., Castiglioni, I., et al. (2016). TCGAbiolinks: An R/Bioconductor package for integrative analysis of TCGA data. Nucleic Acids Res. *44*, e71. 10.1093/nar/gkv1507.

70. Grossman, R.L., Heath, A.P., Ferretti, V., Varmus, H.E., Lowy, D.R., Kibbe, W.A., and Staudt, L.M. (2016). Toward a shared vision for cancer genomic data. N. Engl. J. Med. *375*, 1109–1112. 10.1056/NEJMp1607591.

71. Uhlen, M., Zhang, C., Lee, S., Sjöstedt, E., Fagerberg, L., Bidkhori, G., Benfeitas, R., Arif, M., Liu, Z., Edfors, F., et al. (2017). A pathology atlas of the human cancer transcriptome. Science (80-.). *357*. 10.1126/science.aan2507.

822 72. Ng, S.W.K., Mitchell, A., Kennedy, J.A., Chen, W.C., McLeod, J., Ibrahimova,
823 N., Arruda, A., Popescu, A., Gupta, V., Schimmer, A.D., et al. (2016). A 17-
824 gene stemness score for rapid determination of risk in acute leukaemia. Nature
825 *540*, 433–437. 10.1038/nature20598.

826

827

Figure Legends

Figure 1: Confirmation of the resistance status of the project cell lines. A) Panel of drug-naive (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 (see methods); when drug-naive and drug-adapted cell lines are treated with the respective agent; cisplatin, doxorubicin, eribulin, paclitaxel, gemcitabine, 5-Fluorouracil. Circles indicate drug-naive cell lines, crosses indicate drug-adapted cell lines. Green; MDA-MB-468-derived, blue; HCC38-derived, orange; HCC1806-derived. Data from n ≥ 3, statistics calculated using student t-test and plotted with mean ± SD.

Figure 2: Characterization of drug-adapted cell lines. A) Diagram illustrating the difference 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 *Shared* (purple) variants, right panel; two-fold increase or decrease of shared variants.

Figure 3: Novel candidates with link to therapy failure identified. A) Flow chart of genes that have *de novo* variants observed in ≥2 sublines from >1 parental cell line. B) Venn diagrams of exact *de novo* variants shared between sublines adapted to the same drug. C) Summary of relatedness between sublines adapted from the same parental (%).

Figure 4: Novel gene identified with potential relevance to drug resistance in clinical samples. A) Increased variants and *de novo* and *gained* protein truncating variants as input and screened with known TCGA variants. B) Summary of variants identified and subsequent genes where the Kaplan-Meier graph was statistically significant for high and low gene expression (see methods). C) *ADNP* Kaplan-Meier identified in the doxorubicin, paclitaxel, gemcitabine, and 5-fluorouracil cell lines. D) Left; novel genes associated with drug resistance, right; Kaplan-Meier example of a novel gene KIAA0588. E) Left; genes found to have a role in drug resistance mechanisms, right; Kaplan-Meier example of known gene TOP2A.

Figure 5: Complex sensitivity patterns to cytotoxic and DDR targeted agents. A) Heat map of fold-resistance and collateral sensitivity to cytotoxic agents. B)

Summary of DNA repair pathways targeted by agents used in screening. C) Heat map of fold-resistance and collateral sensitivity to DDR agents.

Figure 6: No trend of sensitivity patterns by Delta (Δ) method. Graphs demonstrating negative correlation; collateral sensitivity in one agent but resistance in to the other (blue), positive correlation; resistance to both agents (red) and no statistical correlation (black) for each group of cells lines belong to MDA-MB-468, HCC38 and HCC1806. For calculation of delta method values see methods.

Supplementary Figure 1: Chemo-naïve cell lines are clinically sensitive to chemotherapy agents. IC_{50} values, of drug-naïve cell lines treated with the respective agent; cisplatin, doxorubicin, eribulin, paclitaxel, gemcitabine, 5-Fluorouracil. Green; MDA-MB-468-derived, blue; HCC38-derived, orange; HCC1806-derived. Black line indicates known Cmax values for chemotherapy agent. Data from n ≥ 3, statistics calculated using student t-test and plotted with mean ± SD.

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

Supplementary Figure 3: *De novo* variant overlaps. The number of *de novo* variants found overlapped in A) drug-resistant cell lines adapted to the same drug, B) drug-resistant cell lines adapted from the same parental cell line.

Supplementary Figure 4: Gene ontology (GO) terms related to gene variants that changed in drug-resistant sublines. A) Number of variants considered increased in drug-resistant sublines (*de novo* variants, *gained* variants and *shared* variants which demonstrated ≥2 increase in variant allele frequency). B) Number of variants considered decreased in drug-resistant sublines (*not called* variants, *lost* variants and *shared* variants which demonstrated ≤2 decrease in variant allele frequency). The number and overlapping GO terms found in increased and

decreased variants were compared between cell lines adapted to the same drug (C, E) and cell lines derived from the same parental cell line (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 (DNA damage response) inhibitors. IC_{50} values, of drug-naïve cell lines treated with the stated drug. Green; MDA-MB-468-derived, blue; HCC38-derived, orange; HCC1806-derived. Black line indicates known Cmax values for DDR agent. Data from n ≥ 3, statistics calculated using student t-test and plotted with mean ± SD

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. Values in table indicate r value of correlation where positive values indicate positive correlation and negative values indicate negative correlation. P values of the correlation are indicated in the blue color scheme, with light blue ($p \leq 0.05$) being the lowest statistical significance, and dark blue ($p \leq 0.00001$) the highest statistical significance.

Supplementary File 1: Mean IC_{50} values, S.D and resistance factor for project panel treated with chemotherapy agents and DNA damage response inhibitors.

Supplementary File 2: Basic variant characterization of the project panel

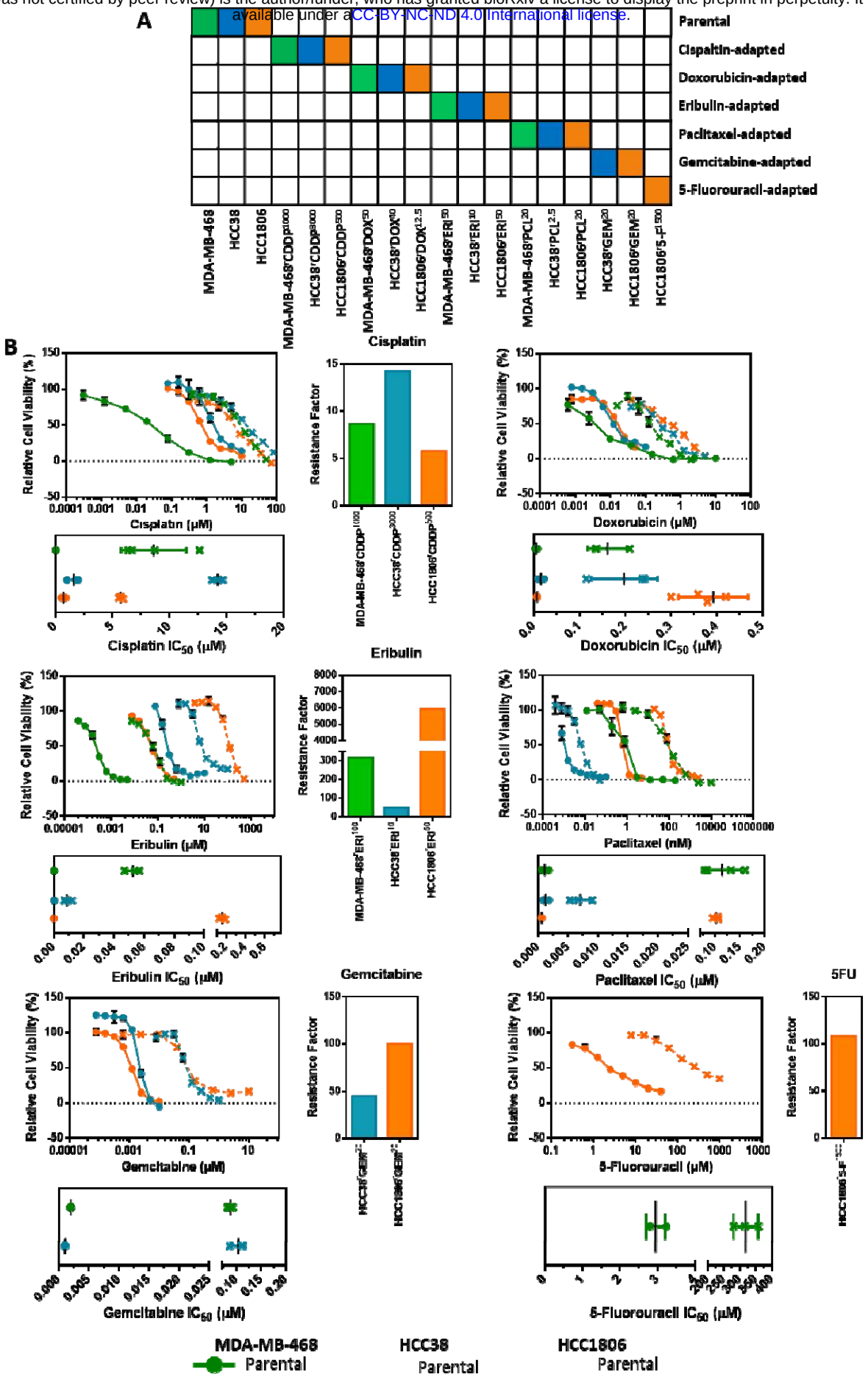
Supplementary File 3: Variants found to be; *de novo*, *gained*, *not called*, *lost* and shared in drug-resistant cell lines

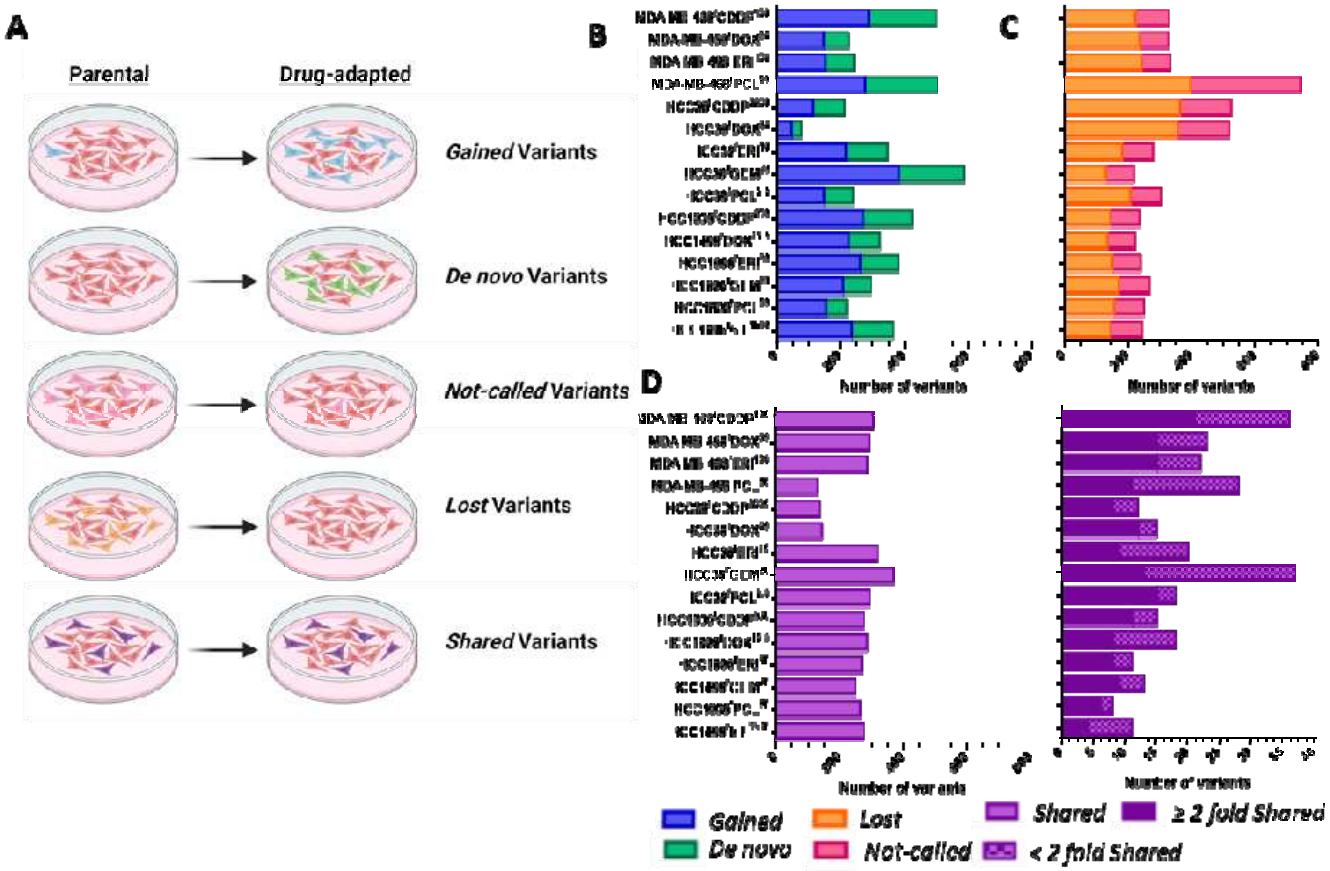
Supplementary File 4: List of genes that have *de novo* variants in ≥2 drug resistant cell lines. Values in the table indicate the variant allele frequency of *de novo* variants identified in stated genes. PMID for genes identified to be previously implicated in cancer and drug resistance.

Supplementary File 5: Exact and same consequence variants Kaplan-Meir graphs when stated gene is expressed high and low and patient is treated with stated drug. Data extracted from the TCGA.

Supplementary File 6: Protein truncating variants Kaplan-Meir graphs when stated gene is expressed high and low and patient is treated with stated drug. Data extracted from the TCGA.

Supplementary File 7: Genes identified in both (i) Exact and same consequence variants and (ii) protein truncating variants analysis. Analysis identifies variant type, cell line and references if gene has been indicated in resistance to the stated drug, or other drug resistance. Green highlighted rows indicate novel resistant candidates.



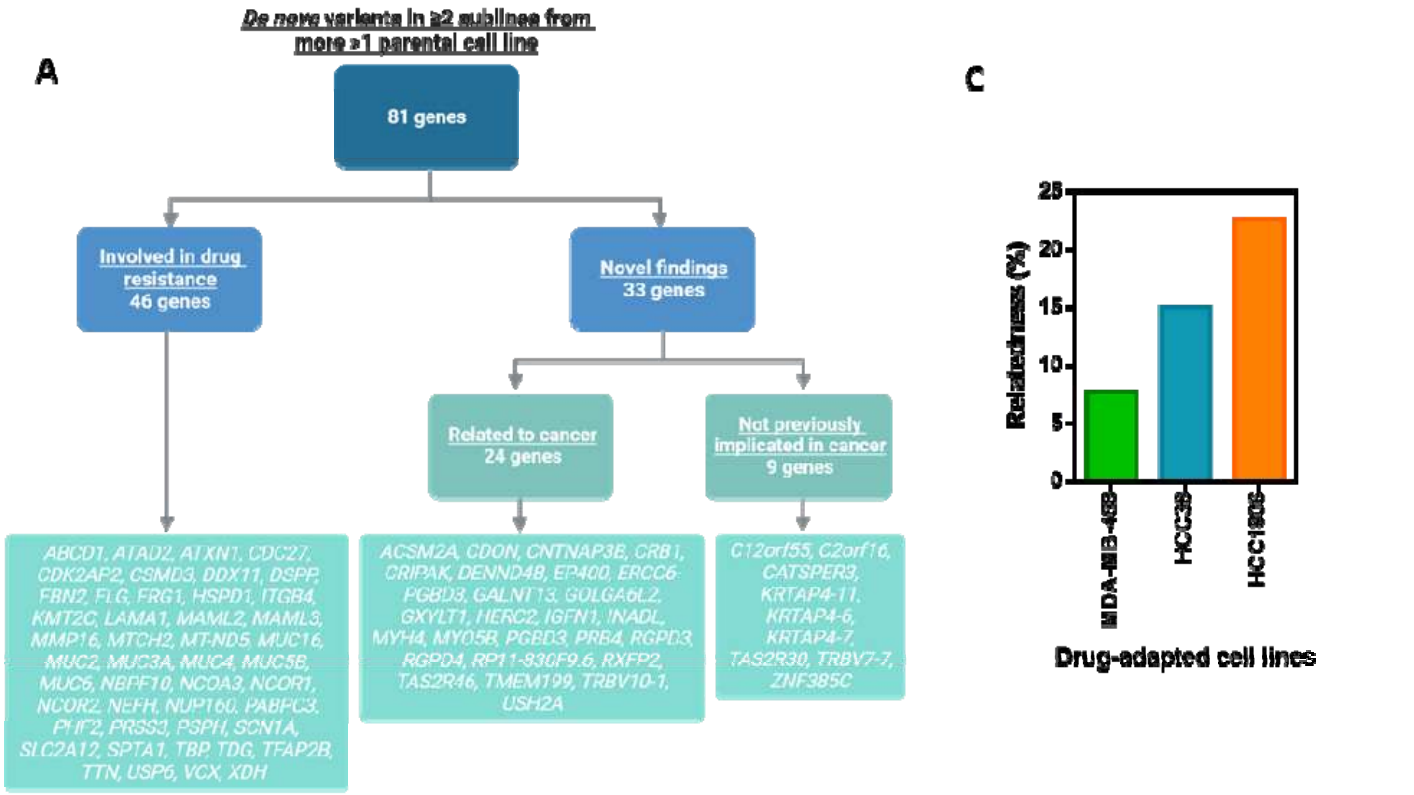


945 **Figure 2**

946

947

948



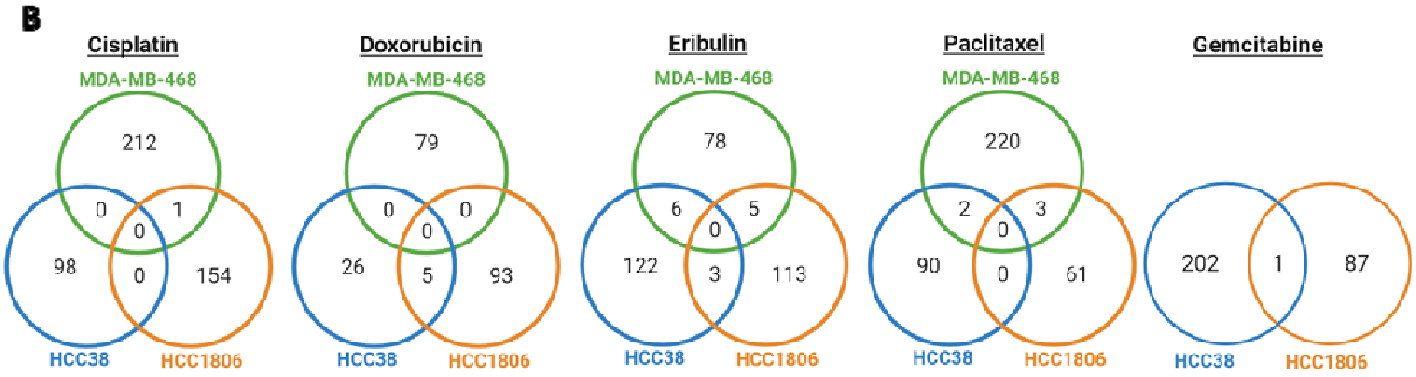
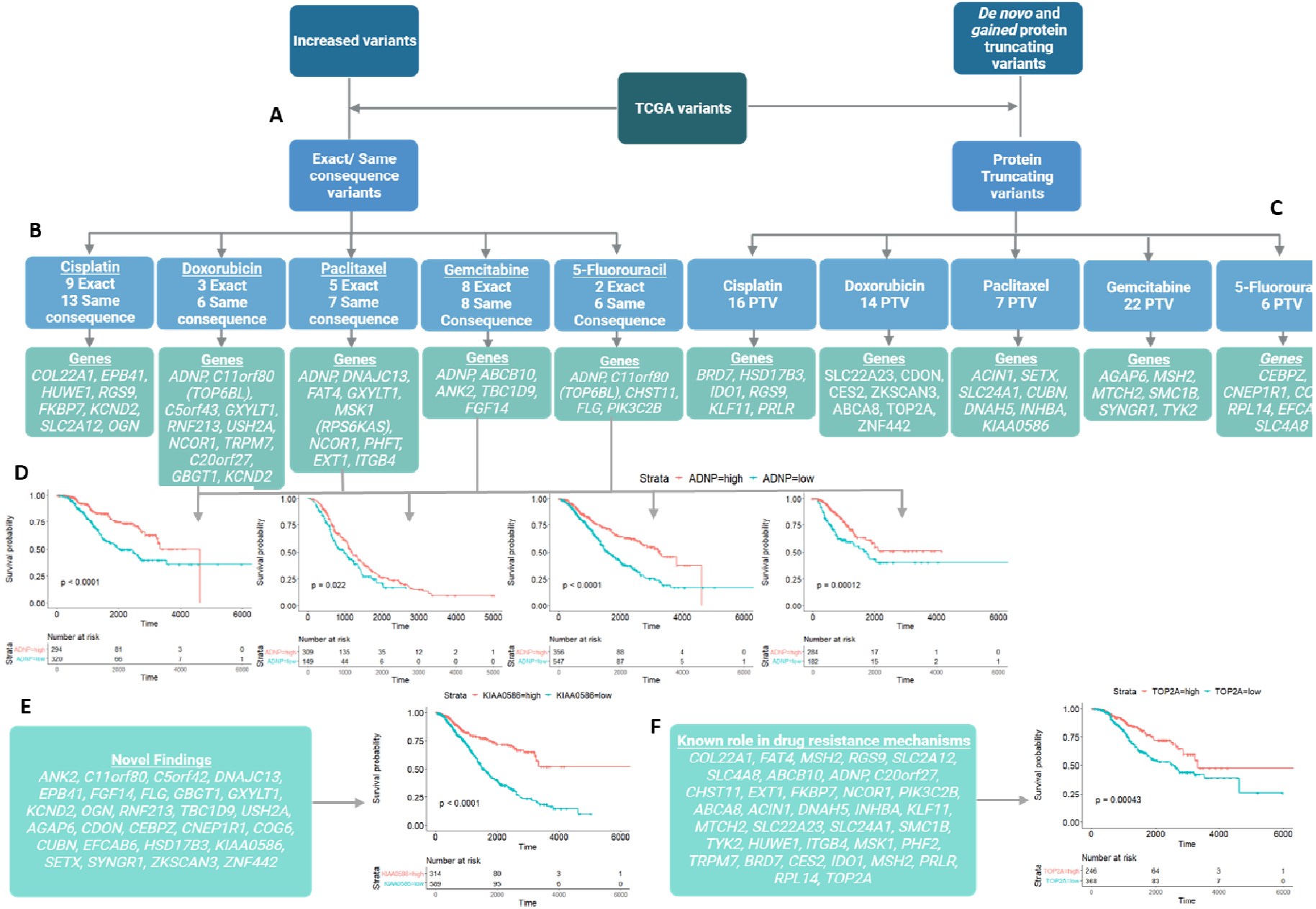


Figure 3



953 **Figure 4**

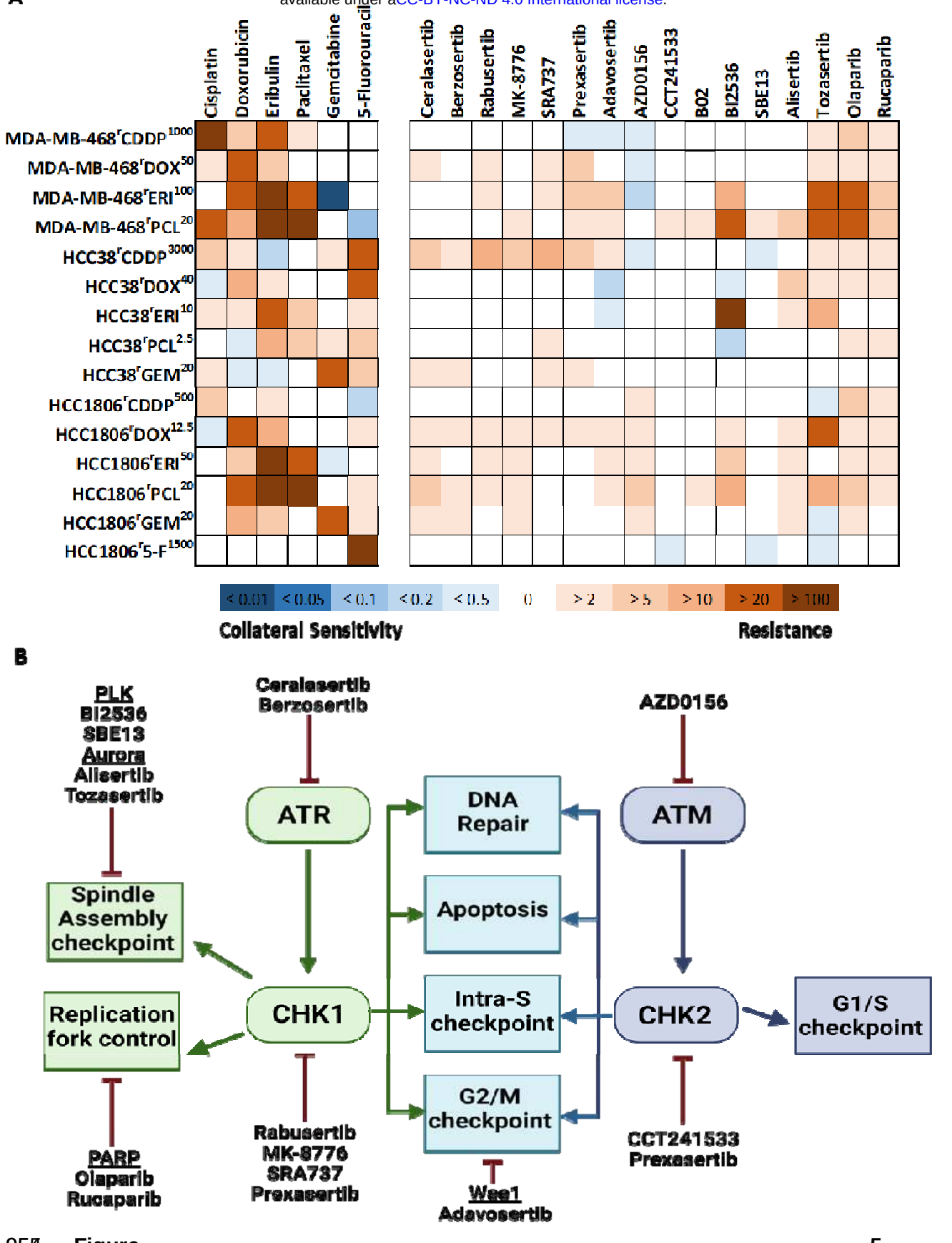


Figure 5

957

958

959

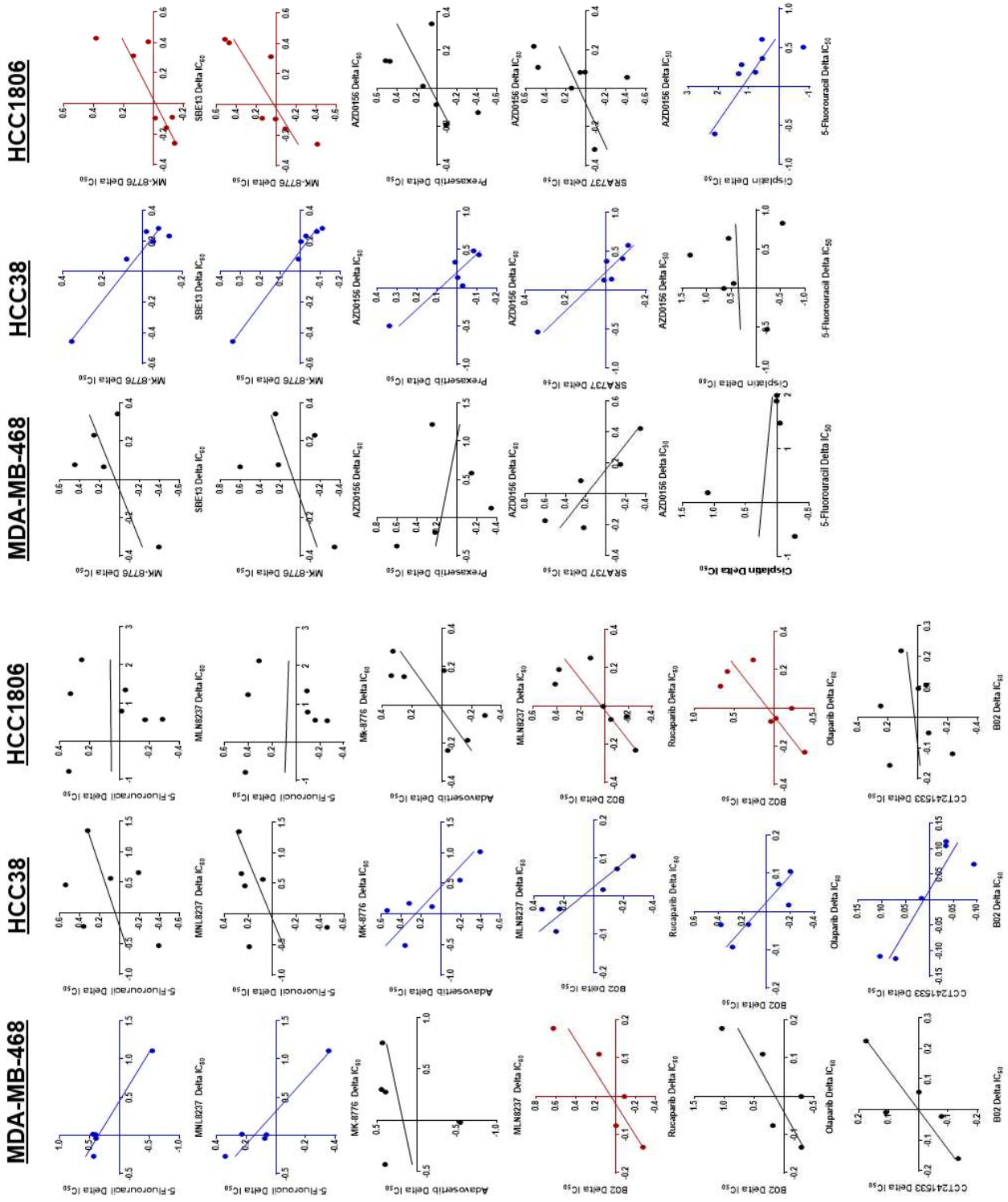
960

961

962

963

964



Figure