Investigating mechanisms of acquired drug resistance in Triple Negative Breast Cancer

Helen Elizabeth Grimsley

November 2020

This thesis is submitted to the University of Kent for the degree of Doctor of Philosophy

Faculty of Sciences

School of Biosciences

University of Kent

Declaration

No part of this thesis has been submitted in support of an application for any degree or other qualification of the University of Kent, or any other University or Institution of learning.

Helen Elizabeth Grimsley

November 2020

Abstract

Triple negative breast cancer (TNBC) is an aggressive, heterogenous, metastatic disease characterised by lack of oestrogen, progesterone and human epidermal growth factor 2 receptors. Due to the lack of druggable targets, chemotherapy remains the mainstay of treatment. Whilst patients initially respond to this therapy, they often relapse due to acquired drug resistance. Given the poor outlook, it is evident that an appropriate second line therapy is required following chemoresistance, as well as clinically relevant biomarkers to identify when this change of therapy is required.

To address this, a panel of six chemotherapeutic agents and 16 inhibitors of the DNA damage response and repair (DDR) pathways were evaluated as potential second line therapy options against a panel of three chemo-naive and 15 chemo-resistant TNBC cell lines. This showed that the PARP inhibitors, olaparib and rucaparib, and the chemotherapeutic agent, doxorubicin, may be ineffective as second treatment strategies for chemo-refractory TNBC, whilst inhibitors targeting CHK2, RAD51 recombinase and PLK1 may be effective. Analysis of the TNBC cell lines exome sequencing data, in combination with data extracted from The Cancer Genome Atlas, identified 70 genes as candidate biomarkers of chemoresistance. This included a loss of function frameshift variant in *TOP2A* within the doxorubicin resistant cell line, HCC1806^{r}DOX$^{12.5}$, as a candidate biomarker of doxorubicin resistance. SiRNA mediated knockdown of TOP2A in the chemo-naive cell line, HCC1806, confirmed a doxorubicin resistance phenotype.

In conclusion, this thesis provides a novel insight into the use of chemotherapeutic agents and DDR inhibitors as a potential second line therapy options after the emergence of chemoresistance in TNBC. It identified 70 clinically relevant candidate biomarkers of chemoresistance, and provides new avenues of research for further exploration of these findings. The work presented in this thesis has the potential to advance understanding of chemoresistance in the clinic and improve the outcome of patients with chemo-refractory TNBC.

Acknowledgements

I would first and foremost like to thank my supervisors Prof. Michelle Garrett (“thanks a bunch”, haha), Prof. Martin Michaelis and Dr Mark Wass for their guidance, patience and invaluable support. Their advice and encouragement have been key to my development as a scientist. My appreciation goes out to Dr Catherine Harper-Wynne and Ms Karina Cox from Maidstone and Tunbridge Well NHS Trust for their insight and advice. I wish to acknowledge Kent Health, Kent Cancer Trust and the Rosetrees Trust for the funding and resources for this project.

Special thanks go to lab members past and present who have helped me in many different ways. Jasmine for your experimental guidance and daffodils. Nathan for your experimental critic and for being my early morning hood buddy. Daniel for your optimism, and distracting conversations. Katie for your kindness and help in the analysis of the TCGA data. Stuart for your pessimism, terrible puns and joy you brought to the lab (RIP SUSAN). Henry, for puns, podcasts and support. You are still wrong about the chicken cake. Ian, I saw you sometimes. Jake for his coding expertise, teaching me how to play squash, the terrible Scottish accent and the rib challenge. Eithaar for “arts and crafts”, getting me into trouble, and the support when I most needed it. Magda the lionheart – I would never have got through this PhD without you. Thank you for your support both academically and personally. Jo for the help at the start, the middle and end, telling me to believe in myself, the puns, jokes, “would-you-rathers”, adventures and arguments about mash potato. Special thanks must go to Edith Blackburn, our lab technician, and quite frankly the lab hero. Edith’s support has been indispensable in my project and the running of the lab, and her support goes far beyond that of a colleague. I also extend my thanks to previous MSc students who were a joy to work with, and others in the School of Biosciences that have helped me along the way.

Thanks go to the UKC Fencing Club, especially to the alumni, Pat and Carols, for the trips to sweaty sports halls across the country. Thanks to Joe, Sam, Rozanne (you bad devil), Nick, Jasmine, Shane and the Cherry Tree. You’ve provided light relief, late nights, and kept music alive. Many thanks to Matt for the emotional support, and for being my partner in crime for many years. Huge appreciation to Adele and Robert for your love and support throughout. I would not have survived this without my “witches”; Becky, Em and Louise. Thanks for the witch’s weekends, the wine, beer, tequila, the emergency phone calls and the endless “cackles”.

Finally, this PhD thesis is dedicated to my Mum, Dad and my sister Grace, who have supported me on this journey from the very start. None of this would have been possible without their continued love and support. Thank you for believing in me, for that I am eternally grateful.

Table of Contents

List of Figures

List of Tables

Abbreviations

AACR	American Association for Cancer Research
ABC	ATP-binding cassette
ADH7	all-trans retinol dehydrogenase
ADP	adenosine diphosphate
AKT	v-akt murine thymoma viral oncogene homolog
ANP32B	acidic leucine-rich nuclear phosphoprotein 32 family member B
APAC/C	anaphase promoting complex/cyclosome
ARG2	arginase type II
ATM	ataxia telangiectasia mutated
ATP	adenosine triphosphate
ATR	Ataxia-Telangiectasia and Rad3 related
ATRIP	ATR-interacting protein
AURK	aurora kinases
BAM	Binary Alignment Map
BCA	Bicinchoninic acid
BCF	Binary Variant Call Format
BCLAF1	Bcl-2-associated transcription factor 1
BCRP	breast cancer resistance protein
BER	base excision repair
BIC	Bayesian Information Criterion
BLM	DNA2-Bloom syndrome protein
BRCA1	breast cancer type 1 susceptibility protein
BRCA2	breast cancer type 2 susceptibility protein
BRD4	bromodomain 4
BSA	bovine serum albumin
C2CD3	calcium-dependent domain containing 3
CA-125	carbohydrate antigen -125
CCLE	Cancer Cell Line Encyclopaedia
CDC25	cell division cycle 25
CDK	cyclin dependent kinase
CDKN1A	cyclin-dependent kinase inhibitor 1
cDNA	complementary DNA
CEP152	centrosomal protein 152 kDa
CHK1	checkpoint kinase 1
CHK2	checkpoint kinase 2
CHO	Chinese Hamster Ovary
CLOCK	circadian locomotor output cycles kaput
CRISPR-cas9	clustered regularly interspaced short palindromic repeats – associated protein 9
cNHEJ	Classical NHEJ
COSMIC	Catalogue of Somatic Mutations In Cancer
CRISPR	clustered regularly interspaced short palindromic repeats
CSCs	cancer stem cells

CSNK1D	casein kinase 1, delta
CTCs	circulating tumour cells
ctDNA	circulating tumour DNA
CTRP	Cancer Therapeutics Response Portal
DAPKs	death-associated protein kinases
DBS	Doublet base substitution
DD	death domain
DDR	DNA damage response and repair
DDR2	*discoidin domain containing receptor-2*
DED	death effector domain
dFdCDP	gemcitabine diphosphate
dFdCTP	gemcitabine triphosphate
DHT	5α-dihydrotestosterone
DIDS	4,4'-diisothiocyanatostilbene-2,2'-disulfonic acid
DISC	death inducing signalling complex
D-loop	displacement loop
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
DNA-PK	DNA-dependent protein kinase
DPD	dihydropyrimidine dehydrogenase
DSBs	DNA double strand breaks
dsDNA	double stranded DNA
dTMP	deoxythymidine monophosphate
dUMP	deoxuridine monophosphate
dUTP	deoxyuridine triphosphate
ECL	enhanced chemiluminescence
EGFR	epidermal growth factor receptor
ELISA	enzyme-linked immunosorbent assay
EMT	epithelial-to-mesenchymal transition
ER	oestrogen receptor
ERCC1	excision repair cross-complementation group 1
ERCC6	DNA excision repair protein ERCC-6
ETAA1	Ewing's tumour-associated antigen 1
FADD	Fas-associated death domain
FANC	Fanconi anaemia complementation group
FBS	fetal bovine serum
FDA	Food and Drug Administration
FDR	false discovery rate
FdUMP	fluorodeoxyuridine monophosphate
FdUTP	flurodeoxyuridine triphosphate
FdUTP	flurodeoxyuridine triphosphate
FER	tyrosine-protein kinase FER
FGF	fibroblast growth factor
FGFR	fibroblast growth factor receptor
FISH	fluorescence *in situ* hybridization
FUTP	fluorouridine triphosphate
FZR1	fizzy and cell division cycle 20 related 1

GalNAc	cetylgalactosamine
GAPDH	glyceraldehyde 3-phosphate dehydrogenase
GATK	GenomeAnalysisTK
GCH1	GTP cyclohydrolase 1
GDSC	The Genomics of Drug Sensitivity in Cancer
GI_{50}	half-maximal growth inhibitory concentration
GLOBOCAN	Global Cancer Observatory
gnomAD	genome aggregation database
GO	gene ontology
GOF	gain of function
GRCH37	Genome Reference Consortium human reference genome
GSH	glutathione
GSTπ	glutathione S-transferase
HER2	human epidermal growth factor receptor-2
HER2+	human epidermal growth factor receptor 2 positive
HERC2	E3 ubiquitin-protein ligase HERC2
HNSCC	head and neck squamous cell cancer
HR+	Hormone receptor positive
HRR	homologous recombination repair
Hsp90	heat shock protein 90
HTT	huntingtin
ICGC	International Cancer Genome Consortium
ICIs	immune checkpoint inhibitors
ICL	interstrand cross-links
ID	Small Insertion and Deletion
IDO1	Indoleamine 2,3-dioxygenase 1
IGF1R	insulin-like growth factor 1 receptor
IHC	immunohistochemistry
IMDM	iscove's modified dublbecco's medium
INDELs	inserts/deletions
INHBA	inhibin beta A chain
INSIGhT	individualised screening trial of innovative glioblastoma therapy
JAK1	interferon-receptor-associated Janus kinase 1
JAK2	Janus kinase 2
kDA	kilo Dalton
KIF18B	kinesin-like protein KIF18B
KIF2C	kinesin family member 2C
KLF11	krueppel-like factor 11
LAR	luminal-androgen
LIN54	protein lin-54 homolog
LOF	loss of function
MDR1	multidrug-resistance protein 1
MET	mesenchymal-to-epithelial transition
MMR	mismatch repair
MRN	MRE11-RAD50-NBS1
mRNA	messenger ribonucleic acid

MSL	mesenchymal-stem like
MTT	3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide
MudPIT	multidimensional protein identification technology
NAC	neoadjuvant chemotherapy
NCI	National Cancer Institute
NER	nucleotide excision repair
NF1	neurofibromin
NF-κB	nuclear factor kappa-light-chain-enhancer of activated B cells
NHEJ	non-homologous end joining
NICE	National Institute for Health and Care Excellence
NIN	ninein
NO	nitric oxide
NSCLC	non-small-cell lung cancer
NT	non-targeting
NT5E	5-nucleotidase, alternative name CD73
nuclear factor kappa-light-chain-enhancer of activated B cells	NF-κB
OCTs	organic cation transporters
OPRT	orotate phosporibosyl transferase
p21	$p21^{CIP1/WAF1}$
PARP	poly(ADP-ribose) polymerase
PARP1	poly[ADP-ribose] polymerase 1
PBS	phosphate buffered saline
PCAWG	pan-cancer analysis of whole genomes
pCR	pathological complete response
PCR	polymerase chain reaction
PCWAG	Pan-Cancer Analysis of Whole Genomes
PD-1	programmed cell death protein 1
PD-L1	programmed death-ligand 1
Penstrep	penicillin streptomycin
PI3K	phosphoinositide 3-kinases
PIKKs	PI3K related kinases
PKA	protein kinase A
PLK1	polo-like kinase 1
POLB	DNA polymerase β
PolyPhen	polymorphism phenotyping
PR	progesterone receptor
PTEN	Phosphatase and tensin homolog
PTGS2	prostaglandin G/H synthase 2
PTPN3	Tyrosine-protein phosphatase non-receptor type 3
PTV	protein truncating variants
PVDF	polyvinylidene fluoride
RAD51	DNA repair protein RAD51 homolog 1
RAE1	mRNA export factor
Rb	retinoblastoma protein
RB1	Retinoblastoma
RCCL	Resistant Cancer Cell Line collection

RF	resistance factor
RNA	ribonucleic acid
RNA-seq	RNA sequencing
ROS	reactive oxygen species
RPA	replication protein A
rRNA	ribosomal ribonucleic acid
RT-Q-PCR	reverse transcriptase quantitative polymerase chain reaction
SBS	Single Base Substitution
SCLC	small cell lung cancer
SDHA	succinate dehydrogenase [ubiquinone] flavoprotein subunit, mitochondrial
SDS	sodium dodecyl-sulfate
SDS-PAGE	sodium dodecyl-sulfate - polyacrylamide gel electrophoresis
shRNA	short hairpin RNA
SIFT	sorting intolerant from tolerent
siRNA	small interfering ribonucleic acid
SLC22A16	solute carrier family 22 member 16
SLCs	solute carrier superfamily
SMARCAL1	SWI/SNF-related matrix-associated actin-dependent regulator of chromatin subfamily A-like protein 1
SMT	somatic mutation theory
SNVs	Single nucleotide variants
SPECC1L	cytospin-A
SRB	sulforhodamine B
SSBs	single strand breaks
ssDNA	single-stranded DNA
TBST	Tris Buffered Saline Tween Buffer
TCGA	The Cancer Genome Atlas
TDRD9	ATP-dependent RNA helicase TDRD
TEMED	tetramethylethylenediamine
TILs	tumour-infiltrating lymphocytes
TLR5	toll-like receptor 5
TNBC	triple negative breast cancer
TOFT	tissue organisation field theory
TOP2A	topoisomerase 2 alpha
TOPBP1	DNA topoisomerase 2-binding protein 1
topo I	topoisomerase I
topo II	topoisomerase II
TP	thymidine phosphorylase
TRADD	Tumour necrosis factor receptor type 1-associated DEATH domain
tRNA	transfer ribonucleic acid
TSGs	tumour suppressor genes
UP	uridine phosphorylase
v/v	volume/volume
VAF	variant allele frequency
VCF	Variant Calling Format

VEP	variant effect predictor
w/v	weight/volume
WDR76	WD repeat-containing protein 76
WEE1	wee1-like protein kinase
WES	whole exome sequencing
WGS	whole genome sequencing
XRN2	5’-3’ exoribonuclease 2
ZBTB17	zinc finger and BTB domain -containing protein 17
ZRANB3	DNA annealing helicase and endonuclease ZRANB3

Chapter 1

Introduction

1. Introduction

1.1 Introduction to Cancer

Cancer is a major worldwide public health problem, and is the second leading cause of premature death globally after cardiovascular disease (Ma and Yu, 2006; Bray *et al.*, 2018). It was estimated that there were 18.1 million new cancer cases and 9.6 million cancer deaths in 2018 with high prevalence of lung, female breast, colorectal and prostate cancer (Bray *et al.*, 2018). According to Cancer Research UK; it has been predicted that there will be 27.6 million new cases of cancer each year by 2040 (Cancer Research UK, 2016). It is therefore clear that a deeper understanding of cancer is required for better cancer prevention, detection, diagnosis and treatment worldwide.

Cancer cannot be described as a single disease state, but comprises of over 200 different diseases, which commonly demonstrate a phenotype of uncontrolled cell growth and proliferation, which can result in invasion of surrounding local tissue, and metastasis to secondary sites. There are several theories, which propose an explanation for the progression of cancer. The tissue organisation field theory (TOFT) states that cancer arises from the disruption of chemical signals, mechanical forces and bioelectric changes of a cell with the adjacent tissue (Baker, 2015). The metabolic theory considers cancer to arise as a mitochondrial metabolic disease through respiratory insufficiency, whilst the cancer stem cell theory considers cancer to arise as result of genomic instability in stem or differentiated cells, termed cancer stem cells (CSCs) (Seyfried, 2015; Tomasetti and Vogelstein, 2015; Tran *et al.*, 2016; Afify and Seno, 2019). The most widely accepted theory, and the one that is assumed through this thesis, is the somatic mutation theory (SMT). SMT states that the development of genetic and epigenetic alterations in a single somatic cell can extricate the cells from the homeostatic mechanisms, which control normal cell proliferation, resulting in the cancer cell phenotype (Blagosklonny, 2005).

SMT considers the development of cancer as a Darwinian process, whereby the accumulation of mutations in the genome over successful clonal expansions result in the mutant oncogenic phenotype (Hanahan and Weinberg, 2011). The majority of somatic mutations that occur are considered passenger mutations, which have no effect on the cell, whilst driver mutations result in an oncogenic progression (Stratton, 2009). Driver mutations are commonly found in oncogenes or tumour suppressor genes (TSGs), which can positively or negatively regulate the promotion of cancer progression respectively. Proto-oncogenes often obtain activating mutations, such as amplification, translocation or missense variants, which are often dominant (Grandér and Grandér, 1998). An example of an oncogene is *EGFR*, whereby activating mutations in EGFR results in

increased cell proliferation, survival and differentiation (Zandi *et al.*, 2007; Sigismund *et al.,* 2018). TSGs are known to have a role in protecting the genetic integrity of the cell, and deleterious loss of function variants in these genes can lead to oncogenesis (Grandér and Grandér, 1998). An example of a commonly mutated TSG in cancer is *TP53*, which has a crucial role in sensing DNA damage and halting cell cycle progression. Loss of function mutations in *TP53* are found in approximately 50% of cancers (Rivlin *et al.*, 2011; Perri, 2016). Genetic aberrations in oncogenes and TSGs affect the genomic stability of the cell, which can result in a higher risk of acquiring further genetic aberrations (Fletcher and Houlston, 2010).

Both internal and external genotoxic factors such as; reactive oxygen species (ROS), tobacco smoke and ultraviolet light, can result in DNA damage (Jackson and Bartek, 2009). Although DNA repair mechanisms are in place to repair the generated somatic mutations, if repaired incorrectly before cell replication; a permanent alteration is fixed in the DNA, further contributing to oncogenesis (Lord and Ashworth, 2012). It has been found that cancer cells often have increased rate of DNA damage and genetic aberrations compared to normal cells, which contributes to the development of oncogenesis (Negrini, 2010). It is thought that inherited germline mutations in oncogenes, such as those in *KRAS* in non-small cell lung cancer patients, can contribute to the development of oncogenesis in combination with somatic mutations (Román *et al.*, 2018).

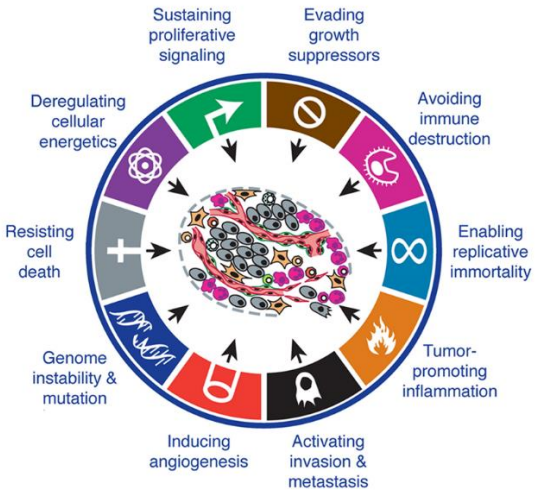


Figure 1.1 The hallmarks of cancer
Schematic diagram highlighting the hallmarks of cancer. Taken from Hanahan and Weinberg 2011.

The sequential development of driver somatic mutations in both oncogenes and TSGs results in an oncogenic phenotype that have been defined as traits, and considered to drive the cancer progression; termed the hallmarks of cancer (Hanahan and Weinberg, 2011). Initially, six hallmarks were identified; with the cancerous cells able to sustain proliferative signalling, evade growth

suppressors, activating invasion and metastasis, enabling replicative immortality, inducing angiogenesis and resisting cell death (Hanahan and Weinberg, 2000). A reviewed update identified two new emerging hallmarks; deregulating cellular energetics and avoiding immune destruction, as well as two enabling characteristics; tumour-promoting inflammation and genome instability and mutation (Figure 1.1) (Hanahan and Weinberg, 2011). Different therapeutic strategies have been developed which target these characteristics of the cancer phenotype. One of the earliest methods to be developed, and which is still a frontline treatment for many cancer types, is chemotherapy.

1.2 Chemotherapeutics

Cancer therapy is predominantly dependent upon the use of surgery, radiotherapy and cytotoxic chemotherapeutic agents. Whilst both surgery and radiotherapy are still commonly used, these methods are reliant on accessibility of the tumour site, and the stage of the tumour. Cytotoxic chemotherapy agents were introduced, which can provide an effective broad-acting therapy across all tissue types and can be used at any tumour stage. These cytotoxic agents are used to induce severe DNA damage, with the rationale that this will result in higher levels of replication stress or mitotic catastrophe and activation of apoptosis in the rapidly dividing cancerous cells compared to normal cells. There are several classes of chemotherapeutic agents based on their mechanism of action, which include, (but not exclusive to); alkylating agents, anti-metabolites, topoisomerase inhibitors and anti-mitotic agents, which are considered in the scope of this thesis due to their use in the treatment of the cancer investigated here (Huang *et al.*, 2017).

1.2.1 Alkylating agents

Alkylating agents are antineoplastic compounds, which chemically react to biological molecules such as; nucleic acids, proteins, amino acids and nucleotides. These agents alter the structure of DNA through cross-links, which can lead to DNA fragmentation. Alkylating agents, such as sulphur mustard gas, were initially used as chemical warfare in World War I, before nitrogen mustards were harnessed for the treatment of cancer. The nitrogen mustard alkylating agents, such as chlorambucil, reacts with nucleophilic sites in DNA through its electron deficient alkyl groups, and was used for the treatment for chronic lymphocytic leukaemia (Siddik, 2005; Vidal *et al.*, 2016). Considered alkylating-like are the platinum analogues; cisplatin and carboplatin, which bind to the N7 position of guanine and/or adenine bases in DNA via a platinum atom (Oronsky *et al.*, 2012). The platinum agents are commonly used as either single agents, or in combination with another drug class for the treatment of cancers of the; breast, testicular, ovarian, cervical, prostate, head

and neck, bladder, lung and refractory non-Hodgkin's lymphomas (Tsimberidou *et al.*, 2009; Dhar *et al.*, 2011; Dasari and Bernard Tchounwou, 2014).

Cisplatin introduces both inter-and intra-strand cross-links in the DNA, which require repair (Pascoe and Roberts, 1974). This results in the inhibition of DNA synthesis and cell division and the induction of apoptotic cell death. Cisplatin has also been found to induce mitochondrial oxidative stress, with an increase of ROS, which results in the loss of mitochondrial protein sulfhydryl group, calcium uptake inhibition and reduction of mitochondrial membrane potential (Marullo *et al.*, 2013a). Furthermore, increased ROS levels in the cell can damage proteins, lipids and further damage DNA adding to the induction of apoptosis.

1.2.2 Antimetabolites

Antimetabolites are folic acid, pyrimidine or purine analogues and have a similar structure to molecules required for DNA or RNA synthesis. Generally, antimetabolites are incorporated into DNA or RNA during S-phase of the cell cycle, or inhibit the enzymes that are required for nucleic acid production, resulting in replication stress and the induction of apoptosis (Huang *et al.*, 2017).

One of the first antimetabolites to be developed for solid tumours was 5-Fluorouracil in 1957, and it remains an essential chemotherapeutic agent, both as a single agent and in combinations with other drug classes for the treatment of colorectal, breast and head and neck cancers (Heidelberger *et al.*, 1957; Arruebo *et al.*, 2011). 5-Fluorouracil is an analogue of uracil, in which a fluorine atom has replaced a hydrogen atom at the C-5 position (Rutman, Cantarow and Paschkis, 1954). It enters the cell via the same facilitated transport mechanism of uracil, before undergoing intracellular conversion into multiple active metabolites, which include; fluorodeoxyuridine monophosphate (FdUMP), flurodeoxyuridine triphosphate (FdUTP) and fluorouridine triphosphate (FUTP) (Wohlhueter, McIvor and Plagemann, 1980; Miura *et al.*, 2010). FdUMP inhibits thymidylate synthase, which is required for the conversion of deoxuridine monophosphate (dUMP) to deoxythymidine monophosphate (dTMP), thereby reducing the nucleotide pool of thymidine required for DNA synthesis (Longley, Harkin and Johnston, 2003a). The accumulation of dUMP can lead to increased levels of deoxyuridine triphosphate (dUTP)(Aherne *et al.*, 1996). Both dUTP and the metabolite FdUTP have been found to be misincorporated into DNA resulting in DNA damage. Furthermore, FUTP has been found to be incorporated into RNA, which disrupts the normal RNA processing and function (Aherne *et al.*, 1996). Together, the active metabolites of 5-Fluorouracil induce DNA damage, and replication stress, which results in the induction of apoptosis.

Gemcitabine is a pro-drug cytidine analogue, which is commonly used as either a single agent, or in combination with another drug class for the treatment of a variety of solid tumours, including; non-small cell lung, pancreatic, bladder and breast cancer (Toschi *et al.*, 2005). Upon influx via nucleoside transporters, gemcitabine undergoes an intracellular conversion to the nucleotides gemcitabine diphosphate (dFdCDP) and triphosphate (dFdCTP), which both have downstream action inducing gemcitabine cytotoxicity. The nucleotide dFdCTP competes with deoxycytidine triphosphate as an inhibitor of DNA polymerase and is incorporated in the DNA. To potentiate this effect, the nucleotide dFdCDP inhibits ribonucleotide reductase, which results in the depletion of deoxyribonucleotide pools required for DNA synthesis (Mini *et al.*, 2006). Furthermore, it has been found that dFdCTP can incorporate into RNA, which can disrupt normal RNA processing, and maturation of ribosomal, transfer and messenger RNA, (rRNA, tRNA, mRNA) precursors (Mojardín *et al.*, 2013). The incorporation of the cytidine analogue into DNA and the reduction of the nucleotide pools results in a disruption of DNA synthesis, increasing replication stress, and the induction of apoptosis.

1.2.3 Topoisomerase inhibitors

Topoisomerase inhibitors target either topoisomerase I or II (topo I, topo II), which are well characterised enzymes involved in the unwinding of DNA during replication and transcription. Chemotherapeutic agents have been developed to target both topoisomerases with camptothecin targeting topo I, and etoposide and doxorubicin targeting topo II (Liang *et al.*, 2019).

Doxorubicin is an anthracycline topo II targeted inhibitor, which has been used as a single agent, or in a combination therapy with other drug classes for the treatment of several cancers including; breast, gastric, lung, ovarian, thyroid, non-Hodgkin's and Hodgkin's lymphoma, sarcoma, multiple myeloma and paediatric cancers (Thorn *et al.*, 2011). The mechanisms of action associated with doxorubicin are multi-factorial, but all lead to increased DNA damage resulting in apoptotic cell death. Doxorubicin has been found to intercalate into DNA, which can result in the inhibition of DNA replication, and disruption of DNA repair mechanisms, which is believed to be mediated by topo II, although the mechanism is not fully known (Yang *et al.*, 2014). Topo II is an ATP-dependent enzyme, which consist of two isoforms; TOP2A and TOP2B. Topo II is known to bind to entangled DNA and DNA supercoils, before breaking both strands of one DNA duplex, passing the other duplex through the gap and resealing the break, in order to reduce torsional stress produced from DNA replication and transcription (Pommier *et al.*, 2010; Yang *et al.*, 2014). It has also been found that Topo II is required for decatenation of DNA during mitosis, and for normal cytokinesis (Carpenter

and Porter, 2004; Yang *et al.*, 2014). It is thought that doxorubicin binds and traps topo II at breakage sites, stabilising the DNA cleavage complex and this prevents the resealing of DNA breaks, resulting in increased DNA damage (Nitiss, 2009). Furthermore, doxorubicin has also been found to oxidize to an unstable metabolite, semiquinone, before being converted back, which releases ROS. The ROS in turn can contribute to the DNA damage and result in cell death (Doroshow, 1986; Kim *et al.*, 2006; Thorn *et al.*, 2011). In addition, doxorubicin has been found to be associated with enhancement of nucleosome turnover around promoter regions, which is thought to be attributed to its ability to intercalate DNA (Yang, Kemp and Henikoff, 2013; Taymaz-Nikerel *et al.*, 2018).

1.2.4 Anti-mitotic agents

Anti-mitotic agents target the dynamics of microtubules, which can disrupt spindle formation and chromosome orientation resulting in mitotic arrest. Microtubules are predominantly formed during interphase and crucial for the completion of mitosis. If the cells remain in a prolonged arrest state, this results in a subsequent apoptosis induction or a senescence-like G_1 state (Mitchison, 2012). Anti-mitotic drugs are divided into two classes based on their mechanism of action; microtubule destabilising agents or microtubule stabilising agents (Kavallaris, 2010).

Destabilising agents inhibit the polymerisation of microtubules and include vinca alkaloids such as; vincristine, vinorelbine and eribulin, which bind to the vinca domain located at the interface between α – and β-tubulin (Jordan and Kamath, 2007; Dumontet and Jordan, 2010; Smith *et al.*, 2010). Eribulin has been approved for patients with metastatic breast cancer (O'Shaughnessy, Kaklamani and Kalinsky, 2019). The binding of eribulin at the vinca domain is selectively at the plus ends of microtubules as well as to soluble tubulin subunits. This prevents the addition of new subunits without affecting normal subunit loss, and results in net microtubule depolymerisation (Natarajan *et al.*, 2012; Lu, Pokharel and Bebawy, 2015).

Stabilising agents promotes the polymerisation of microtubules and include the drug groups; taxanes and epothilones, which are found to bind to the inner surface of the microtubule at the taxoid-binding site on β-tubulin (Altmann, 2001; Jordan and Kamath, 2007). Paclitaxel, a common taxane agent, has been used both as a single agent, or in combination with other drug classes for the treatment of breast, ovarian, colorectal, lung, head and neck cancer (Zhu and Chen, 2019). Paclitaxel binds to the taxoid binding site on the inner surface of the microtubule lattice, along its entire length and promotes microtubule stability and suppression of the microtubule shortening

events, which results in net microtubule polymerisation (Derry, Wilson and Jordan, 1995; Snyder *et al.*, 2001; Jordan and Wilson, 2004; Jordan *et al.*, 2005).

1.2.5 Problems with chemotherapeutic agents

The use of chemotherapeutic agents for the treatment of cancer does not come without problems. For example, the therapeutic window for targeting cancer cells is small, with the assumption that increasing DNA damage through chemotherapy in cancerous cells is attributed to their ability to rapidly divide. However, some normal cells, such as those found in the mouth, stomach, bowel and hair follicles are also rapidly dividing, which can result in undesirable side effects (Falzone, Salomone and Libra, 2018). Long term effects, such as cardiotoxicity, have also been identified with treatment of chemotherapeutic agents, such as doxorubicin, which is more predominant in children and adolescents who receive chemotherapy at a young age (Mancilla, Iskra and Aune, 2019). Furthermore, patients can become refractory to the treatment as resistance develops to the chemotherapeutic agents resulting in therapy failure and patient relapse.

1.3 Resistance to chemotherapeutic agents

1.3.1 The challenge of drug resistance

Resistance to chemotherapeutic agents, and to therapies designed to be selective for specific molecular targets, remains the biggest challenge to the treatment of cancer in the clinic. Most patients with advanced cancer die as the cancer either already exhibits, or develops drug resistance to the therapy, and also to other existing therapies through multi-drug resistance (Garraway and Jänne, 2012; Konieczkowski *et al.,* 2018). The development of drug resistance can result in an increase in tumour mass, invasion of nearby tissues or metastasis to distant tissues, which results in patient death. Approximately 90% of failures to chemotherapy are during the invasion and metastasis of cancer related to drug resistance (Mansoori *et al.*, 2017a).

Drug resistance can be divided into two categories; intrinsic or acquired. Intrinsic resistance indicates that prior to therapy, pre-existing resistance exist in the patient's tumour cells, rendering the therapy ineffective. Acquired resistance develops during drug treatment in tumour cells that have previously demonstrated sensitivity to the drug. This can be through genetic or epigenetic changes that occur in the tumour cells with the application of drug, or the expansion of a small residual population of tumour cells, which are not killed, enabling a regrowth of tumour that no longer responds to the drug (Holohan *et al.*, 2013; Cree and Charlton, 2017). The latter is thought

to be a result of a high degree of molecular cell heterogeneity often found in tumours, resulting in the expansion of resistant minor subpopulations (Swanton, 2012).

Tumour heterogeneity presents a challenge in the study of drug resistance. It has been found that cell heterogeneity can be spatial; within a single tumour and amongst multiple metastases, as well as temporal; in which it is a result of a selective pressure induced by drug therapy (Konieczkowski *et al.,* 2018). It has also been found that simultaneous development of resistance can occur in multiple metastases suggesting that the resistant subpopulations may pre-exist before tumour dissemination (Wagle *et al.*, 2011; Sanborn *et al.*, 2015; Konieczkowski *et al.,* 2018).

Another challenge is also found in the models used to analyse drug resistance. The use of cell line models is still considered a widely accepted method underpinning drug resistance research (Garraway and Jänne, 2012). However, whilst these models are derived from patients, they may not be entirely chemo-naive. These historically obtained cell lines are often derived from patient biopsies with very little information as to what treatment had been received, if any, at the time of the biopsy. Investigation of the year of the biopsy and the country it was obtained from, may provide some useful insight as to what stand of care treatment was given at the time. However, lack of this historical information can hinder the understanding of drug resistance.

A further challenge is that drug resistance is multifactorial. Resistance has been found to develop not only though changes in the cancer genome, such as somatic mutations and chromosomal rearrangements, but also through non-genetic mechanisms such as protein phosphorylation and changes in gene expression. The way in which the resistance effectors develop, or become dysregulated can be extensive for any tumour context (Konieczkowski *et al.,* 2018). In order to overcome refractory tumour phenotypes, an understanding of the mechanisms driving resistance to anti-cancer agents is important.(Garraway and Jänne, 2012).

1.3.2 Mechanisms of drug resistance

Six hallmarks of drug resistance have been proposed to describe common resistance mechanisms, which include; drug pumps, alteration of drug targets, detoxification mechanisms, increased DNA damage repair, reduced apoptosis and altered proliferation (highlighted in figure 1.2) (Cree and Charlton, 2017). These hallmarks, which underly the emergence of drug resistance, can be applied to both chemotherapeutic agents and to molecularly targeted drugs, although chemotherapeutic agents will be the focus of the discussion in the following subsections.

1.3.2.1 Drug pumps

Alterations in drug pumps can result in an increase of drug efflux or a decrease in drug uptake, which ultimately leads to an observed drug resistance phenotype. Increased expression of the ABC transporter family has shown to efflux drugs from the cell at a rate that exceeds their entry in a relatively promiscuous manner. These transporters have been shown to efflux structurally unrelated chemotherapeutic agents, small molecule targeted drugs and xenobiotics, which results in a multi-drug resistance phenotype. For example, the ABCB1 transporter/multi-drug resistance transporter (MDR1) has been found to efflux common chemotherapeutic agents including; doxorubicin, eribulin and paclitaxel (Chen *et al.*, 2016; *Oba et al.,* 2016). The ABCG2/ breast cancer resistance protein (BCRP), has been found to efflux inhibitors such as irinotecan, gefitinib and imatinib (Tsuruoka *et al.*, 2002). Given the problem ABC transporter activity has upon the efficacy of a drug, new drug development programmes often screen compounds to determine their substrate activity for the transporters before continuing drug development (Montanari and Ecker, 2015).

The solute carrier superfamily (SLCs) of membrane transporters contain; the organic anion transporting polypeptide, organic anion transporters and the organic cation transporters (OCTs). Decrease in the expression of SLCs can result in a reduced cellular uptake of chemotherapeutic agents resulting in a drug resistance phenotype (Zhou *et al.*, 2017; J. Zhou *et al.*, 2020). For example, Gao *et al.*, 2019, showed that the decrease of the organic transporter 2 (OCT2/SLC22A2) results in a reduced cellular accumulation of cisplatin. Furthermore, it has been shown that doxorubicin is imported into the cell via the solute carrier family 22 member 16 (SLC22A16/OCT6) in a sodium-independent manner (Okabe *et al.*, 2005a, 2005b; Muley *et al.*, 2020).

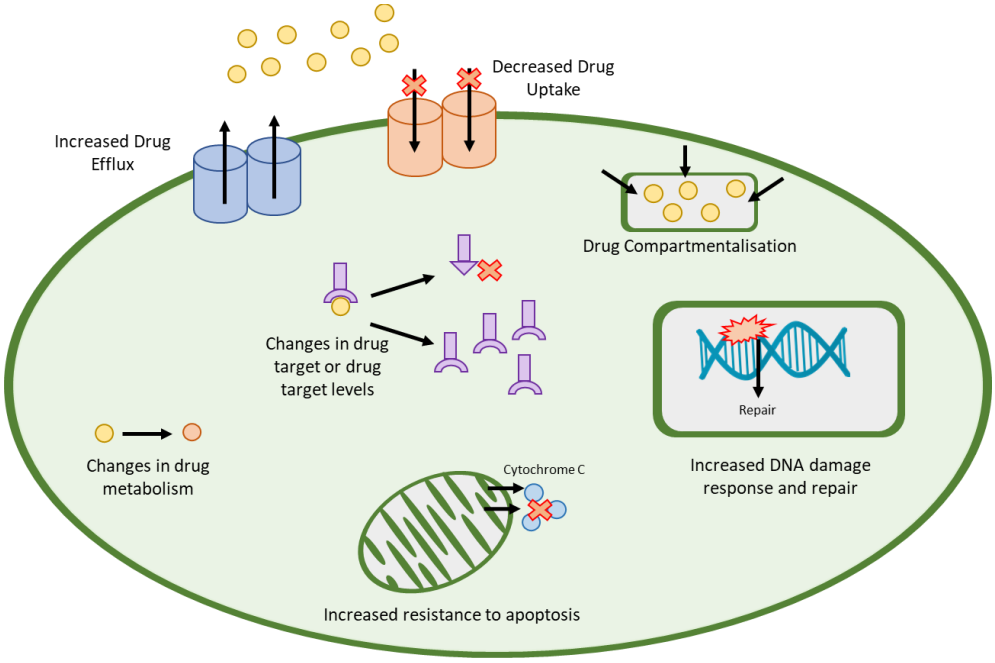


Figure 1.2 Mechanisms of drug resistance in cancer cells
Schematic diagram identifying the common mechanisms of drug resistance in cancer cells. This includes increased drug efflux, and decreased drug uptake, drug compartmentalisation, changes in drug target or drug target levels, changes in drug metabolism, increased resistance to apoptosis and increased DNA damage response and repair. Adapted from (Mansoori *et al.*, 2017).

1.3.2.2 Detoxification mechanisms

If uptake of the anti-cancer agents is efficient, there are many detoxification mechanisms that reduce the cytotoxicity of anti-cancer agents, which occur within the cell. These mechanisms can involve directly changing the drug, prevention of required drug metabolism needed to form active metabolites or drug compartmentalisation.

Inactivation of platinum agents have been found associated with conjugate formation between the drug and the thiol glutathione (GSH) catalysed by glutathione S-transferase (GSTπ). This has been found to result in a detoxification of the molecule and also increases the affinity to ABC transporters (Delou *et al.*, 2019) (see sub- section 1.3.2.1 Drug pumps). Upregulation of GSTπ and its activity has been implicated as a detoxification resistance mechanism in several cancer types (Cazenave *et al.*, 1989; T *et al.*, 1992; Di Nicolantonio *et al.*, 2005). Furthermore, GSH can act as an antioxidant, which can inhibit oxidative stress, such as ROS, which is found increased during treatment with cisplatin. Both cisplatin resistant lung and ovarian cancer cell lines have been shown to have higher levels of GSH, counteracting the high levels of ROS induced by cisplatin (Chen and Kuo, 2010; Catanzaro *et al.*, 2015, 2018).

Some chemotherapy agents, such as gemcitabine and 5-Fluorouracil, undergo complex metabolism pathways to form active metabolites, as previously discussed (sub-section 1.2.2 Antimetabolites). Altered pathways of drug metabolism can prevent the formation of the active metabolite, detoxifying the agent and resulting in an observed resistant phenotype. For example, an increase of expression of the 5-Fluorouracil activating enzymes; thymidine phosphorylase (TP), uridine phosphorylase (UP), and orotate phosphoribosyl transferase (OPRT), have been associated with 5-Fluorouracil sensitivity (Houghton and Houghton, 1983; Schwartz *et al.*, 1985; Evrard *et al.*, 1999; Longley and Johnston, 2005). Furthermore, 5-Fluorouracil can be catabolised by dihydropyrimidine dehydrogenase (DPD) (Longley and Johnston, 2005). Diasio and Harris, 1989, have reported that overexpression of DPD in cancer cell lines results in resistance to 5-Fluorouracil, which was also confirmed by Takebe *et al.*, 2001, who showed that high levels of DPD mRNA expression in colorectal tumours correlated with 5-Fluorouracil resistance.

Sequestration of drugs away from the drug target and into cellular compartments is another mechanism that results in drug detoxification. This drug compartmentalisation is often as a result of defects in lysosomal and protein trafficking (Vadlapatla *et al.*, 2013). One such example, is where cisplatin is sequestered within the vesicle structures of the lysosome, golgi and secretory compartments before being effluxed from the cell (Katano *et al.*, 2004; Liang *et al.*, 2005; Safaei *et al.*, 2005; Vadlapatla *et al.*, 2013). The ABC transporter, ABCA3, has been implicated in exosome biogenesis and the efflux of vesicle structures in a drug resistance context (Safaei *et al.*, 2005; Chapuy *et al.*, 2009; Vadlapatla *et al.*, 2013; Overbeck *et al.*, 2017).

1.3.2.3. Alteration of drug targets

Another mechanism of resistance to anti-cancer drugs is alteration of the drug target, which can be a result of either a mutation or changes in the expression levels due to epigenetic alterations (Wang, Zhang and Chen, 2019). Whilst this mechanism is commonly associated with newer small molecule targeted agents, the truism is associated with chemotherapeutic agents too (Cree and Charlton, 2017). As previously mentioned, part of the cytotoxicity induced by 5-Fluorouracil is the inhibition of thymidylate synthase. Increased expression of thymidylate synthase has been implicated as a mechanism of resistance to the treatment of 5-Fluorouracil (Longley, Harkin and Johnston, 2003b). Mutations in the drug target have been seen in topo I, which results in the inability of camptothecin to effectively bind and carry out the intended cytotoxic mechanism, resulting in an observed resistance phenotype (Larsen and Skladanowski, 1998).

1.3.2.4 Changes in DNA damage repair

Commonly, chemotherapeutic agents introduce DNA damage as their mechanism of action, which can ultimately lead to cell death. To circumvent this, a frequent resistance mechanism is to upregulate the DNA repair pathways to remove the exogenous DNA damage. This can be through either aberrant expression or mutation of DNA repair components or regulators of DNA repair. Nucleotide excision repair (NER) is a highly conserved DNA repair pathway that repairs DNA lesions such as those introduced by cisplatin. The excision repair cross-complementation group 1 (ERCC1) protein plays a key role in NER. Olaussen *et al.*, 2006, showed through an international adjuvant lung cancer trial, which retrospectively conducted an immunohistochemical analysis of ERCC1 in tumour samples of 761 patients with metastatic lung cancer, that there was a statistically significant survival benefit in patients with low levels of ERCC1 when receiving platinum-based chemotherapy, compared to those that had high levels of ERCC1. Similar results were seen in retrospective clinical trials with ovarian and colorectal cancer (Dabholkar *et al.*, 1992; Metzger *et al.*, 1998; Kang *et al.*, 2006; Martin, Hamilton and Schilder, 2008). An alternative example is seen when a DNA repair pathway is switched off to increase tolerance to exogenous DNA damage. Normally, mismatch repair (MMR) pathway recognises cisplatin-induced DNA adducts, which results in successive repair cycles that ultimately triggers apoptosis. However, a loss of function in the MMR pathway results in reduced recognition of DNA damage, cell death is not efficient, which thereby promotes tolerance to cisplatin, and also increases the development of mutations, which can lead to further carcinogenic transformation (Povey *et al.*, 2002; S. *et al.*, 2013). This latter example is also a common resistance mechanism for 5-Fluorouracil, which has been established in colorectal cancer (Mark Meyers *et al.*, 2001).

1.3.2.5 Reduced apoptosis

As mentioned in the previous sub-section, increased DNA damage is seen with the application of many chemotherapeutic agents, which ultimately result in cell death via the induction of apoptosis. Therefore, another common resistance mechanism is to increase cell pro-survival signalling, and reduce signalling through pathways, which result in apoptosis. It must be noted, that other forms of cell death can also be triggered by chemotherapeutic agents, which include necrosis, necroptosis and autophagy, and changes in these pathways can also lead to a resistant phenotype (Kroemer *et al.*, 2009; Cree and Charlton, 2017). One example is the loss of function of a critical tumour suppressor gene, *TP53*, which has multiple downstream targets that can induce apoptosis (Fridman and Lowe, 2003). Whilst p53 has been found to be mutated in more than 50% of cancer patients,

mutated p53 has also been implicated in driving chemoresistance due to loss of function on activating transcription of PUMA, a key pro-apoptotic protein (Huang *et al.*, 2019).

Another example, is the activation of nuclear factor kappa-light-chain-enhancer of activated B cells (NF-κB) pathway, which has been found to rescue cancer cells from the apoptotic pathway, promotes survival and proliferation and prevents cell death (Xia *et al.*, 2018; Delou *et al.*, 2019). Korber *et al.*, 2016, found a chemoresistance phenotype to 5-Fluorouracil in colonic carcinoma cells lines, which was strongly dependent on NF-κB activation. Furthermore, anti-apoptosis proteins including Bcl-XL and Bcl-2, have been found upregulated via the activation of NF-κB during the emergence of chemoresistance in invasive pancreatic cancer (Greten *et al.*, 2002; Y. Li *et al.*, 2016; Li *et al.*, 2018).

1.3.2.6 Altered proliferation

Activation of the phosphatidylinositol 3-kinase (PI3K) / serine/threonine kinase (AKT) pathway can result in increased cellular functions including survival, proliferation, migration and differentiation (West, Castillo and Dennis, 2002). Upregulation of this pathway, results in increased proliferation and has been implicated in resistance to anti-cancer agents including chemotherapeutic agents such as doxorubicin (Christowitz *et al.*, 2019). Both *in vitro* and *in vivo* studies have shown that combination of small molecule inhibitors of PI3K/AKT pathway with a chemotherapeutic agent are successful in attenuating chemotherapeutic resistance (West, Castillo and Dennis, 2002). Altered proliferation can become a self-perpetuating mechanism, whereby cancer cells with a high rate of proliferation are genetically unstable, resulting in the development of further drug resistant mechanisms through genetic aberrations (Mansoori *et al.*, 2017c).

1.3.3 Overcoming resistance to chemotherapeutic agents

Given the clinical implications resistance to chemotherapeutic agents have on the outcome of patient survival, it is important to consider methods of overcoming resistance. One such approach, is to use both a small molecule inhibitor and a chemotherapeutic agent in combination therapy. For example, several ABC transporter inhibitors have been developed, such as zosuquidar, which when combined with a known substrate chemotherapeutic agent, attenuates the chemotherapeutic resistance in cells (Nanayakkara *et al.*, 2018). However, although the use of ABC transporter inhibitors has led to success in MDR reversal in preclinical studies, little impact was seen on clinical outcome (Falasca and Linton, 2012). Many ABC transporter inhibitors have been clinically tested over the last forty years, but the use in the clinic is impeded by severe toxicities, and as such no

effective agent has been developed and approved to date (Choi and Yu, 2014; J. M. A. Delou *et al.*, 2019). However, fourth-generation ABC transporter inhibitors have emerged, which are based on either natural or semi-synthetic compounds, and are under investigation (Karthikeyan and Hoti, 2015; J. M. A. Delou *et al.*, 2019).

Another approach to overcome drug resistance, is to determine an appropriate second line therapy. For a second line therapy to be effective, the patient must not demonstrate cross-resistance, and to this end, the drugs selected are normally of a different drug class to the first line therapy. Within a patient setting, investigation of new drug classes as an appropriate second line therapy is difficult, as patients entering clinical trials have often undergone several treatment options before entering a clinical trial, often as a last treatment resort. A new drug class of small molecule inhibitors of the DNA damage response have been developed, many of which are undergoing clinical trials both as monotherapies or in combination with other drugs. However, they have not been investigated as an appropriate second line therapy after the development of chemo-resistance.

1.4 DNA damage response

1.4.1 Signalling in the DNA damage response

In order to preserve genomic integrity after DNA damage through endogenous or exogenous stress, the cells must identify and repair the damaged DNA. This is conducted through the DNA damage response (DDR), which can be broken down into three integral steps; 1) recognition of the DNA lesions 2) downstream cascade of DNA repair signalling and cell cycle arrest and 3) induction of apoptosis if the damage is irreparable (Blackford and Jackson, 2017).

1.4.1.1 Recognition of DNA lesions

The DDR is driven by signal transduction with downstream cascades of protein phosphorylation. Initial detection of DNA damage and cellular signalling is predominantly instigated through three phosphoinositide 3-kinase (PI3K) related kinases (PIKKs); Ataxia-Telangiectasia Mutated (ATM), Ataxia-Telangiectasia and Rad3 related (ATR), and DNA-dependent protein kinase (DNA-PK) (Blackford and Jackson, 2017). Both ATM and DNA-PKcs are active in the recognition of DNA double strand breaks (DSBs), whilst ATR is predominantly activated by DNA replication stress, or lesions, which result in single-stranded DNA (ssDNA). Whilst DNA-PKcs are thought to be limited to the repair of DSBs at the site of the lesion, both ATR and ATM have both a local and global cellular responses through the phosphorylation of the downstream effectors; checkpoint kinase 1 (CHK1) and checkpoint kinase 2 (CHK2) respectively (Blackford and Jackson, 2017). It must be noted that

ATM, ATR and DNA-PKcs have overlapping signalling pathways and substrates, and ATR and ATM have been found to substitute for each other, suggesting some role redundancy (Brown and Baltimore, 2003; Wang *et al.*, 2004).

Each of the DDR kinases are recruited and activated to the site of the damaged DNA through individual co-factors dependent on the type of DNA damage that has occurred. The ATR-interacting protein (ATRIP) has been found to bind to the heterotrimeric replication protein A (RPA), a protein complex that associates with ssDNA, and ATRIP recruits ATR to the site of ssDNA (Cortez *et al.*, 2001; Zou and Elledge, 2003). Further protein recruitment to the ssDNA site is required for full ATR activation. Ewing's tumour-associated antigen 1 (ETAA1) is recruited to RPA-coated ssDNA sites, where it binds to ATR through the ATR-activation domain, stimulating ATR (Bass *et al.*, 2016; Haahr *et al.*, 2016). DNA topoisomerase 2-binding protein 1 (TOPBP1) has also been found to contain an ATR-activation domain, and the activation of ATR by TOPBP1 is a crucial step in the initiation of ATR dependent signalling, but the recruitment and mechanism are not yet clear (Kumagai *et al.*, 2006; Blackford and Stucki, 2020). Both ATM and DNA-PKc are activated through their recruitment to the DNA lesions by their co-factors. ATM has been found to be recruited and activated by the MRE11-RAD50-NBS1 (MRN) complex, which can recognise and bridge broken double stranded DNA (dsDNA) ends (Lee and Paull, 2004, 2005). DNA-PKc is known to be recruited and activated to dsDNA ends by a heterodimeric complex, Ku, which contains two subunits; Ku70 and Ku80 (Gottlieb and Jackson, 1993).

One of the key roles of ATM and ATR during DNA lesion recognition, is the activation of cell cycle checkpoints in order to arrest the cells at G1/S or G2/M boundaries to allow for DNA repair (Kastan and Bartek, 2004). This is predominantly driven through activating phosphorylation of the cell cycle checkpoint proteins; CHK1 and CHK2. Although cross-talk has been identified in these activation pathways, the end result is the inhibition of the cyclin -dependent kinase (CDK) activity, which is known to drive cell cycle progression (Blackford and Jackson, 2017). Upon recognition of DNA damage, ATM/ATR rapidly activate CHK1/CHK2 respectively, which inactivate the cell division cycle 25 (CDC25) family of phosphatases, which counteracts the inhibitory phosphorylation of cyclin dependent kinases (CDKs) by the wee1-like protein kinase (WEE1)(Squire *et al.*, 2005). Furthermore, ATM/ATR can both phosphorylate p53, which results in a slower transcription of CDKs by $p21^{CIP1/WAF1}$, which is key for the progression of the G1/S checkpoint (Blackford and Stucki, 2020). Activation of CHK1/CHK2 also promotes DNA repair and activate apoptosis pathways (Patil, Pabla and Dong, 2013). With the DNA lesion recognised and progression of the cell cycle halted (Figure 1.3), effective DNA repair can occur.

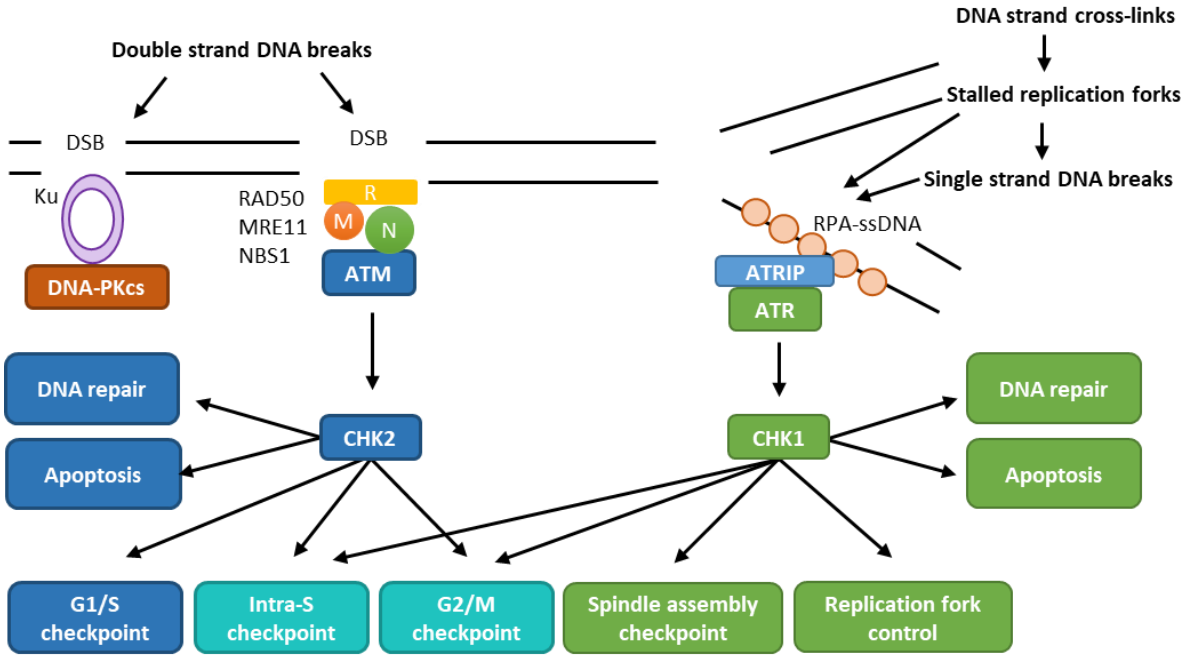


Figure 1.3 DNA lesion recognition and cell cycle halt
Exogenous or endogenous DNA damage result in the recruitment of the three PIKKs to recognise the DNA lesions. DSBs damage results in the recruitment and activation of DNA-PKcs by Ku-bound to DSB ends, and ATM to DSBs by the MRE11-RAD50-NBS1 (MRN) complex. ATR is recruited to RPA-coated ssDNA by its binding partner ATRIP after damage, which produces ssDNA such as DNA strand crosslinks or stalled replication forks. ATR and ATM activate the downstream CHK1 and CHK2 kinases to halt the cell cycle to prevent entry into the S-phase, delay progression through S -phase and stop the cells from entering mitosis. Activation of checkpoint kinases also initiates DNA repair and if repair fails or there is prolonged activation of the checkpoints, the kinases trigger apoptosis. Figure adapted from both Garrett and Collins, 2011 and Blackford and Jackson, 2017.

1.4.1.2 DNA repair signalling

There are many different agents that can induce DNA damage, and their mechanisms results in the formation of distinct types of damage. Multiple cellular DNA repair mechanisms exist to address the damage, and the selection of the appropriate DNA repair pathway is crucial after damage has been detected. Table 1.1 outlines a summary of the different repair pathways available for the repair of damaged DNA by common DNA damaging agents (adapted from a figure by Chatterjee and Walker, 2017). DNA repair pathways are available to repair bulky lesions, or incorrect base incorporation, but two of the most common DNA repair mechanisms are the repair of DSBs and single strand breaks (SSBs). It must be noted that the selection of the distinct DNA repair pathways is often dependent on the stage of the cell cycle (Hustedt and Durocher, 2017).

Table 1.1 Summary of repair mechanisms for damaged DNA

		DNA damage	DNA repair pathways
DNA damaging agents	Toxins, alkylating agents, Base deamination, replication errors	Mismatches, uracil, abasic sites, adducts	Mismatch repair, base excision repair
	Oxidative damage, eletrophiles	Lesions, single strand break, double strand break	Base excision repair, single strand break repair, double strand break repair
	Ionizing radiation, UV radiation, crosslinking agents, aromatic compounds, heat cold hypoxia	Bulky lesions, intra-interstrand crosslink repair, single strand break repair, double strand break repair, translesion synthesis	Nucleotide excision repair, interstrand crosslink repair, single strand break repair, double strand break repair, translesion synthesis

DSBs are commonly repaired through two distinct mechanisms; non-homologous end joining (NHEJ) and homologous recombination repair (HRR). Classical NHEJ (cNHEJ; to distinguish from alternative end joining, which functions in the absence of cNHEJ proteins), is a rapid high capacity pathway, which works to ligate two broken DNA ends without needing a repair template, predominantly during G1 phase of the cell cycle (Branzei and Foiani, 2008; Scully *et al.*, 2019). Briefly, cNHEJ is initiated by Ku and DNA-PKcs binding to the DSBs, promoting DNA-end tethering (Graham, Walter and Loparo, 2016). Subsequently, additional cNHEJ core factors are recruited to allow for the ends to be closely aligned and ligated, which includes, XRCC4, XLF and DNA ligase IV, and stabilised by PAXX (Blackford and Jackson, 2017). Some DSBs require additional proteins such as DNA polymerases and nucleases in order for repair to occur (Blackford and Jackson, 2017). DSBs are mostly repaired by cNHEJ, and although it has been previously described as error prone, in relation to the development of CRISPR-Cas9 assays, it is also described as efficient and mostly accurate (Bétermier, Bertrand and Lopez, 2014; Hsu, Lander and Zhang, 2014; Blackford and Jackson, 2017).

The HRR pathway relies on the presence of a homologous donor template and is predominately used to repair DSBs at replication forks during the S phase and G2 phase of the cell cycle (Beucher *et al.*, 2009; Karanam *et al.*, 2012; Scully *et al.*, 2019). One of the key determinants of the DSB repair pathway of choice is DNA-end resection in order to produce ssDNA, a perquisite required for HRR (Mladenov *et al.*, 2019; Scully *et al.*, 2019). Briefly, the two blunt ended DSB ends are converted to ssDNA tail through the endonuclease activity of MRN, EXO1 and the DNA2-Bloom syndrome protein (BLM). The activity of MRN displaces any Ku present, committing to HRR as the repair pathway. With available ssDNA overhangs, RPA binds and this activates ATR/ATRIP, which phosphorylates CHK1 initiating HRR (Sørensen *et al.*, 2003). RPA is then replaced with recombinase DNA repair protein RAD51 homolog 1 (RAD51) (Heyer, Ehmsen and Liu, 2010). RAD51 is loaded onto ssDNA,

forming a nucleoprotein filament, which can then form synaptic complexes that contain a three-stranded DNA helix intermediate supporting the formation of the hetero-duplex DNA. This is composed of the invading strand and the complementary strand (Chen, Yang and Pavletich, 2008; Scully *et al.*, 2019). If base-pairing is successful, the synapse is stabilised, and the non-base-paired strand of the invading molecule is displaced to form the displacement loop (D-loop). A DNA polymerase binds to the invading strand to extend the invading strand using the donor DNA molecule as a template (Scully *et al.*, 2019).

One ended DSBs are created when a replication fork collapses, or when a replication fork collides with a SSB, and requires activation of HRR (Saleh-Gohari *et al.*, 2005). Complex signalling networks are required to stabilise the replication fork and to delay cell cycle progression to repair the damage (Liao *et al.*, 2018). Briefly, stalled replication forks are protected from degradation by SWI/SNF-related matrix-associated actin-dependent regulator of chromatin subfamily A-like protein 1 (SMARCAL1), DNA annealing helicase and endonuclease ZRANB3 (ZRANB3) or HRR factors (Schlacher, Wu and Jasin, 2012; Poole and Cortez, 2017). Fork reversal can be mediated by poly[ADP-ribose] polymerase 1 (PARP1) to prevent formation of DSBs by protecting the replication fork from colliding with SSBs (Liao *et al.*, 2018). PARP1 can also induce the synthesis of poly(ADP-ribose) chains, which can act as an signal for other DNA repair proteins. If the fork collapses, SMARCAL1 will be recruited to RPA-ssDNA at the replication forks to stabilise (Puccetti *et al.*, 2019). As before, RPA-ssDNA at stalled replication forks activates the ATR-CHK1 axis. CHK1 is recruited to the replication fork where it stabilises the replication fork, regulates the activation of replication origins, regulates elongation and delays S-phase progression (Meyer *et al.*, 2020). During the stabilisation of the replication forks, CHK1 also activates HRR through an activating phosphorylation of RAD51 and BRCA2 (Sørensen *et al.*, 2003).

If replication stress is too high and fork progression is delayed too long, or the DNA damage is not resolved, as a last resort, the cells activate the mechanisms, which result in cellular death. This programmed cell death is to ensure DNA fidelity is maintained.

1.4.1.3 Induction of apoptosis

Apoptosis is a secondary response to DNA damage with the main goal of protecting a multicellular organism against a damaged cell. If DNA damage is too severe to effectively repair, or the cell has been arrested for a considerable amount of time, apoptosis is induced (Roos, Thomas and Kaina, 2016). Apoptosis is a series of coordinated signalling events of programmed cell death which are

activated through two major signalling pathways; extrinsic (death receptor) pathway or intrinsic (mitochondrial) pathway.

Briefly, the extrinsic pathway is mediated by the activation of cell surface receptors, which transmit apoptotic signals after binding with specific ligands, termed; death receptors (Elmore, 2007; Nowsheen and Yang, 2012a). Signalling is then mediated by the cytoplasmic part of the death receptor, which through the death domain (DD) conserved sequence, can bind with adaptor proteins, such as Fas-associated death domain (FADD) protein or Tumour necrosis factor receptor type 1-associated DEATH domain (TRADD), forming the death inducing signalling complex (DISC) which further disseminates the signal (Elmore, 2007; Nowsheen and Yang, 2012a). FADD also contains the death effector domain (DED), which can sequester procaspase-8 to the DISC, resulting in an autocatalytic activation due to autoproteolysis, which releases caspase-8 activating the effector caspases, and resulting in cell death (Elmore, 2007; Nowsheen and Yang, 2012a).

The intrinsic pathway is predominantly triggered by cellular stress, specifically mitochondrial stress caused by DNA damage and heat shock (De Zio, Cianfanelli and Cecconi, 2013). The inner mitochondrial transmembrane potential is disrupted, resulting in permeability and the release of proapoptotic proteins from the mitochondrial intermembrane space into the cytoplasm (Elmore, 2007; Nowsheen and Yang, 2012a). This includes cytochrome c, which activates the apoptosome, resulting in the activation of the caspase cascade. Activation of caspase-9 and the subsequent proteolytic effector procases-3, -6 and -7 cleaves protein substrates which result in both the mediation and amplification of the death signal, resulting in cell death (Elmore, 2007; Nowsheen and Yang, 2012a). The pathways of DDR and apoptosis are closely regulated and converge (Nowsheen and Yang, 2012b). This introduction will only discuss key links of this convergence, however there are many other ways cellular death can be triggered, which are beyond this scope.
The tumour suppressor protein, p53, is a key protein for the mediation of cellular response to stress, with its ability to initiate DNA repair, cell-cycle arrest, senescence and apoptosis (Nowsheen and Yang, 2012b). It is considered to be the balance between survival and death following DNA damage (Roos, Thomas and Kaina, 2016). In the event of DNA repair remaining unresolved, p53 can initiate apoptosis by transactivating the pro-apoptotic proteins, including BAX, BID, NOXA and PUMA, which results in the permeabilization of the mitochondrial membrane and the release of pro-apoptotic factors (Fridman and Lowe, 2003). Both the extrinsic or the intrinsic apoptotic pathways can be stimulated via p53 induction depending on the DNA damage. Furthermore, p53 has been shown to bind to the outer mitochondrial membrane and antagonise the anti-apoptotic functions of BCL2 and BCL-XL (Nowsheen and Yang, 2012b). Through transcription, p53 can also control the

permeabilization of the mitochondrial membrane by activating the pro-apoptotic protein BAX, or neutralising the anti-apoptotic proteins BCL2 and BCL-XL (Nowsheen and Yang, 2012b). Importantly, ATM, ATR, CHK1, CHK2 and DNA-PK can phosphorylate p53 on Ser15, Ser37, Thr18 and Ser20, which results in the uncoupling of p53 from its inhibitory MDM2 binding partner to carry out its role in resolving cellular stress induced by apoptosis (Roos, Thomas and Kaina, 2016).

DNA-PK has also shown to have another role in triggering of apoptosis in response to severe DNA damage (Puccetti *et al.*, 2019). DNA-PK can undergo proteasomal degradation during the apoptotic process, which results in the suppression of pro-survival signals (Burma and Chen, 2004). The Ku70 binding partner of DNA-PK has been shown to suppress apoptosis by sequestering BAX from the mitochondria, whilst acetylation of Ku70 can disrupt this interaction (Sawada *et al.*, 2003; Cohen *et al.*, 2004; Nowsheen and Yang, 2012b).

BRCA1 has been shown to enhance p53-independent apoptosis when present in the cytoplasm (Gu *et al.*, 2010). Overexpression of BRCA1 induces apoptosis and it is thought to be linked to BRCA1 nuclear export during DNA damage events, and to the c-Jun N-terminal kinase pathway (Harkin *et al.*, 1999; Nowsheen and Yang, 2012b) . Furthermore, BRCA1 has been found in the mitochondria where it can promote BCL2 mediated apoptosis. This depletes BRCA1 from the nucleus, which results in decreased mediated-HRR (Laulier *et al.*, 2011; Nowsheen and Yang, 2012b). ATR has also been shown to mediate the phosphorylation of BRCA1 after UV DNA damage, demonstrating a convergence of p53 and BRCA1 mediated apoptosis (Deng, 2006).

1.4.2 Targeting the DNA damage response for therapeutic benefit

An early event of cancer progression is the abrogation of the DDR, which can result in further increased genomic instability and increased mutation rate, further facilitating cancer progression. Many genetic instability disorders, such as Fanconi Anaemia, or Li-Fraumeni syndrome, have a high cancer incidence due to mutations in genes involved in the DDR (Martin and Smith, 2007). Further to this, as discussed earlier (1.3.2.4), dysregulation of the DDR can promote drug resistance and result in patient therapy failure. The rationale of targeting the DDR pathways for therapeutic benefit follows the concept that the cancer cells have defects or dysregulation of these pathways, and have been shown to demonstrate a greater dependence on the remaining functional DDR processes for both cancer progression and drug resistance (Garrett and Collins, 2011). To this end, many small-molecule inhibitors targeting the DDR have been developed with the reasoning that in combination with DNA-damaging chemotherapeutic agents, inhibition of the remaining DDR

pathways will result in an increase of drug induced cytotoxicity (Garrett and Collins, 2011). Evidence has also been seen whereby inhibitors of the DDR can also work as a monotherapy, with the rationale that pre-existing defects in the DDR in the cancer cells in combination with a DDR inhibitor can result in synthetic lethality (Martin and Smith, 2007).

The principle of synthetic lethality is based on the idea that simultaneous perturbation of two genes, or in this case DNA repair pathways, results in cellular death (Nijman, 2011). This can be achieved through genetic or epigenetic perturbation of one DNA repair pathway in combination with an inhibitor of another (Figure 1.4). An example of this scenario, is the combination of a deleterious mutation in *BRCA1* or *BRCA2* gene in combination with a PARP inhibitor. PARP inhibitors, such as olaparib and rucaparib, function by trapping PARP on the DNA during the repair of SSBs resulting in replication fork stalling and the formation of DSBs. In *BRCA1*/*BRCA2* proficient cells, DSB repair occurs via HRR and the replication forks restart, resulting in cell survival. In *BRCA1*/*BRCA2* deficient cells, HRR is impaired, which results in DSB accumulation and cellular death (Dziadkowiec *et al.*, 2016). Treatment with the PARP inhibitor, olaparib, is now approved for BRCA-deficient ovarian cancer patients (Montemorano, Michelle and Bixel, 2019).

The concept of synthetic lethality for the development of new drug targets and applications has led to high-through-put genome-wide screens using CRISPR-Cas9. Screening of CRISPR-Cas9 gene knockouts against panels of drugs have generated new synthetic lethal combinations, and this is expected to become a more common place approach as CRISPR-Cas9 gene knockout cell lines become more accessible (O'Neil, Bailey and Hieter, 2017). Recently Wang *et al.*, 2019, identified through a genome-wide CRISPR screen that deficiency of RNASEH2 is synthetically lethal with the inhibition of ATR.

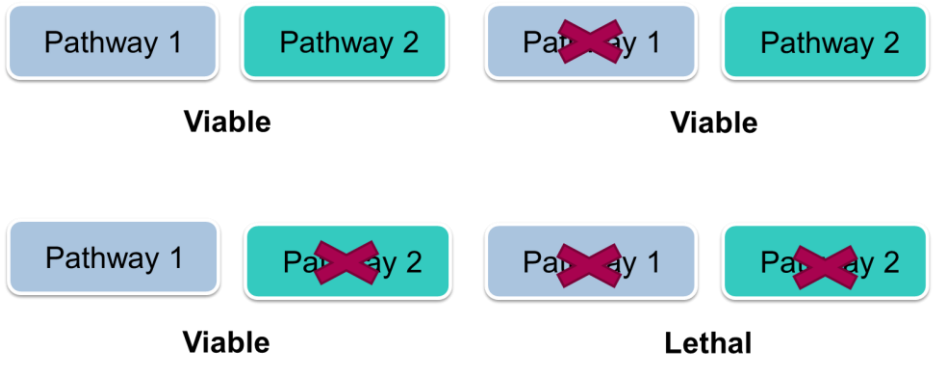


Figure 1.4 Synthetic lethality of DNA repair pathways
Schematic representation of synthetic lethality. Two pathways are synthetic lethal only when their simultaneous inactivation results in cellular death. In this example, failure of pathway 1 or pathway 2 does not affect viability whereas inactivation of both at the same time is lethal. This can occur through genetic and/or epigenetic changes in combination with a small molecular targeted inhibitor.

As previously discussed, ATR, ATM and DNA-PK orchestrate the recognition of DNA lesions and subsequent DNA repair signalling cascades. Several drugs have been developed in order to target and inhibit these proteins to result in increased DNA damage and cell death. Several inhibitors of ATR are in clinical trials including M6620, BAY1895344, berzosertib and ceralasertib, and when used as a monotherapy can increase levels of DNA replication stress (Mei *et al.*, 2019). Of note, ceralasertib is currently in twenty-two trials both as a single agent or in a drug combination. This includes a phase II trial in combination with the PARP inhibitor olaparib, as inhibition of ATR and PARP have been reported to be synthetically lethal (Turner *et al.*, 2008; Mei *et al.*, 2019; Warren and Eastman, 2020). The ATM inhibitor; AZD0156 has shown to exhibit activity, which can be potentiated in combination with agents, which induce DSBs, such as chemotherapeutic agents and irradiation (Riches *et al.*, 2020). AZD0156 is currently in a phase I study to determine preliminary efficacy of ascending doses of the drug either a monotherapy or in combination with chemotherapies in patients with advanced malignancies (NCT02588105)(Riches *et al.*, 2020). DNA-PK inhibitors have also shown to sensitise tumours to DSB inducing agents, such as chemotherapy and radiation, and combination phase I clinical trials are currently underway for several DNA-PK inhibitors including, M3814 (NCT02516813) and CC-122 (NCT01421524) (Mohiuddin and Kang, 2019). In addition a dual DNA-PK and mTOR inhibitor has just completed a phase II study individualised screening trial of innovative glioblastoma therapy (INSIGhT)(NCT02977780) (Alexander *et al.*, 2017; Mohiuddin and Kang, 2019).

Inhibitors have also been developed targeting the ATR/ATM downstream effectors; CHK1 and CHK2. Inhibitors of CHK1/CHK2 can modulate the response of the cell to DNA damage, and can lead to the abrogation of the cell cycle checkpoints, inhibit DNA repair and change the regulation of apoptosis (Garrett and Collins, 2011). Many CHK1 inhibitors have been developed, such as MK-8776, prexasertib, rabusertib and SRA737, and have shown to kill tumour cells that present high levels of replication stress. Of note, prexasertib, which was considered to be a dual CHK1/CHK2 inhibitor entered phase II studies, but was terminated in April 2019, likely due to a high rate of observed toxicity (Warren and Eastman, 2020). The CHK1 inhibitor, SRA737, is the only CHK1 inhibitor currently undergoing further clinical development and has just completed two phase I/II trials, one as a monotherapy and the other in combination with gemcitabine (Banerji *et al.*, 2019; Plummer *et al.*, 2019; Warren and Eastman, 2020). Currently, there are no compounds, which target only CHK2 in clinical trials. However, Cancer Research UK are looking for a commercial partner for the further development of a CHK2 inhibitor, CCT241533, which has shown a good selectivity profile for CHK2 and of which has demonstrated activity that potentiates the cytotoxicity of PARP

inhibitors. More medicinal chemistry development is required to obtain appropriate pharmacological properties based on CCT241533 (Anderson *et al.*, 2011; Caldwell *et al.*, 2011; *Chk2 Inhibitor Programme | Commercial Partnerships | Cancer Research UK*, 2020).

Another protein critical for DSB repair through HRR, and critical for DNA replication, is RAD51. RAD51 forms nucleoprotein filaments at damaged DNA sites, or stalled replication forks, in order to allow for break repair and replication fork start (Mills *et al.*, 2017). A handful of RAD51 inhibitors have been developed including B02, CYT01B and CYT-0851 (Huang *et al.*, 2012; Mills *et al.*, 2017). Of note, a phase I/II clinical trial is currently recruiting to test CYT-0851 in B-cell malignancies and advanced solid tumours (NCT03997968).

WEE1 is another attractive target involved in the DDR and is found down stream of CHK1 where it regulates the progression of the cell cycle at the G2/M boundary. WEE1 is a negative regulator of CDK1, whereby the inhibitory phosphorylation at tyrosine 15, deactivates CDK1 and prevents mitotic entry (De Witt Hamer *et al.*, 2011; Koh *et al.*, 2018). Inhibition of WEE1 has been shown to augment the effects of DNA-damaging agents (Rajeshkumar *et al.*, 2011; Kausar *et al.*, 2015). Of note, the WEE1 inhibitor, adavosertib, is involved in active phase I/II clinical trials, as well as trials that are currently recruiting. Most of the trials are in combination with chemotherapy agents or radiotherapy, but it is interesting to note a phase II trial, which is recruiting to use adavosertib as a monotherapy in advanced solid tumours that have a mutation in *SETD2* (NCT03284385). It has been shown that WEE1 selectively kills *SETD2* deficient cancers through deoxynucleoside triphosphate starvation (Pfister *et al.*, 2015).

Also downstream of CHK1 are the polo and aurora kinases. It has been observed that CHK1 negatively regulates polo-like kinase 1 (PLK1), which is required for mitotic progression, and CHK1 contributes to the activation of aurora B at the spindle checkpoint in response to tension-lacking kinetochores (Tang, Erikson and Liu, 2006; Yu, 2007). It is also important to note there is interaction between polo kinases and aurora kinases, and the pathways have been found to converge. One example is that in which PLK1 gets activated through Bora and Aurora A kinase (Chopra *et al.*, 2010). With their role in cell cycle progression at the G2/M boundary and within the M phase at the spindle checkpoint, inhibitors targeting polo and aurora kinases are also termed second generation mitotic drugs, with the first generation being the anti-mitotic agents discussed early in this chapter (1.2.4).

Polo kinases play a crucial role in the progression of mitosis, and as such have several functions including; activation of CDK1, bipolar spindle formation, regulation of anaphase-promoting complex, centrosome maturation, chromosome segregation and execution of cytokinesis (Chopra *et al.*, 2010). PLK1 also is known to directly phosphorylate WEE1, which results in its degradation and therefore entry of cells into mitosis (Chopra *et al.*, 2010). Inhibition of PLK1 prevents bipolar spindle formation, which results in cells arrest and with prolonged arrest, results in apoptosis (Chopra *et al.*, 2010). Several polo kinase inhibitors started clinical trials, such as volasertib and BI2536, but these showed inhibitory off-target effects, including the death-associated protein kinases (DAPKs), which counteracted the cell death induced by PLK inhibition (Raab *et al.*, 2014; Abdelfatah *et al.*, 2019). However, a phase I study with the PLK1 inhibitor, CYC140 is currently recruiting (NCT03884829) and a phase I trial is underway for CFI-400945, a PLK4 inhibitor (NCT01954316) (Moureau *et al.*, 2016; Veitch *et al.*, 2019).

Aurora kinase activity and protein expression are cell cycle regulated, and peak during mitosis in order to carry out critical mitotic processes, which includes; chromosome alignment, chromosome segregation and cytokinesis (Bavetsias and Linardopoulos, 2015a). Further to this, aurora kinase A provides an important link between the DDR and the cell cycle (Figure 1.5). Once HRR machinery is active for the repair of DSBs, the cell cycle is halted by the activation of the CHK1 via ATM or ATR. Cell cycle arrest is achieved through the regulation of CDC25 phosphatases, by WEE1 and PLK1 (Bavetsias and Linardopoulos, 2015a). Importantly, two pathways converge to maintain inhibition of CDK1 preventing cell cycle progression. During recovery from the DNA damage checkpoint, PLK1 becomes dominant and stimulates cell cycle progression (Bavetsias and Linardopoulos, 2015a). During this late phase and during unperturbed cell cycle, Aurora kinase A is an upstream activator of PLK1. Furthermore, Aurora A inhibits RAD51 recruitment to DNA DSBs, decreasing DSB repair by HRR (Bavetsias and Linardopoulos, 2015a).

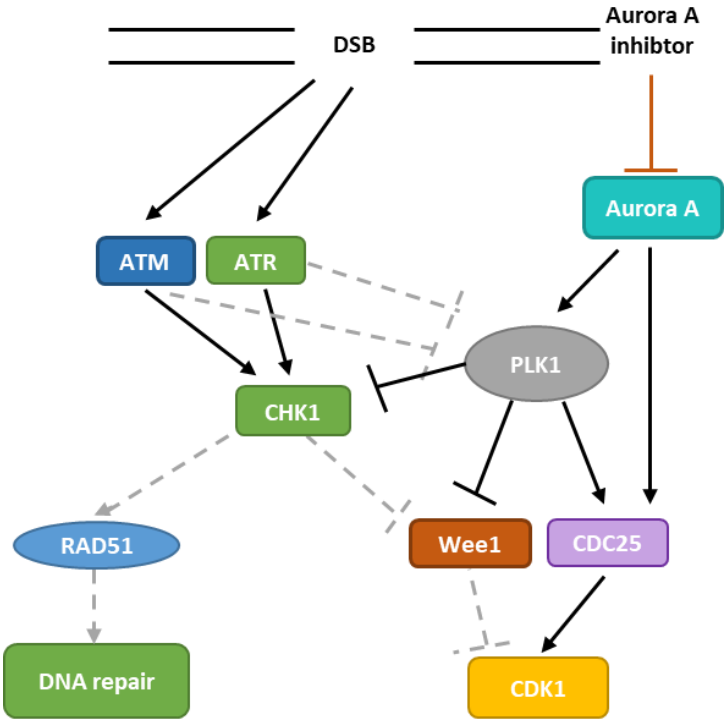


Figure 1.5 Schematic representation of Aurora inhibitor targeting both DDR and cell cycle pathways
Aurora A is directly involved in DNA repair pathways and cell cycle pathways. Inhibition of Aurora A results in increases cellular sensitivity to death. Figure adapted from Bavetsias and Linardopoulos, 2015.

Several aurora kinase inhibitors have been through clinical trials including barasertib, alisertib, tozasertib, danusertib, AT9283, PF-03814735, and AMG 900, which initially aimed to target solid tumours, but were often terminated due to limited efficacy observed in patients (Bavetsias and Linardopoulos, 2015a, 2015b; Borisa and Bhatt, 2017). Very few clinical trials are underway using aurora kinase inhibitors at present, however there is a phase I study of a combination of MLN0128, an mTOR inhibitor, and alisertib in patients with advanced solid tumours and metastatic TNBC recruiting (NCT02719691).

In order for each of these drugs to be used successfully in the clinic, patients need to be identified in subgroups of those that will be likely to get the greatest benefit from treatment with these inhibitors. As with the PARP inhibitor example, patients are selected, which have shown to have either a *BRCA1* or *BRCA2* mutation. In this case, this is considered a predictive biomarker for treatment with PARP inhibitors, and as such, more biomarkers are required to be found for patient stratification for new drug classes.

1.5 Biomarkers of resistance

With the progression of novel drug classes as new therapeutic options, a method is required to determine when drug resistance has occurred, which can be indicative of when a change of therapy is required. One such method is through the identification of biomarkers, a portmanteau of "biological marker", which was initially, broadly defined by the National institute as "a characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic process, or pharmacologic responses to a therapeutic intervention" (Atkinson *et al.*, 2001; Strimbu and Tavel, 2010). The US Food and Drug Administration (FDA) have called for the investigation of fit-for-purpose validated biomarkers in order to evaluate patient response and stratification to cancer therapy and drug resistance (Garnett *et al.*, 2012).

1.5.1 Biomarker discovery

One of the most harnessed methods of identifying biomarkers, and a classic paradigm in diagnostic pathology, is immunohistochemistry (IHC) (Chae *et al.*, 2017). This method can detect specific antigens in histologic specimens through the use of target-specific antibodies with colorimetric or fluorescent reagents, allowing for the identification of proteins or phosphoproteins in the sample (Chae *et al.*, 2017). This method is commonly used to detect the progesterone receptor (PR) and oestrogen receptor (ER) for the diagnosis of breast cancer (Ronchi *et al.*, 2020). Another common method is the use of fluorescence *in situ* hybridization (FISH). FISH uses fluorescently tagged DNA probes which can anneal to target sequence genes and be detected through fluorescence microscopy, allowing for the detection of genetic aberrations in the DNA sample (Chae *et al.*, 2017). FISH is commonly used, alongside IHC, to evaluate the human epidermal growth factor receptor-2 (HER2) status to determine if the receptor is overexpressed and/or amplified, as a diagnostic tool of breast cancer (Chae *et al.*, 2017).

The development of next generation sequencing techniques, such as whole exome sequencing (WES), has allowed for the identification of changes in the genomic landscape of cancer cells. Measuring changes in DNA and RNA can identify mutations, which may be driving cancer, or drug resistance. These mutations can be identified in both pre-clinical models and patient tumour samples and can be used to stratify patients for treatment. One such method using WES in the clinics is to monitor circulating tumour DNA (ctDNA) or circulating tumour cells (CTCs) in the patient's blood (Chen and Zhao, 2019). This liquid biopsy technology can be used to signal both patient relapse and emergence of resistance to the patient's treatment regime (Chen and Zhao,

2019). Although initially, it was hard to detect the ctDNA with enough sensitivity, recent advancements in this technology can detect as low as 90-150 base pair of ctDNA fragments, increasing the biomarker detection in the patient's ctDNA (Mouliere and Rosenfeld, 2015; Mouliere *et al.*, 2018).

Large-scale efforts have led to the development of patient tumour databases, which can be used for the discovery of biomarkers. Large projects have accumulated patient genomic data, which has been collected and stored and made publicly available for analysis. The Cancer Genome Atlas (TCGA) is one such publicly available database, and includes sequenced patient tumour samples, along with gene expression data, and the response of these tumour samples to the treatment of a panel of anti-cancer drugs (Weinstein, 2013). The Genomics of Drug Sensitivity in Cancer (GDSC) database includes cell lines from the Catalogue of Somatic Mutations in Cancer (COSMIC) database, interrogated with a panel of anti-cancer drugs (Yang *et al.*, 2013; Forbes *et al.*, 2017). Also of note is the Broad Institutes Cancer Therapeutics Response Portal (CTRP), which screened 130 clinical and preclinical drugs against a panel of cancer cell lines (Garnett *et al.*, 2012). Comparison of cancer cell lines response to anti-cancer agents, in a known genomic context, can identify biomarkers driving cancer, as well as biomarkers, which could be indicative or driving a drug resistant phenotype.

With the vast amount of public data now available, and the rapid changes in molecular technology, studies identifying biomarkers are occurring at a high frequency (Henry and Hayes, 2012). However, with the overabundance of biomarker information, this is not necessarily translatable, or useful to the clinic. The discovery of biomarkers through next generation sequencing techniques must be experimentally validated and demonstrate clinical relevance (Figure 1.4) (Goossens *et al.*, 2015).

1.5.2 Experimental validation

Biomarkers identified through next generation sequencing techniques in pre-clinical models, need to undergo experimental validation. Investigation of biomarkers associated to drug resistance are often conducted in pre-clinical cancer cell models in which a cancer cell line with acquired drug resistance is interrogated against the paired drug-sensitive cell line. Gene knockdown studies in the drug-sensitive cell line, using siRNA, shRNA or CRISPR methods can identify if the target gene is involved in the resistant phenotype. Further evidence can complement this with the introduction of the functional gene, or overexpression of the target gene into the resistant cell line to observe any re-gain of drug sensitivity.

1.5.3 Biomarker clinical relevance and validity

A biomarkers relevance is indicative of its ability to provide clinically relevant information. Cancer biomarkers can be broadly classified into three categories based on how they are used; diagnostic, prognostic and predictive (Goossens *et al.*, 2015). Diagnostic biomarkers are used to identify whether a patient has a specific cancer or disease. One example is elevated levels of carcinoembryonic antigen in the tissue or the hormone calcitonin in the serum as a diagnosis of a thyroid medullary carcinoma patient (DeLellis, Wolfe and Rule, 1987; Chatterjee and Zetter, 2005). Prognostic biomarkers are not directly linked to or used to make therapeutic decisions, but can inform clinicians with regards to the risk of clinical outcomes, such as cancer recurrence or disease progression. Mutations in the tumour suppressor gene, Phosphatidylinositol 3,4,5 trisphosphate 3 phosphates (PTEN), are routinely used as a biomarker of poor patient prognosis in cancers such as glioblastoma, malignant melanoma, endometrial, prostate, breast, colorectal and pancreatic cancer (Goossens *et al.*, 2015; McCabe, Kennedy and Prise, 2016; Bazzichetto *et al.*, 2019). Predictive biomarkers can be used for patient stratification as they can be indicative as to whether the patient will respond to a given drug. For example, HER2 positive patients have been found to respond to the treatment of trastuzumab in breast cancer, whilst KRAS-activating mutations in colorectal cancer patients are found to be intrinsically resistant to EGFR inhibitors (Slamon *et al.*, 2001; Romond *et al.*, 2005; Van Cutsem *et al.*, 2009; Goossens *et al.*, 2015).

The validity of a biomarker is the measure of its effectiveness or utility. In addition, the use of a biomarker must be reproducible and done so with accuracy (Chatterjee and Zetter, 2005). Often, cancer biomarkers lack specificity and sensitivity. For example, Alpha Fetoprotein (AFP) is considered a gold standard diagnostic biomarker for liver cancer, however it is also known to occur in patients with chronic hepatitis infection (Conti *et al.*, 2015; Luo *et al.*, 2018; Lv *et al.*, 2019). Due to the heterogenous nature of many cancers, some cancer biomarkers that are reported are considered to have low sensitivity in which they are only found in a small subset of patients of a much larger cancer group. These biomarkers are not useful for patient screening however, they could be useful when detecting cancer recurrence within the subset of patients in which this biomarker is detected (Chatterjee and Zetter, 2005). One such example is the biomarker carbohydrate antigen -125 (CA-125), which is found in a subset of ovarian cancers. Of that subset, postsurgical elevation of CA-125 has been found to been indicative of cancer recurrence (Chatterjee and Zetter, 2005; Sharma, 2009). Furthermore, when a biomarker is used to predict a patient's response to a drug, the biomarker must strongly correlate with the resistance or sensitivity accordingly.

The majority of studies focus on biomarkers, which are indicative of a patients response to a drug, but not on biomarkers that indicate when a therapy has stopped working (Michaelis *et al.*, 2020). Identification of predictive biomarkers, which can indicate the emergence of drug resistance and therapy failure, will be particularly useful in cancers that frequently metastasise and develop resistance to the patient's treatment regime. One such cancer where biomarkers indicating early therapy failure would be useful, is the highly aggressive metastatic triple negative breast cancer.

1.6 Triple Negative Breast Cancer

The Global Cancer Observatory (GLOBOCAN), an interactive web-based platform presenting global cancer statistics, estimated the worldwide incidence of new breast cancer in 2018 as 2,088,849, and was considered the highest incident tumour for women (Bray *et al.*, 2018; da Silva *et al.*, 2020). Triple negative breast cancer (TNBC) accounts for approximately 10-15% of all breast cancers, and has been found to frequently affect younger patients (< 50 years) and is more prevalent in African-American women (Reis-Filho and Tutt, 2008; Lund *et al.*, 2009). TNBC tumours are often larger and less differentiated, and are approximately 2.5 fold more likely to metastasize within five years of initial diagnosis, compared to other breast cancer sub-types (Dent *et al.*, 2007; Lehmann *et al.*, 2011; Lee and Djamgoz, 2018).

TNBC is defined by the lack of the oestrogen receptor, progesterone receptor and human epidermal growth receptor 2 (Ryu *et al.*, 2011). Due to the lack of these targets, patients with TNBC do not benefit from hormonal manipulation, which are commonly used in the treatment of other breast cancer subtypes (Wahba and El-Hadaad, 2015). Previously, surgery remained the frontline treatment option for TNBC, with Freedman *et al.*, 2009 showing that the local recurrence rate after surgery is not high in TNBC compared to other breast cancer sub-types (Wahba and El-Hadaad, 2015). Radiotherapy is often applied following surgery, although there is controversy on this issue (Wahba and El-Hadaad, 2015). However, despite adjuvant treatment, the risk of relapse remains high in the first two years (Wahba and El-Hadaad, 2015). More commonly now, clinicians adopt neoadjuvant chemotherapy as the risk of systemic relapse is high, and if pathological complete response (pCR) is achieved it results in improved long term outcomes (Cortazar and Geyer, 2015; Cai *et al.*, 2020).

Clinicians have frequently adopted an intensive chemotherapeutic approach with sequential anthracycline and taxane regimes, which was determined after retrospective analyses of clinical trials reported prior to 2010 (Bergin and Loi, 2019). Eribulin was FDA-approved in 2010, which

showed, through the 'EMBRACE' trial, that there was an improvement of overall survival by 4.7 months in patients with metastatic TNBC, and the use of platinum agents as a neoadjuvant treatment have also shown an increase in pCR (Cortes *et al.*, 2011; Von Minckwitz *et al.*, 2014; Melinda *et al.*, 2016; Lee and Djamgoz, 2018). Initially, TNBC had been shown to be more sensitive to the chemotherapy regime than hormone receptor positive breast cancer, but there is a high risk of recurrence in patients that do not achieve a pCR (Carey *et al.*, 2007). Currently, re-challenging the recurrent or metastatic TNBC with chemotherapy, remains the predominant treatment strategy, but this often results in a poor response, toxicity and multi-drug resistance (Lee and Djamgoz, 2018). New therapy approaches are under investigation, which target specific characteristics that define TNBC, compared to other breast cancer subtypes.

Immunotherapy has made strong advancements in the last few years, having now been approved by the National Institute for Health and Care Excellence (NICE) for the treatment of metastatic TNBC. Immune checkpoint inhibitors (ICIs) have been shown to result in durable clinical remissions in several metastatic cancers (Keenan and Tolaney, 2020). Metastatic TNBC has been identified as a candidate cancer type to respond to immunotherapy through three key characteristics; 1) TNBC has more tumour-infiltrating lymphocytes (TILs) than other breast cancer subtypes which have been found to correlate with good response to ICIs, 2) TNBC has higher levels of programmed death-ligand 1 (PD-L1) expression on both tumour and immune cells which is a direct target for ICIs, and, 3) TNBC has a large number of nonsynonymous mutations, which result in tumour specific neoantigens, which activate neoantigen-specific T cells to mount an antitumour immune response, strengthening the use of ICIs (Keenan and Tolaney, 2020). Of note are the ICIs, pembrolizumab and atezolizumab, which target programmed cell death protein 1 (PD-1) and PD-L1 respectively. Pembrolizumab was investigated in the KEYNOTE-173 trial in combination with chemotherapy in a neoadjuvant setting, whilst atezolizumab was approved after the Impassion130 trial in combination with nab-paclitaxel (Schmid *et al.*, 2018, 2020; Keenan and Tolaney, 2020). However, Impassion131 phase three trial, a continuation of investigation of atezolizumab and nab-paclitaxel combination, showed the combination did not significantly reduce the risk of TNBC progression and death compared to a placebo and paclitaxel in a PD-L1 positive population (*FDA alerts health care professionals and oncology clinical investigators about efficacy and potential safety concerns with atezolizumab in combination with paclitaxel for treatment of breast cancer | FDA*, 9/08/2020).

Inhibitors targeting the PI3K/AKT pathway, a regulator of cell growth and glucose metabolism, are another therapeutic approach under investigation for the treatment of TNBC. Hormone receptor positive (HR+) and HER2+ breast cancers are found commonly to harbour PI3KCA mutations, which

are found less commonly (10%) in TNBC (Cantley and Neel, 1999; Koboldt *et al.*, 2012). However, TNBC has been shown to have pathological activation of the PI3K/AKT pathway through either the loss of PTEN activity, loss of INPP4B or amplification of PI3KCA. Furthermore, TNBC which gain PI3K activation through the loss of PTEN have also shown more growth inhibition from PI3K inhibitors than with cells with *PIK3CA* mutations (Stemke-Hale *et al.*, 2008). The combination of the highly activated PI3K/AKT pathway, and the susceptibility to PI3K inhibition, through loss of PTEN activity, makes treatment with PI3k/AKT inhibitors a candidate therapy for TNBC compared to other breast cancer subtypes (Massihnia *et al.*, 2016; McCann, Hurvitz and McAndrew, 2019). The experimental AKT inhibitor, ipatasertib, underwent the LOTUS phase II clinical trial, and whilst did not improve patient survival compared to the placebo, those with *PIK3CA*/*AKT1*/*PTEN* alterations did show a significant 4.1-month increase in survival (McCann, Hurvitz and McAndrew, 2019). A phase II clinical trial is currently recruiting, which is investigating the combination of ipatasertib in with an ICI, atezolizumab, along with paclitaxel in patients with locally advanced or metastatic TNBC (NCT04177108).

Cyclin-dependent kinase (CDK) inhibitors are also under investigation as a novel treatment of a sub-type of TNBC. CDK 4/6 help tightly regulate the progression of the cell cycle at the G1-S phase transition (Lundberg and Weinberg, 1999; McCann, Hurvitz and McAndrew, 2019). For the cell cycle to continue, the retinoblastoma protein (Rb) is inactivated through hyperphosphorylation by CDK 4/6. CDK 4/6 inhibitors block the hyperphosphorylation of RB, which result in the inhibition of progression from G1/S phase (Lundberg and Weinberg, 1999; Fry *et al.*, 2004). Cell line studies have shown that luminal type, HR+ and HER2- breast cancers are sensitive to growth restriction through inhibition of a CDK4/6 inhibitor; Palbociclib (Finn *et al.*, 2009). Luminal cancers have upregulation of the Rb pathway, with dependence on the CDK4/cyclin D1/Rb interaction (Matutino, Amaro and Verma, 2018). TNBC cell lines of the luminal-androgen (LAR) and mesenchymal-stem like (MSL) subsets were sensitive to CDK4/6 inhibitors, whilst those that demonstrated resistance had elevated cyclin E1 mRNA and protein levels, or are considered as basal-like TNBC (Asghar *et al.*, 2017; Matutino, Amaro and Verma, 2018; Niu, Xu and Sun, 2019). There are currently three approved CDK inhibitors, palbociclib, ribociclib and abemaciclib for the treatment of HR+ and HER2- advanced or metastatic breast cancer (Matutino, Amaro and Verma, 2018) Clinical trials are underway investigating the use of CDK inhibitors in combination with an additional drug class for the treatment of TNBC. Of note, is the PAveMenT trial which is investigating the combination of palbociclib and an ICI, Avelumab, in metastatic AR+ TNBC (NCT04360941).

Finally, as previously mentioned (section 1.4.2), DDR inhibitors are a new drug class, which are also being investigated as a new therapeutic option for the treatment of TNBC, and alongside chemotherapeutic agents, is the drug class of focus in this thesis. TNBC has been shown to have high chromosomal instability, and this has been attributed to defects in DNA repair pathways, specifically HRR (Meyer *et al.*, 2020). Importantly, the percentage of TNBC patients with a *BRCA1* mutant is 35%, whilst other diagnosed breast cancer patients have < 10% of *BRCA1* or *BRCA2* mutants (Peshkin, Alabek and Isaacs, 2010). This led to the OlympiAD study that compared the use of olaparib versus standard single agent chemotherapy in BRCA-mutated breast cancer patients, including TNBC patients (Mehanna *et al.*, 2019; Robson *et al.*, 2019). Often TNBC tumours may not have the germline mutation in *BRCA1* or *BRCA2*, but have other defects in genes involved in the HRR pathway which result in a similar phenotype, and these are considered as a "BRCAness" type (Turner, Tutt and Ashworth, 2004; Peshkin, Alabek and Isaacs, 2010; Mehanna *et al.*, 2019). For example, defects in genes such as *ATM*, *ATR*, *CHK1*, *CHK2*, *RAD51*, *NBS1* and the Fanconi anaemia complementation group (*FANC*) family of genes have all been reported to demonstrate cellular sensitivity to PARP inhibitors (Lord and Ashworth, 2016). *BRCA1* and *BRCA2* mutants are now considered a biomarker for treatment with PARP inhibitors, with olaparib now approved for this context in ovarian and breast cancers (Montemorano, Michelle and Bixel, 2019). However, given the fraction of TNBC patients that harbour the *BRCA1*/*2* mutation, it is evident that more clinical biomarkers for drug response need to be determined to identify new targeted treatments, or treatments as a second line option after the development of chemotherapy resistance.

1.7 Overview and aims of this thesis

TNBC is an aggressive breast cancer sub-type, in which patients often relapse as a result of the development of acquired drug resistance. Given the poor outlook, therapy-refractory TNBC is a disease of unmet medical need. An appropriate second line therapy is required after chemoresistance has occurred. TNBC has high chromosomal instability, which is attributed to defects in the DNA repair mechanisms, specifically in HRR, with many TNBC patients having mutant *BRCA1*/*2* or demonstrating a "BRCAness" phenotype. Inhibitors of the DDR are a new drug class now entering the clinic, and investigation of their use both as a monotherapy or in combination with another drug type is underway. However, their use after therapy failure due to acquired resistance to chemotherapeutic agents are yet to be determined. Analysis of DDR inhibitors as a second line therapy option after chemotherapeutic resistance, may provide a beneficial therapeutic option. Furthermore, early identification of the emergence of chemoresistance in TNBC is crucial in identifying when a change of therapy is appropriate. Understanding mechanisms of acquired

resistance in TNBC, and identifying biomarkers of resistance that will indicate when a change in therapy is required could potentially improve the outcomes of TNBC patients that develop chemoresistance.

Previous studies have shown that pre-clinical cell lines models can be used to identify cell response to novel drugs, as well as determine candidate biomarkers of resistance. To this end, this project uses chemo-naive MDA-MB-468, HCC38 and HCC1806 TNBC cell lines, and chemo-resistant TNBC sub-lines acquired from the Resistant Cancer Cell Line (RCCL) collection; a set of more than 1300 cancer cell lines that can be used to model acquired resistance in cancer (Michaelis, Wass and Cinatl, 2019). These cell lines will be used to consider inhibitors of the DDR as an appropriate next line therapy option after chemo-resistance has occurred, and the exome sequencing data of these cell lines to identify candidate biomarkers as an indication of chemotherapy failure.

Aims and objectives:

- Characterise chemo-naive and chemo-resistant TNBC cell lines by:
    - Examination of morphological differences between chemo-naive and chemo-resistant cell lines
    - Cross-resistance profiling to other chemotherapeutic agents
- Determine viable next-line therapy options by:
    - Cross-resistance profiling the chemo-naive and chemo-resistant TNBC cell lines against DDR inhibitors.
- Identification of clinically relevant candidate biomarkers or mechanisms of resistance by:
    - Examination of exome variants in chemo-naive and chemo-resistant TNBC cell lines
    - Comparison of identified candidate variants with clinical patient data

Chapter 2

Materials and Methods

2.0 Materials and Methods

2.1 Cell Biology

2.1.1 Cell lines and cell line nomenclature

Fifteen TNBC cell lines were obtained from the Resistant Cancer Cell Line (RCCL) collection and were comprised of chemo-naive and chemo-resistant cell lines (Table 2.1; https://research.kent.ac.uk/industrial-biotechnology-centre/the-resistant-cancer-cell-line-rccl-collection/; (Michaelis, Wass and Cinatl, 2019). For the purpose of this project, the MDA-MB-468, HCC38 and HCC1806 cell lines are assumed to be chemo-naive as historical information (previous treatment, year and country sample was taken) was not available. The chemo-resistant cell line nomenclature is as follows: chemo-naive cell line its derived from, r (to indicate resistance), abbreviation of drug name, concentration of drug cell line is maintained in (ng/mL). Drug abbreviations are as follows; Cisplatin = CDDP, Doxorubicin = DOX, Eribulin = ERI, Gemcitabine = GEM, Paclitaxel = PCL and 5-Fluorouracil = 5-F. For example, HCC38^{r}CDDP^{3000} indicates a cisplatin resistant sub-line derived from HCC38 and maintained in 3000 ng/mL of cisplatin.

Table 2.1 Panel of chemo-naive and chemo-resistant cell lines

	Cell lines			
	Chemo-naive	MDA-MB-468	HCC38	HCC1806
Chemo-resistant	**Cisplatin**	MDA-MB-468^{r}CDDP^{1000}	HCC38^{r}CDDP^{3000}	HCC1806^{r}CDDP^{500}
	Doxorubicin	MDA-MB-468^{r}DOX^{50}	HCC38^{r}DOX^{40}	HCC1806^{r}DOX$^{12.5}$
	Eribulin	MDA-MB-468^{r}ERI^{50}	HCC38^{r}ERI^{10}	HCC1806^{r}ERI^{50}
	Gemcitabine	-	HCC38^{r}GEM^{20}	HCC1806^{r}GEM^{20}
	Paclitaxel	MDA-MB-468^{r}PCL^{20}	HCC38^{r}PCL$^{2.5}$	HCC1806^{r}PCL^{20}
	5-Fluorouracil	-	-	HCC1806^{r}5-F^{1500}

2.1.2 Maintenance of cell lines

All cell lines were maintained in Iscove's Modified Dulbecco's (IMDM; Fisher Scientific, UK) containing L-Glutamine and 25 mM HEPES, supplemented with Fetal Bovine Serum (FBS; Sigma-Aldrich, Germany) and 1% Penicillin-Streptomycin Solution Liquid (Life Technologies, UK; complete IMDM), at 37 °C in a humidified 5% $C0_2$ incubator. The cells were passaged when the flask was approximately 70-80% confluent. Cells in a T25 flask were washed with 2 mL phosphate-buffered saline (PBS) preceding detachment with 1 mL Trypsin-EDTA x10 solution (Sigma-Aldrich, Germany). The detached cells were resuspended in complete IMDM and split at an appropriate concentration

into a new flask. Chemo-resistant cell lines were maintained in the required chemotherapeutic agent at concentrations outlined in the cell nomenclature in Table 2.1. Chemotherapeutic agents used are outlined in Table 2.2. Prior to plating, a cell sample was treated with trypan blue and counted on a haemocytometer.

In order to prevent genetic deviation, cell populations were passaged continuously for no longer than 5 months. Fresh stocks were made from the earliest passage possible from a subpopulation of each cell line. Cells were grown to approximately 80% in a T75 flask, trypsinised and resuspended in complete IMDM before being spun down at 270 x g for five minutes at room temperature. Culture medium was aspirated and the cell pellet was resuspended in 3 mL of freeze-down media (10% DMSO, 40% FBS, 50% complete IMDM) before being aliquoted into three cryovials and cooled in a polystyrene sectioned box in the -80 ˚C overnight. Frozen cell stocks were transferred to a cryostat to increase longevity. For revival, freeze-downs were thawed as quickly as possible by warming in a 37 ˚C water bath. Once thawed, the cells in the freeze down media were transferred to 9 mL of complete IMDM in a T25 flask and incubated as per normal maintenance. Chemo-resistant cell lines were re-introduced to the chemotherapeutic agent after two successful passages, to maintain resistance.

All cell lines have been tested to ensure they were free from *Mycoplasma* contamination using the VenorGeM® Mycoplasma PCR detection kit (Minerva Labs, UK).

2.1.3 Cell seeding density optimisation

Cells were seeded with five technical replicates in seven 96-well plates in 200 µL complete IMDM at cell densities specified. One plate was fixed every 24 hours with 70 µL 10% (w/v) trichloroacetic acid (TCA), stained and analysed as described for SRB growth assay (see section 2.1.4). Raw absorbances were used to generate growth curves in GraphPad Prism 6 (GraphPad Software Inc, USA). The doubling time for each cell line was calculated using the following equation:

$$Doubling\ time = \frac{Duration\ (hours) x \log(2)}{\log(final\ OD) - \log(inital\ OD)}$$

2.1.4 Sulforhodamine growth assay

Sulforhodamine B dye (SRB; Sigma-Aldrich, UK) is a useful tool for measuring proliferation of cells. It has been shown that the dye binds to amino acids which provides an estimation of cellular density (Skehan *et al.*, 1990). Cells were plated out per well as specified in 200 µL complete IMDM and incubated for the stated time. Cells were fixed with 70 µL 10% (w/v) TCA for 30 minutes and washed five times with water. The fixed cells were stained with 0.4% (w/v) SRB solubilised in 1% (v/v) acetic acid (Fisher Scientific, UK) for 30 minutes, and washed five times with 1% (v/v) acetic acid before drying overnight at 37 °C. Bound SRB was solubilised in 100 µL 10 mM Tris base (Sigma-Aldrich, UK). Absorbance was read 490 nm wavelength in a Victor X4 Multilabel Plate reader (PerkinElmer Life Sciences, USA). The half-maximal growth inhibitory concentration (GI_{50}) of the specified drug was determined using GraphPad Prism 6 (GraphPad Software Inc, USA).

2.1.5 MTT assay

3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT; Universal Biologicals UK) is a yellow tetrazole that is metabolically reduced to purple formazan only in living cells. This can be used as a quantitative colorimetric assay which allows for the measurement of cytotoxicity and proliferation (Mosmann, 1983). Cells were plated out per well as specified in 50 µL complete IMDM. When screening for sensitivity to drugs, the desired agents were serially diluted and added to cells in 50 µL complete IMDM per well with the concentrations stated (Table 2.2 and Table 2.3 for list of drugs). After incubation of 120 hours, 25 µL of MTT reagent was added and cells were incubated for 4 hours. The cells were lysed with addition of 100 µL 20% sodium dodecyl sulfate (SDS, Fisher Scientific) and incubated overnight. Absorbance was read at 600 nm wavelength in a Victor X4 Multilabel Plate reader (PerkinElmerLife Sciences, USA). The half-maximal growth inhibitory concentration (GI_{50}) of the specified drug was determined using GraphPad Prism 6 (GraphPad Software Inc, USA). Structures of each of the chemotherapeutic agents and DNA damage response and repair targeted inhibitors are found in Appendix A1.

Table 2.2 List of chemotherapeutic agents

Drug	Supplier	Catalogue Number
Cisplatin	Sigma-Aldrich	P4394
Doxorubicin	Selleckchem	S1208
Eribulin	Eisai	N/A
Gemcitabine	Selleckchem	S1149
Paclitaxel	Cayman Chemicals	10461
5-Fluorouracil	Sigma-Aldrich	F6627

Table 2.3 List of DNA damage response and repair targeted inhibitors

Drug	Alternative names	Targets	Supplier	Catalogue Number
Adavosertib	MK-1775, AZD1775	Wee1	Adooq Bioscience	A10599
Alisertib	MLN8237	Aurora A	Adooq Bioscience	A10004
AZD0156	-	ATM/ATR	Selleckchem	S8375
B02	-	RAD51	Sigma-Aldrich	SML0364
Berzosertib	VE-822, VX970, M6620	ATR	Adooq Bioscience	A13289
BI2536	-	Polo kinase 1/2/3	Selleckchem	S1109
CCT241533	-	CHK2	Institute of Cancer Research, London	N/A
Ceralasertib	AZD6738	ATR	Adooq Bioscience	A15794
MK-8776	-	CHK1	Adooq Bioscience	A11167
Olaparib	-	PARP	Adooq Bioscience	A10111
Prexasertib	LY2606368	CHK1/2	Adooq Bioscience	A13684
Rabusertib	LY2603618	CHK1	Adooq Bioscience	A11036
Rucaparib	-	PARP	Adooq Bioscience	A10045
SBE13	SBE 13 HCL	PLK1	Adooq Bioscience	A14426
SRA737	CCT245737	CHK1	Institute of Cancer Research, London	N/A
Tozasertib	VX-680, MK-0457	Pan-Aurora kinase	Adooq Bioscience	A10981

2.2 Biochemistry

2.2.1 Cell lysis

Cells were plated out into 10cm plates at specified densities and grown for 48 hours, reaching 70% confluency. The culture medium was removed prior to the cells being washed twice with ice cold PBS. 100 µL of lysis buffer (50 mM HEPES pH7.4, 250 mM NaCl, 0.1% NP40, 1 mM DTT, 1 mM EDTA, 1 mM NAF, 10 mM β-Glycerophosphate, 0.1 mM sodium orthovanadate and Complete™ protease inhibitor cocktail (Roche, Switzerland)) was added and the cells were scraped into cold microcentrifuge tubes and incubated on ice for 30 minutes before centrifugation at 13,000 rpm for 10 minutes at 4 °C to remove insoluble material. The lysate was then transferred to a clean microcentrifugation tube and kept on ice (if to be used immediately) or frozen on dry ice and stored at -80 °C.

2.2.2 Determination of protein concentration

Bicinchoninic acid (BCA) is a sodium salt which is known to form an intense purple complex with cuprous ion in an alkaline environment. Proteins can reduce CU^{+2} to CU^{+1}, resulting in a purple colour formation by BCA. This allows for a colorimetric assay for the determination of the protein

concentration from cell lysates (Smith *et al.*, 1985). Cell lysate (as obtained in section 2.4.1) was diluted 20-fold in ddH_2O and 10 µL was added in duplicate to a 96-well plate. 10 µL of diluted (0.1-1 µg mL^{-1}) bovine serum albumin (BSA) protein standards (Sigma-Aldrich, USA) and ddH_2O blanks were included on every plate. At a 1:50 dilution Copper (II) sulfate solution (Sigma-Aldrich, USA) was mixed with BCA, and 200 µL was added to each well. Samples were mixed on a plate shaker prior to incubation at 37 °C for 30 minutes. Absorbances were read at 570 nm wavelength in a Victor X4 Multilabel Plate reader (PerkinElmer Life Sciences, USA). A standard curve was produced from the protein standards of known concentration, and sample protein concentration was determined.

2.2.3 SDS-PAGE

In order to determine, in a semi-quantitative manner, the amount of specific proteins within the cell lysates, the samples were separated by molecular weight under denaturing conditions using sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE). Equal amounts of protein sample (15-45 µg) were diluted into lysis buffer and 3x loading buffer (187.5 mM Tris-base pH 6.8, 6% (w/v) SDS, 30% glycerol, 15% (v/v) β-mercaptoethanol, 0.15% (w/v) bromophenol blue), typically at a 1:2 ratio of sample buffer to lysate. Samples were heated at 95 °C for 5 minutes to allow for denaturation and reduction of the proteins, and centrifuged at 14,000 xg for 1 minute. Samples were loaded onto fixed concentrations gels, as stated in text, with a concentration dependent on the protein molecular weight. Precision Plus Protein Standards (BIO-RAD, USA) were loaded in a separate lane to allow for an estimation of the protein size. Samples underwent electrophoresis through the gel in Tris-glycine running buffer at 150 V for 60-90 minutes.

2.2.4 Western blotting

The proteins separated by SDS-PAGE (section 2.2.3) were transferred to methanol-activated 0.2 µm pore Immobilon-P PVDF membrane (Millipore, USA) at 100 V for 90 minutes in transfer buffer (25 mM Tis-Base, 190 mM glycine, 10% methanol) using the Bio-Rad Mini Transfer Trans-Blot Transfer Cell apparatus (Bio-Rad, USA). After transfer, membranes were re-activated in methanol and incubated in ponceau S solution (0.1% ponceau S in 5% acetic acid) for five minutes and rinsed in ddH_20 to determine the quality of transfer. Membranes were appropriately sliced to separate proteins of interest for probing in primary antibodies. Subsequently the membranes were blocked in Tris Buffered Saline Tween Buffer (TBST; 50 mM Tris pH8.0, 150 nM NaCl, 0.1% Tween-20 (v/v)) containing 5% milk. The membranes were then incubated in primary antibody diluted into 5% milk/TBST at 4 °C overnight (Table 2.4 for list of antibodies). Membranes were washed with TBST

twice for ten minutes before incubation in secondary horseradish peroxidase-conjugated goat, anti-rabbit or anti-mouse antibody (Bio-Rad, USA; primary antibody dependent) at room temperature for one hour. Membranes were washed four times for five minutes in TBST before detection was performed with Enhanced Chemiluminescence (ECL) Western Blotting Substrate (Pierce Biotechnology, USA). Bands were visualised by exposure to Amersham Hyperfilm ECL (GE Healthcare, USA). The film was then scanned on an Epson Expression 1600 (Seiko Epson Inc., Japan) and labelled in Microsoft Powerpoint.

Table 2.4 List of antibodies used in western blotting analysis

Primary Antibody	Species	Dilution	Supplier	Catalogue Number
MDR1	Mouse	1:1000	Santa Cruz	sc-55510
$p21^{CIP1/WAF1}$	Rabbit	1:500	Cell Signalling Technology	#29475
Cyclin E1	Mouse	1:2000	Santa Cruz	sc-481
TOP2A	Rabbit	1:1000	Cell Signalling Technology	12286S
GAPDH	Mouse	1:1000	Merck Millipore	MAB374
β- tubulin	Rabbit	1:2000	Cell signalling Technology	#2146
β- actin	Mouse	1:4000	Santa Crus, USA	Sc-47778
Secondary Antibody	**Species**	**Dilution**	**Supplier**	**Catalogue Number**
Anti-mouse HRP conjugate	Goat	1:10,000	Bio-Rad, USA	170-6516
Anti-rabbit HRP conjugate	Goat	1:10,000	Bio-Rad, USA	170-6515

2.3 Molecular Biology

2.3.1 Reverse transcription quantitative polymerase chain reaction (RT-qPCR)

Cells were plated out in 10 cm plates at specified seeding densities and grown for 48 hours, reaching 70% confluency. The culture medium was removed prior to the cells being washed twice with ice cold PBS. RNA was extracted using the QIAGEN RNeasy Mini kit (QIAGEN, Germany) following the manufacturer's instructions. The concentration of RNA was determined using a NanoDrop 2000C Spectrophotometer (ThermoFisher, USA).

In order to reverse transcribe RNA into complementary DNA (cDNA), GoScript reverse transcriptase (PROMEGA, USA) was used. Purified RNA was mixed with Oligo(dT)ls primer (PROMEGA) and incubated at 70 °C for five minutes. The reaction mixture was prepared as follows; 4 µL GoScript 5x reaction buffer (PROMEGA), 4 µL $MgCl_2$ (PROMEGA), 1 µL PCR Nucleotide Mix (PROMEGA), 2 µL

RNasin Ribonuclease inhibitor recombinant (PROMEGA), 1 µL GoScript reverse transcriptase (PROMEGA), and 3 µL nuclease free water. The samples were incubated on the following programme, 25 °C for five minutes, 42 °C for one hour, 70 °C for 15 minutes and cool to 4 °C until ready to use. The cDNA was then diluted in RNAase free water in a ratio of 1 part cDNA, 14 parts RNAase free water.

Polymerase chain reaction (PCR) was performed using the cDNA as a template and Sybr Green Real-Time PCR Master Mixes® (ThermoFisher USA). A 10 µL reaction was set up containing 4.4 µL of diluted cDNA, 5 µL Sybr Green Real-Time PCR Master Mixes®, 0.3 µL forward and 0.3 µL reverse primer. Primer sequences are shown in Table 2.5. The PCR cycling conditions were 95 °C for 10 minutes followed by 40 cycles of denaturation at 95 ˚C for 10 seconds and annealing at 58 °C for 15 seconds, and elongation at 72 °C for 25 seconds. A melting curve was then produced by incubating in the following cycle; 95 °C for 15 seconds, 60 °C for one minute and 95 °C for one second. Data was analysed in QuantStudio™ Design & Analysis Software v1.4.3 (ThermoFisher, USA).

Table 2.5 PCR and sequencing primers

Forward primers	Sequence (5'-3')
CDKN1A_F1	TGG AGA CTC TCA GGG TCG AAA
CDKN1A_F2	CAT GTG GAC CTG TCA CTG TCT TGT A
GAPDH	GTC ATC CAT GAC AAC TTT GGT A
Reverse primers	**Sequence (5'-3')**
CDKN1A_R1	GGC GTT TGG AGT GGT AGA AAT C
CDKN1A_R2	GAA GAT CAG CCG GCG TTT G
GAPDH	GGA TGA TGT TCT GGA GAG C

2.3.2 Lipid mediated reverse transfection siRNA knockdown

Transient knockdown of target gene expression was achieved using lipid-mediated reverse transfection of small interfering RNA (siRNA) oligonucleotides. Lipofectamine (Life Technologies, USA) transfection reagent and siRNA oligonucleotides (QIAGEN, USA) were used at final concentrations as indicated in the text. Appropriate death and non-targeting control oligonucleotides were included to establish transfection efficiency and toxicity. The oligonucleotide sequences used to deplete gene expression are listed in Table 2.6.

Oligonucleotide and transfection reagent were complexed in OptiMEM (Life Technologies, USA) by incubation at room temperature for 15 minutes. Meanwhile, cells were prepared by trypsinisation

and cell counting. For 96-well plates, 50 µL of oligonucleotide/lipid complex was added to each well before addition of cells at numbers stated in 50 µL complete IMDM. For analysis of protein knockdown levels by western blot, cells were transfected in 6-well plate format with 1.5 mL oligonucleotide/lipid mixture and an equal volume of cells at a count stated in the results. After incubation for 48 hours, cells were harvested and analysed by western blot as described in section 2.2.4.

Table 2.6 siRNA target sequences

siRNA	Target Gene	Target Sequence (5'-3')
AllStars Negative Control	Non-targeting	Sequence validated by QIAGEN
AllStars Hs Cell Death Control	Death Control	Sequence validated by QIAGEN
CDKN1A_7	CDKN1A	CTGGCATTAGAATTATTTAAA
CDKN1A_6	CDKN1A	CAGTTTGTGTGTCTTAATTAT
CDKN1A_5	CDKN1A	AAGACCATGTGGACCTGTCAC
CDKN1A_9	CDKN1A	AAGAAGGGCACCCTAGTTCTA
TOP2A_1	TOP2A	GCTTTCCTTACCAATTGAT
TOP2A_2	TOP2A	CTACTTGAGACGTCCGATT
TOP2A_3	TOP2A	CCTCTTCTAATATGTACAT
TOP2A_4	TOP2A	CCATTGAGGAACTTTCATT

2.3.3 Whole exome sequencing

Whole exome sequencing (WES) was performed on the TNBC cell lines by the Genomic Core Facility, Philipps-University, Marburg, Germany, using a transposase-based method utilising the Illumina "Nextera Exome Enrichment Kit". 50 ng of fragmented and adapter tagged DNA was amplified via PCR protocol and sequencing indexes were added. The indexed libraries were pooled, denatured to ssDNA before hybridisation to biotin-labelled custom oligonucleotide capture probes, specific to targeted regions. Addition of Streptavidin beads, which binds to biotinylated probes, allowed the bound DNA fragments to be magnetically pulled down and eluted from solution before amplification by PCR. 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.

2.4 Bioinformatics

2.4.1 Next-generation sequencing variant alignment and annotation pipeline

2.4.1.1 Quality control of FASTQ files

FASTQC was used to perform quality control checks on the raw sequence data to identify low-quality reads and contaminants including duplicates, adapters and PCR primers (Andrews S, 2018). Modular set of analyses allowed for identification of problems with per base sequence quality, per sequence quality scores, per base sequence content, per base/sequence GC content, per base N content, sequence length distribution, sequence duplication levels, overrepresented sequences and Kmer content.

2.4.1.2 Trimming

Trimmomatic was used for the removal of sequencing adaptors from the raw sequence data. Parameters for this tool were used as follows: NexteraPE-PE.fa:2:30:10 LEADING:3 TRAILING:3 SLIDING WINDOW: 4:15 MILEN:36 (Bolger, Lohse and Usadel, 2014).

2.4.1.3 Mapping raw FASTQ reads to reference genome

The Burrows-Wheeler Alignment (v.0.7.17) was used to align the raw FASTQ files to Genome Reference Consortium human reference genome (GRCh37) outputting a Sequence Alignment Map (SAM) format, which includes the read-group information (Burrows and Wheeler, 1994; Li *et al.*, 2009; Church *et al.*, 2011). Parameters were used at the tools default settings; -M -R. Here only paired reads were used, and Samtools flagstat used to print statistics throughout each of the subsequent steps.

2.4.1.4 Sorting SAM file and conversion to BAM file

In order to achieve a fast retrieval of alignments in overlapping specified chromosomal regions, the SAM files were inputted 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).

2.4.1.5 Marking duplicate PCR reads

In order to mitigate potential biases on variant calling algorithms, Picard Tools MarkDuplicates (v2.17.10) was used to tag PCR duplicates, which were subsequently removed. (Picard Toolkit. 2019. Broad Institute, GitHub Repository. http://broadinstitute.github.io/picard/; Broad Institute). The

output consists of two files; a BAM file containing SAM flags for each of the reads, and another identifying the number of duplicates for the paired end reads.

2.4.1.6 Merging BAM files

As the sequencing was performed in two lanes for each cell line, it was required to merge the two BAM files together. The two BAM files were then run together through Picard Tools MarkDuplicates, in order to remove any PCR duplicates after merging the files with the output as one merged BAM file.

2.4.1.7 Building BAM index

To ensure a faster search of data through the BAM file, an index of the BAM file was created and sorted in coordinate order Picard Tools BuildBamIndex (v2.17.10) to output an indexed BAM file (BAMi).

2.4.1.8 Sequence realignment

Single nucleotide variants (SNVs) may mistakenly be identified in the individual genome with respect to the reference genome due to poor local realignment. Insertion or deletion of bases (INDELs) increase the number of mismatching reads highlighting the requirement for these regions to be realigned. Both the BAM file and the GRCh37 reference genome were input into GenomeAnalysisTK-3.7.0 (GATK) with RealingerTargetCreator to create realignment targets. GATK IndelRealinger was then used to execute the realignment of the listed targets (McKenna *et al.*, 2010).

2.4.1.9 Build recalibration model

The sequencer which estimates the quality score of each base call can have systematic errors. GenomeAnalysisTK-3.7.0 was used to perform base score recalibration. It analyses patterns of covariation in the sequence datasets, and applies the recalibration to the sequence data (McKenna *et al.*, 2010).

2.4.1.10 Variant calling

SAMtools mpileup was used to generate Binary Variant Call Format (BCF) files from the BAM files in order to compute the genotype likelihoods for each read using the default parameters (Li, 2011). These were then input into BCFtools to call the SNVs and INDELS to generate a Variant Calling Format (VCF) file containing information about the variants position in the genome.

2.4.1.11 Variant effect prediction

The Ensembl Variant Effect Predictor (VEP) was used to annotate the called variants to determine which genes, and region of the gene, the variant is located in (i.e. coding, regulatory, non-coding regions etc), the prediction of consequence of a called variant (i.e. frameshift, stop-gain, missense etc) and if the variant is considered synonymous or non-synonymous. VEP predicts protein function of the called variants by labelling with pathogenicity scores by programmes; SIFT, Polyphen and ClinVar (Ng and Henikoff, 2003; Landrum *et al.*, 2014; Adzhubei, Jordan and Sunyaev, 2015; McLaren *et al.*, 2016).

2.4.2 Variant filtering

Variants called in the individual genome, were filtered to identify high confidence somatic variants. Called variants were checked for quality and coverage. Variants were removed if the Phred quality score is < 30 or the variants have less than 10 reads supporting the base call. If < 3 reads at the base call did not support the variant, these were subsequently removed from the dataset. In order to remove common germ-like variants, variants found at a frequency of ≥ 0.001% in the genome aggregation database (gnomAD) were removed (Karczewski *et al.*, 2019). However, if variants were found to be in ≥ 3 samples in The Cancer Genome Atlas (TCGA), or ≥ 10 samples in Catalogue Of Somatic Mutations In Cancer (COSMIC), these were re-added to the called variants list (Bamford *et al.*, 2004; Weinstein, 2013; Ghandi *et al.*, 2019). Variants were also removed if they were not found in the protein sequence, such as upstream variants, as these are considered outside of the confident sequencing scope. This left only high confidence somatic variants in the VCF file.

2.4.3 Computational analysis

2.4.3.1 Variant comparative studies

Comparative studies of the called variants in VCF files were conducted through the use of scripts made in the Python language and graphs and plots were made using GraphPad Prism 6 (GraphPad Software Inc, USA). Density plots were made using the online tool DensityPlotter (Spencer, Yakymchuk and Ghaznavi, 2017). Mutational patterns, kataegis and mutational signatures were analysed using the online tool Mutalisk (Lee et al., 2018).

2.4.3.2 Gene Ontology

Two online tools were used to analyse gene ontology (GO). GO functional enrichment analysis was conducted using G:profiler (Raudvere *et al.*, 2019). This tool maps genes to known functional

information sources, such as Ensembl, and identifies statistically significant enriched terms (McLaren *et al.*, 2016). KEGG BRITE pathway was used to label the gene lists with biological and cellular functions. KEGG BRITE is the reference database for BRITE mapping in KEGG Mapper and is a collection of hierarchical classification systems which incorporates different types of relations including; genes and proteins, compounds and reactions, drugs, diseases and organisms and cells (Kanehisa, 2000; Kanehisa *et al.*, 2019).

2.4.3.3 Variant effect predictions

Three variant effect predictors were used to identify if a missense variant in a gene is to be considered damaging to the structure or function of the protein it encodes. SIFT (Sorting intolerant from tolerant), is a variant predictor which assess whether an amino acid substitution affects protein structure based on sequence homology and the physical properties of amino acids (Ng and Henikoff, 2003). PolyPhen (Polymorphism Phenotyping) is a tool which predicts possible impact of an amino acid substitution on the structure and function of a protein using both physical and comparative considerations. These include analysis of protein secondary structure, including surface area, and Phi-psi dihedral angles, sequence alignment and phylogenetic and structural information to characterise the variant (Adzhubei, Jordan and Sunyaev, 2015). Mutational assessor predicts the functional impact of amino acid substitutions in proteins through assessments based on evolutionary conservation in protein homologs, and identifies if the variant is in binding domains required for protein-protein interaction, DNA/RNA or small molecule binding (Reva, Antipin and Sander, 2011).

2.4.3.4 Extraction of data from Sanger Genomics of Drug Sensitivity in Cancer (GDSC)

Pan-cancer data was extracted from the GDSC which had *CDKN1A* expression data available, and a response to treatment of the following drugs (measured as area under the curve); cisplatin, AZD7762 (CHK1 inhibitor), 681640 (WEE1/CHK1) inhibitor, QL-VIII-58 (ATR/MTOR inhibitor), and KU-55933, CP466722 (ATM inhibitors) (Yang *et al.*, 2013). *CDKN1A* expression data was divided into high or low expression based on the mean expression. Box plots were created of the distribution of response to the drugs (measured by area under the curve), vs high and low expression using GraphPad Prism 6 (GraphPad Software Inc, USA).

2.4.3.5 TCGA analysis

Variant data was extracted via the GDSC Data portal and the Bioconductor R package *TCGAbiolinks* was used to obtain clinical data (Colaprico *et al.*, 2016; Grossman *et al.*, 2016). Chromosomal locations of patient variants were remapped from GRCh38 to GRCh37 using the NCBI Genome Remapping service. Pan-cancer gene expression and survival data was extracted for each chemotherapeutic agent. Survival analyses were conducted to determine the response of the patient treated with the chemotherapeutic agent for when the gene expression was high or low. Cox proportional hazards regression was used to calculate the hazard ratio for cohorts expressing high vs low expression levels of the given gene. The '*surv*_cutpoint' function of the package *survminer* in R allowed for the identification of the optimal expression cut-ff point to give the lowest p-value for high vs low expression. The cut-off selected was between the 20^{th} and 80^{th} percentiles of gene expression values as previously described by Uhlen *et al.*, 2017. The calculations used overall survival as the measure of clinical outcome. Overall survival is defined as days to last medical follow up or death as was previously described by Ng *et al.*, 2016. The calculations were performed using the R *survminer* and *survival* packages. From this Kaplan-Meier survival curves were generated using the R package *ggsurvplot*. Statistical analysis using the Wald test (or log rank (Mantel-Cox) test was performed to obtain p-value of significance for each Kaplan-Meier graph. Hazard ratios were also calculated which refer to values for "low" (below median) expression for each given gene, with values >1 indicative of increased hazard (a reduced overall survival) and values < 1 are indicative of decreasing hazard (an increased overall-survival).

Chapter 3

Characterisation of chemo-resistant Triple Negative Breast Cancer cell lines

3. Characterisation of chemo-resistant Triple Negative Breast Cancer cell lines

3.1 Introduction

Resistance to chemotherapy agents is a bottleneck in the treatment of cancer leading to poor drug responses and patient survival. Often in breast cancer, this is due to genomic instability which facilitates the tumour resistance to cytotoxic therapies (Kalimutho *et al.*, 2019). Understanding the underlying mechanisms of both intrinsic and acquired resistance experimentally, can aid in the development of therapeutic strategies to overcome drug-resistance in the clinic. Methods of investigating drug resistance in patient samples are becoming increasingly popular, especially with the development of liquid biopsy technology which has the potential to both confirm mechanisms of drug-resistance or identify drug-resistance emerging in a patient. By detecting circulating tumour DNA (ctDNA) or circulating tumour cells (CTCs), samples can be measured for changes in heterogeneity and can be used to identify emerging biomarkers of drug-resistance (Rolfo *et al.*, 2014). Tang *et al.*, 2016 showed that ctDNA could be used to monitor clonal evolution during routine management of non-small-cell lung cancer (NSCLC), and identified emerging resistance to tyrosine kinase inhibitors through observed epidermal growth factor receptor (EGFR) T790M mutations. However, although given the wide scope of research liquid biopsy can provide, it is not without technological difficulties. For example, the concentration of ctDNA and CTCs depends on the localisation of the tumour tissuse, with primary or metastatic brain lesions difficult to assess via blood analyses (Heidrich *et al.*, 2020). Due to this, surgical biopsies, to obtain patient tumour samples, continue to dominate the clinic. These can facilitate research on samples pre and post-drug treatment, or between primary and metastatic sites. However, the patient sample material is often limited and retrospective, and the samples are commonly fixed in formalin and embedded in paraffin which can reduce their usefulness in molecular analysis (Herwig *et al.*, 2011).

The use of cell line models is still a widely accepted method underpinning pre-clinical drug-resistance research. Cell line models can allow for an understanding of mechanisms driving drug-resistance before its emergence in the clinic, and to identify ways to overcome this resistance phenotype. A number of methods and models have been successfully developed which have determined important resistance mediators (Garraway and Jänne, 2012a).

A common way of addressing drug-resistance is through large-scale screening of cancer cell lines with known genetic background, to identify biomarkers of intrinsic sensitivity or resistance to panels of drugs. One study, which used part of the Cancer Cell Line Encyclopaedia (CCLE), was conducted by Barretina *et al.*, 2012 where the genetic background of 479 cell lines was coupled

with pharmacological profiles to 24 anticancer drugs. The study confirmed previously established activating mutations in *BRAF* and *NRAS* as predictors of sensitivity to the treatment of the MEK inhibitor; PD-0325901. Further to this, the study found elevated *AHR* gene expression was identified as an additional predictor of sensitivity in cell lines with the *NRAS* mutation.

Another large-scale screen was conducted by Garnett *et al.*, 2012 describing the response of 600 cancer cell lines, with a known genetic background, to a panel of 130 inhibitors, which revealed markers of resistance that correlated with genetic mutations, or cancer type. One example identified, was Ewing's sarcoma cells harbouring the EWS-FLI1 gene translocation, which showed sensitivity to PARP inhibitors. A further study went on to combine the data from the studies by Barretina and Garnett, conducted by Nichols *et al.*, 2014, which showed, through a head and neck squamous cell cancer (HNSCC) specific study, that HNSCC cell lines harbouring *PIK3CA* mutations conferred sensitive to treatment with PI3K inhibitor; AZD6482. This supported the hypothesis that *PIK3CA* can be used as a biomarker for treatment by PI3K inhibitors.

Large scale siRNA or shRNA screens have been useful to determine if knockdown of target genes identify genes regulating sensitivity to drugs. An example is seen when Campbell *et al.*, 2016 used a series of siRNA screens that identified kinase genetic dependencies in 117 cancer cell lines. Examples highlighted in the study, was an increased sensitivity to fibroblast growth factor receptor (FGFR) inhibitors in osteosarcoma cell lines, and to mitotic inhibitors in *SMAD4* mutant cells. Recently, a genome editing approach is being more frequently adopted which uses the bacterial CRISPER-Cas9 system, thus avoiding several pitfalls associated with siRNA screens. A large scale CRISPR-Cas9 screen was performed by Dev *et al.*, 2018 on *BRCA1*-deficient breast cancer cells treated with PARP inhibitors. Two previously uncharacterised proteins were identified; C20orf196 and FAM35A, whose inactivation correlated with PARP inhibitor resistance.

Another well-founded method to study drug-resistance *in vitro*, is to establish a resistant cell line by exposing drug-sensitive cells to the drug of interest. This can result in the generation of resistant clones which can undergo extensive characterisation to determine the cause or mechanism of resistance (Garraway and Jänne, 2012a). This can be achieved through short-term pulsing of cells with a high drug concentration, allowing for recovery in drug-free media, thereby mimicking the cycling doses of an intravenously administered drug in a clinical setting. However, resistance in this method is often low and transient (McDermott *et al.,* 2014). A preferred method of generating resistant cell lines is long-term chronic exposure to the drug. The cells can undergo dose-escalation with high concentrations of a compound, or a single high concentration dose of the compound is

administered until a resistant population emerges. These methods have been shown to generate higher levels of resistance, and a much more stable phenotype than the short-term methods (McDermott *et al.,* 2014). The *in vitro* models with acquired drug resistance has been shown to be beneficial in understanding the mechanisms of resistance, as seen in the recent works of Michaelis *et al.,*. Michaelis and colleagues have co-developed a large Resistant Cancer Cell Line (RCCL) collection and shown that resistance to nutlin, an MDM2 inhibitor, in neuroblastoma cell lines (UKF-NB-3) is due to the formation of *de novo* p53 mutations, which has now been confirmed in clinical liposarcoma using liquid biopsy methods (Michaelis *et al.*, 2011, 2012; Jung *et al.*, 2016).

Triple negative breast cancers (TNBC) are highly metastatic aggressive breast cancers which are generally susceptible to chemotherapy initially, however, the early complete response does not correlate with overall patient survival. Often TNBC patients relapse within three to five years due to resistance to the standard chemotherapy treatment being administered, and so there is an urgent need for second, or even third line treatments (Lehmann *et al.*, 2011). Extensive studies have shown that once drug-resistance has emerged, cross-resistance to several, structurally, and functionally unrelated drugs can occur (J. Wang *et al.*, 2017). This multi-drug-resistance (MDR) phenotype may be developed through non-drug specific resistance mechanisms, such as the expression of drug efflux pumps in the cell membrane that reduce intracellular drug levels to less than therapeutic concentrations (Pluchino *et al.*, 2012). This generates a difficult ongoing clinical problem that still needs to be overcome. An alternative strategy proposed by Hall *et al.,* 2009, was to identify an “Achilles heel” in which the emerged resistance to one drug may have conferred a hypersensitivity to another drug. The term “collateral sensitivity”, or acquired vulnerability, was coined to described this phenomenon, and the first full report of this was seen in the work of Bech-Hansen *et al.,* 1976, where the group used a series of MDR sub-lines derived from the Chinese Hamster Ovary (CHO) in increasing concentrations of colchicine to understand cross-resistance. This phenomenon can be harnessed for patient treatment in the clinic, and by determining the patterns of cross-resistance or acquired vulnerability to clinically relevant drugs, this will help inform clinicians on the next line of therapy in treating TNBC.

This chapter characterises a panel of chemo-resistant TNBC cell line models which have acquired resistance to the chemotherapeutic agents; cisplatin, doxorubicin, eribulin, gemcitabine, paclitaxel and 5-Fluorouracil. The cell line models were obtained from the RCCL collection in which the development of acquired drug resistance had been induced through chronic exposure to the drug. The chapter also sought to determine any cross-resistance or acquired vulnerability phenotypes to the panel of chemotherapeutic agents.

3.2 Results

3.2.1 Growth characteristics of chemo-naive and chemo-resistant TNBC cell lines

A panel of chemo-naive and chemo-resistant TNBC cell lines were selected from the Resistant Cancer Cell Line (RCCL) Collection for analysis of chemo-resistant mechanisms. TNBC is known to be a heterogeneous cancer type, therefore three chemo-naive cells lines were initially selected to reflect a range of tumour location, stage, grade and ethnicity as shown in Table 3.1.

Table 3.1. Origin of chemo-naive TNBC cell lines

	MDA-MB-468	HCC38	HCC1806
Tumour Location	Breast carcinoma	Breast, Ductal carcinoma	Breast, Ductal, squamous cell carcinoma
Stage	-	III	II
Grade	-	II	II
Age	51	50	60
Gender	Female	Female	Female
Ethnicity	Black	Caucasian	Black
Cell line source	Metastasis	-	Primary

The chemo-resistant cell lines were previously developed through long-term incubation of the chemo-naive cells in increasing concentrations of the stated chemotherapeutic agent, as previously described by Cinatl *et al.*, 1999. The chemo-resistant cell lines chosen for this project have been developed to have a resistant phenotype to chemotherapy agents common or historical to the treatment of TNBC. Along-side the three chemo-naive TNBC cell lines, a total of fifteen chemo-resistant sub-lines were selected for analysis (Table 3.2).

Table 3.2. Chemo-naive and chemo-resistant TNBC cell lines selected from the RCCL

		MDA-MB-468	HCC38	HCC1806
	Chemo-naive	X	x	X
Chemo-resistant cell lines	**Cisplatin**	X	X	X
	Doxorubicin	X	X	X
	Eribulin	X	X	X
	Gemcitabine	-	X	X
	Paclitaxel	X	X	X
	5-Fluorouracil	-	-	X

First, the morphology of each of the TNBC cell lines was established. Microscopy was used to determine differences in shape, structure and form, and images were taken for each of the TNBC cell lines (Figure 3.1).

The MDA-MB-468 cell line consisted of small, round, convex cells, grape-like in structure, which often grew on top of each other forming chains away from the flask surface. MDA-MB-468^{r}CDDP^{1000} looked similar to that of MDA-MB-468, except the formation of chains of cells were more apparent. MDA-MB-468^{r}DOX^{50} were similar to the MDA-MB-468 in morphology. MDA-MB-468^{r}ERI^{50} had a mixture of cell morphology, with a sub-population likened to the morphology of MDA-MB-468, and another sub-population containing small, elongated, round cells. MDA-MB-468^{r}PCL^{20} had a morphology distinct from MDA-MB-468. The cells were observed to be elongated, spindly cells, which could be considered to be mesenchymal in shape. Furthermore, the cells did not grow on top of each, or form chains, but would pack tightly together on the surface of the flask.

HCC38 consisted of large, flat, thin cells, which were almost transparent on certain microscopic planes. These were similar in shape to flat endothelial cells, and would often pack tightly together on the surface of the flask. There did not appear to be a notable difference between HCC38 and HCC38^{r}DOX^{40}, HCC38^{r}PCL$^{2.5}$ or HCC38^{r}GEM^{20}. Both HCC38^{r}CDDP^{3000} and HCC38^{r}ERI^{10} showed to be slightly smaller than HCC38, and would pack into discrete islands when growing.

HCC1806 consisted of small, polygonal, raised cells, similar to the shape of epithelial cells. These cells grew in discrete islands on the flask surface, leaving areas of the flask empty. There was no notable difference between the morphology of HCC1806 and the HCC1806 derived chemo-resistant cell lines, with the exception to HCC1806^{r}PCL^{20} and HCC1806^{r}5-F^{1500}, where the cells would grow on top of each other in layers.

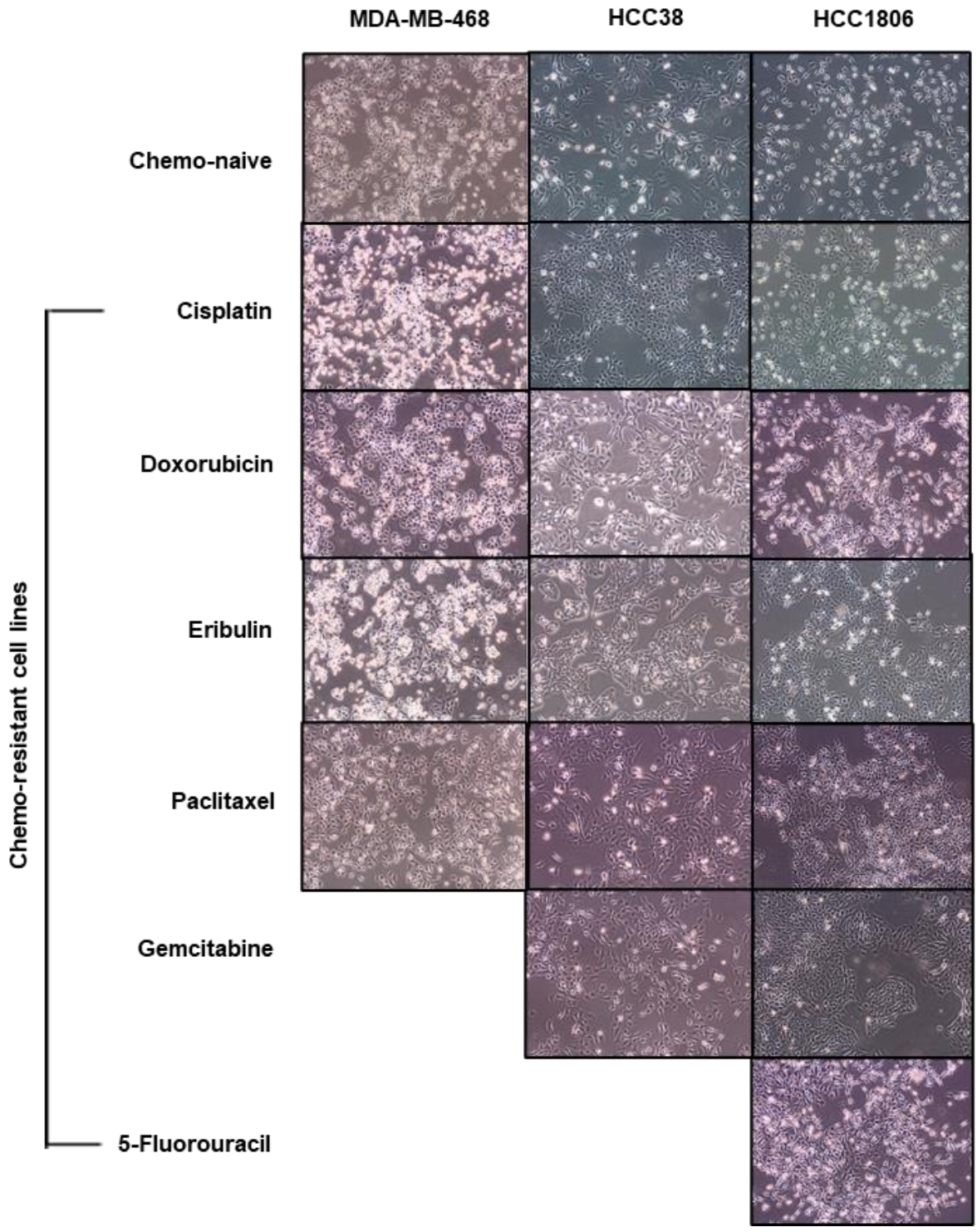


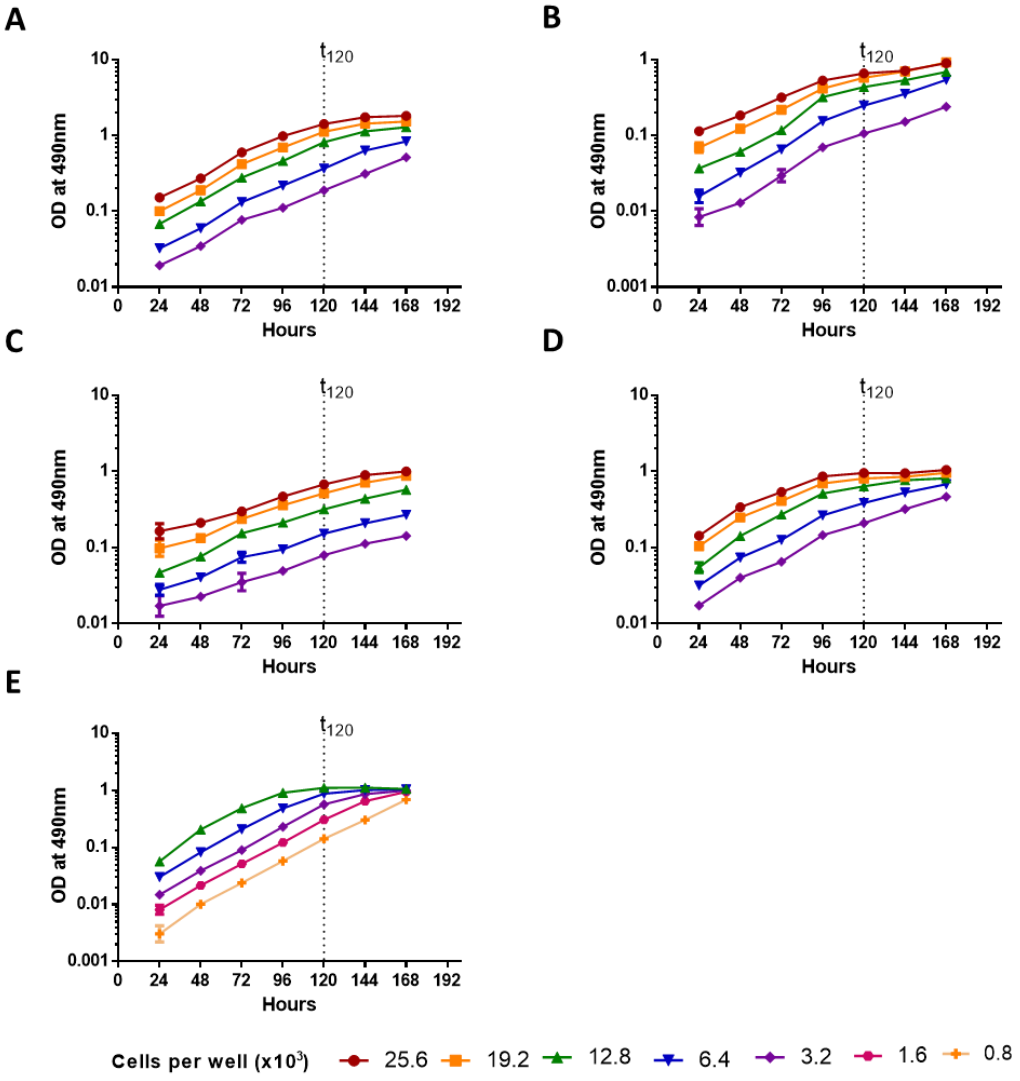
Figure 3.1 Morphology of chemo-naive and chemo-resistant TNBC cell lines
Morphology of each of the chemo-naive and chemo-resistant TNBC cell lines were analysed using microscopy. Cells were visualised at 40x magnification, and images taken. Images are representative of ≥ 3 independent experiments.

Next, the growth characteristics of each of the chemo-naive and chemo-resistant TNBC cell lines were established. In order to determine the optimal conditions for the 120-hour MTT assay, each cell line was plated at five different densities across seven 96-well plates in 200 µl complete IMDM. Every 24 hours, one plate was fixed and stained as described for the SRB growth assay (Section 2.1.4). Once all seven were stained, the data was analysed and raw absorbances were used to generate growth curves in GraphPad Prism 6. The ideal plating density is considered to be when the cells remain in exponential growth during exposure to the drug within the 120 hours of the MTT experiment. This allows for consistency between the cell lines in later experiments, while having optimal uptake and response to the drug during the assays. Further to selecting the ideal plating density, the logarithmic doubling time for each of the TNBC cell lines could be calculated. The doubling times was averaged across at least three independent experiments in each cell line, and calculated as described in the Materials and Methods (section 2.1.3).

Both MDA-MB-468^{r}DOX^{50} and MDA-MB-468^{r}ERI^{50} demonstrated an increase in doubling time from MDA-MB-468 with 48.49 ± 0.87 and 42.69 ± 7.05 hours respectively (Figure 3.2). MDA-MB-468^{r}CDDP^{1000} showed a slightly faster doubling time compared to MDA-MB-468 with 34.87 ± 2.25 hours, whilst MDA-MB-468^{r}PCL^{20} showed a much faster doubling time with 31.27 ± 2.9 hours (Figure 3.2). For each of the cell lines, 12.8 x10^{3} was determined at the optimal plating density, with the exception of MDA-MB-468^{r}PCL^{20}, where a lower optimal cell density of 6.4 x10^{3} cells per well was chosen.

HCC38^{r}CDDP^{3000} had a faster doubling time than HCC38 with a time of 36.7 ± 9.19 hours (Figure 3.3). HCC38^{r}DOX^{40}, HCC38^{r}ERI^{10}, HCC38^{r}GEM^{20} and HCC38^{r}PCL$^{2.5}$ all had a slower doubling time than HCC38 with times; 45.19 ± 3.37, 45.18 ± 5.52, 42.6 ± 6.80, and 44.72 ± 1.04 respectively (Figure 3.3). The HCC38 and HCC38 derived chemo-resistant cell lines optimal plating density was 12.8 x10^{3} with the exception to HCC38^{r}CDDP^{3000} and HCC38^{r}DOX^{40} where it was chosen to plate at a higher density of 19.2 x10^{3} cells per well (Figure 3.3).

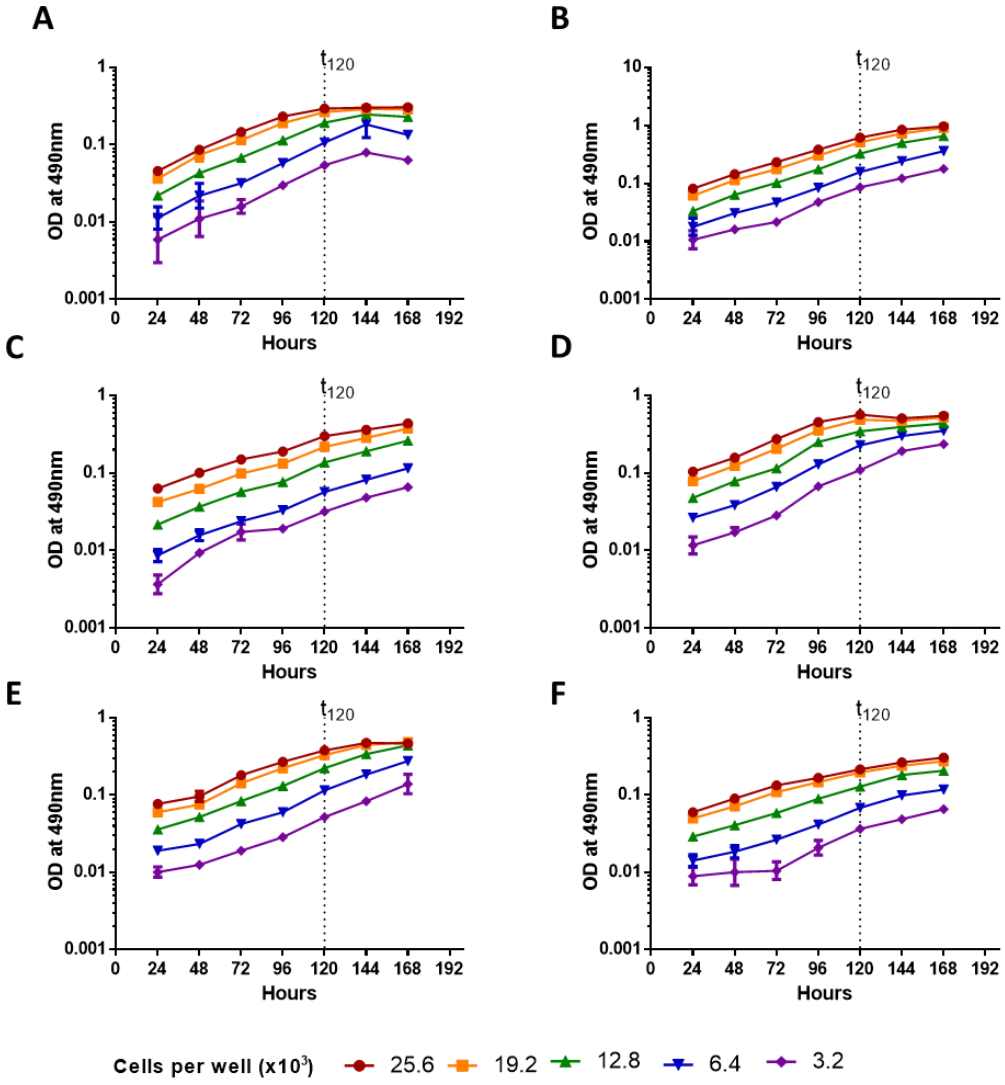
Both HCC1806^{r}DOX$^{12.5}$ and HCC1806^{r}5-F^{1500} had a faster doubling time compared to HCC38 with a time of 30.35 ± 1.09 and 32.15 ± 0.65 respectively (Figure 3.4). HCC1806^{r}CDDP^{500}, HCC1806^{r}ERI^{50}, HCC1806^{r}GEM^{20} and HCC1806^{r}PCL^{20} had a slower doubling time compared to HCC38 with 39.10 ± 5.32, 44.45 ± 3.74, 43.09 ± 3 and 39.82 ± 6.43 respectively (Figure 3.4). Each of this set were plated at 12.8 x10^{3} cells per well.



F

	Doubling time (hours)	Optimal Plating density (cells per well x10³)
MDA-MB-468	36.07 ± 2.25	12.8
MDA-MB-468^{r}CDDP^{1000}	34.87 ± 3.19	12.8
MDA-MB-468^{r}DOX^{50}	48.49 ± 0.87	12.8
MDA-MB-468^{r}ERI^{100}	42.69 ± 7.05	12.8
MDA-MB-468^{r}PCL^{20}	31.27 ± 2.9	6.4

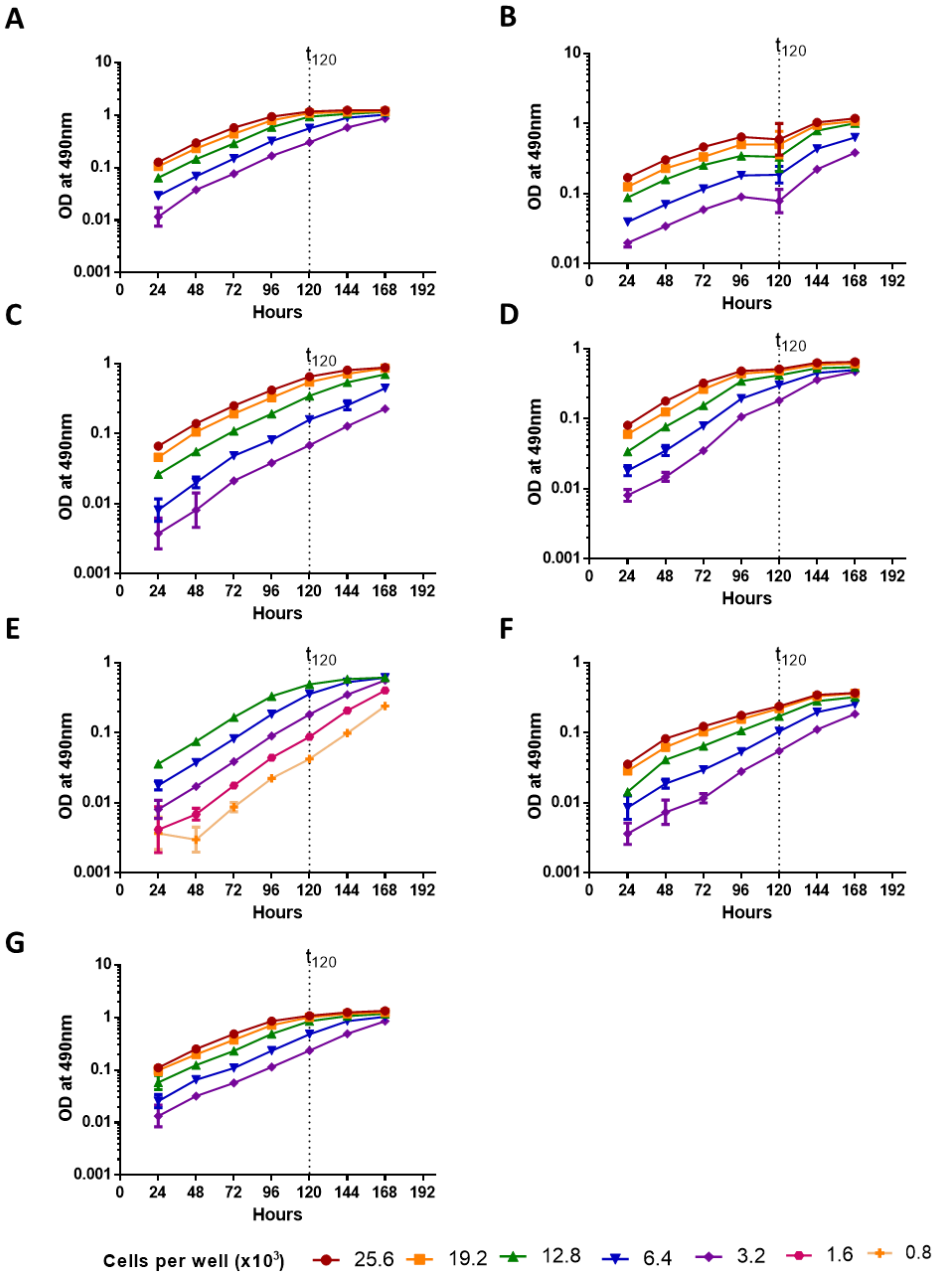
Figure 3.2. Growth characterisation of MDA-MB-468 and chemo-naive and chemo-resistant cell lines
Cells were seeded with five technical replicates in seven 96-well plates at the cell number indicated. One plate was fixed every 24 hours and analysed by SRB assay. Growth curves A) MDA-MB-468 B) MDA-MB-468^{r}CDDP^{1000}, C) MDA-MB-468^{r}DOX^{50}, D) MDA-MB-468^{r}ERI^{100}, E) MDA-MB-468^{r}PCL^{20}, were generated using GraphPad Prism 6. Data points represent mean ± SD where graphs are a representation of ≥3 independent experiments. Dotted line at t_{120} represents endpoint of MTT assay. Variable y-axis between graphs. F) Table shows calculated mean ± SD doubling time from ≥3 independent experiments and the chosen optimal plating density.



G

	Doubling time (hours)	Optimal Plating density (cells per well x10³)
HCC38	38.58	12.8
HCC38^{r}CDDP^{3000}	36.7 ± 9.19	19.2
HCC38^{r}DOX^{40}	45.19 ± 3.37	19.2
HCC38^{r}ERI^{10}	45.18 ± 5.52	12.8
HCC38^{r}GEM^{20}	42.6 ± 6.80	12.8
HCC38^{r}PCL$^{2.5}$	44.72 ± 1.04	12.8

Figure 3.3. Growth characterisation of HCC38 and chemo-naive and chemo-resistant cell lines
Cells were seeded with five technical replicates in seven 96-well plates at the cell number indicated. One plate was fixed every 24 hours and analysed by SRB assay. Growth curves A) HCC38 B) HCC38^{r}CDDP^{3000}, C) HCC38^{r}DOX^{40}, D) HCC38^{r}ERI^{10}, E) HCC38^{r}GEM^{20}, F) HCC38^{r}PCL$^{2.5}$, were generated using GraphPad Prism 6. Data points represent mean ± SD where graphs are a representation of ≥3 independent experiments. Dotted line at t_{120} represents endpoint of MTT assay. Variable y-axis between graphs. G) Table shows calculated mean ± SD doubling time from ≥3 independent experiments and the chosen optimal plating density.



H

	Doubling time (hours)	Optimal Plating density (cells per well x10³)
HCC1806	34.63 ± 4.54	12.8
HCC1806^{r}CDDP^{500}	39.10 ± 5.32	12.8
HCC1806^{r}DOX$^{12.5}$	30.35 ± 1.09	12.8
HCC1806^{r}ERI^{50}	44.45 ± 3.74	12.8
HCC1806^{r}GEM^{20}	43.09 ± 3	12.8
HCC1806^{r}PCL^{20}	39.82 ± 6.43	12.8
HCC1806^{r}5-F^{1500}	32.15 ± 0.65	12.8

Figure 3.4. Growth characterisation of HCC1806 and chemo-naive and chemo-resistant cell lines
Cells were seeded with five technical replicates in seven 96-well plates at the cell number indicated. One plate was fixed every 24 hours and analysed by SRB assay. Growth curves A) HCC1806 B) HCC1806^{r}CDDP^{500}, C) HCC1806^{r}DOX$^{12.5}$, D) HCC1806^{r}ERI^{50}, E) HCC1806^{r}GEM^{20}, F) HCC1806^{r}PCL^{20} G) HCC1806^{r}5-F^{1500} were generated using GraphPad Prism 6. Data points represent mean ± SD where graphs are a representation of ≥3 independent experiments. Dotted line at t_{120} represents endpoint of MTT assay. Variable y-axis between graphs. H) Table shows calculated mean ± SD doubling time from ≥3 independent experiments and the chosen optimal plating density.

3.2.2 Confirmation of resistant phenotype in chemo-resistant cell lines

Having determined the optimal plating densities for each of the TNBC cell lines, the response of the chemo-resistant cell lines to the chemotherapeutic agents they have been adapted to, were studied. Dose-response curves were generated to a serial dilution of each compound using a 120-hour MTT assay allowing for the calculation of the half-maximal growth inhibitory concentration (GI_{50}) and a confirmation of the resistant phenotype. The MTT assays were conducted on both the chemo-naive cell lines and the chemo-resistant cell lines to allow for the calculation of the resistance factor (RF), a ratio of the chemo-resistant cell line GI_{50} to the chemo-naive cell line GI_{50}. The RF is used to define the level of resistance seen in the chemo-resistant cell lines, whereby a RF ≥ 2 is indicative of a resistance phenotype, whilst an RF ≤ 0.5 is indicative of a sensitive phenotype. The RF is calculated as per the following equation:

$$Resistance\ factor\ (RF) = \frac{Chemo-resistant\ cell\ line\ \mathrm{GI50}}{Chemo-naive\ cell\ line\ GI50}$$

Analysis of the cisplatin resistant cell lines confirmed a resistance phenotype is seen relative to the chemo-naive counterpart (Figure 3.5). MDA-MB-468 demonstrated the most sensitivity to cisplatin, compared to the two other chemo-naive cell lines, with a GI_{50} value of 0.027 µM. The highest RF was observed in MDA-MB-468^{r}CDDP^{1000} with an RF of 319, whilst HCC38^{r}CDDP^{3000} and HCC1806^{r}CDDP^{500} demonstrated a RF of 8.88 and 7.71 respectively.

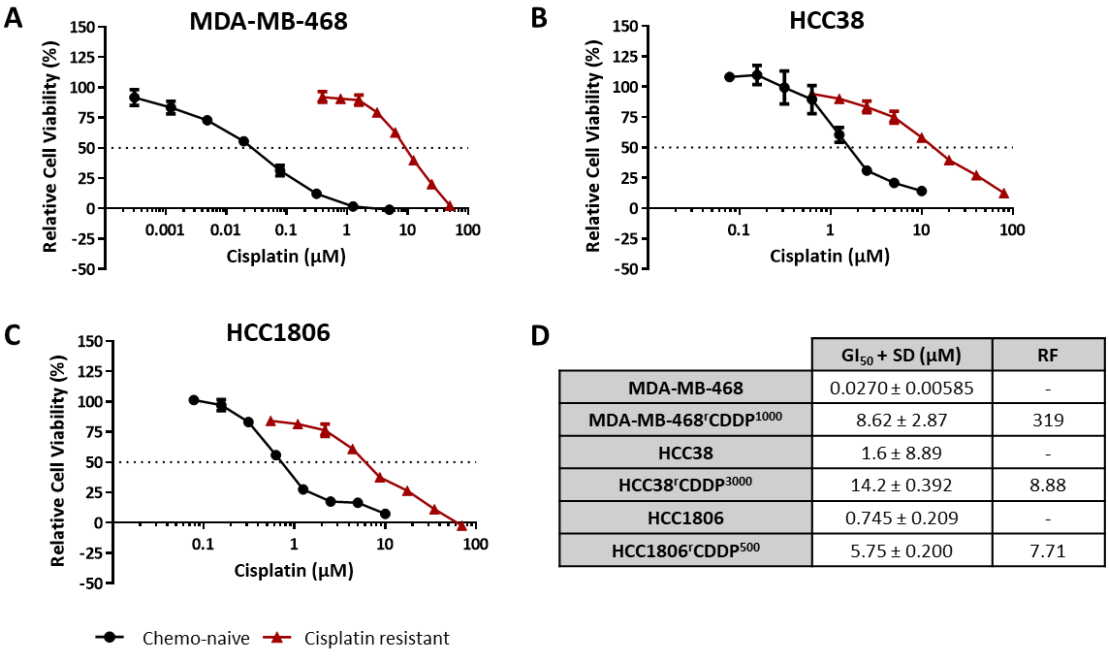
Examination of the doxorubicin resistant cell lines confirmed a resistance phenotype is seen relative to the chemo-naive counterpart (Figure 3.6). The MDA-MB-468 cell line demonstrated the most sensitivity to doxorubicin compared to the two other chemo-naive cell lines with a GI_{50} value of 0.00518 µM. HCC1806^{r}DOX$^{12.5}$ demonstrated the highest RF of 58.7, whilst MDA-MB-468^{r}DOX^{50} and HCC38^{r}DOX^{40} had RF's of 31 and 12.9 respectively.

Each of the eribulin resistant cell lines demonstrated a resistance phenotype relative to the chemo-naive counterpart (Figure 3.7). MDA-MB-468 cell line demonstrated the most sensitivity to eribulin compared to the two other chemo-naive cell lines with a GI_{50} value of 0.000167 µM. It was noted that each of the cell lines showed nM sensitivity to eribulin. HCC1806^{r}ERI^{50} had the highest RF of 5920, whilst MDA-MB-468^{r}ERI^{50} and HCC38^{r}ERI^{10} had RF's of 314 and 46.9 respectively.

Both of the gemcitabine resistant cell lines demonstrated a resistance phenotype compared to the chemo-naive counterparts (Figure 3.8). A similar sensitivity to gemcitabine was observed in HCC38 and HCC1806 with GI_{50} values of 0.00197 µM and 0.00104 µM respectively. HCC1806^{r}GEM^{20} demonstrated the greatest fold resistance with an RF of 99.7, whilst HCC38^{r}GEM^{20} had an RF of 44.3.

Analysis of the paclitaxel resistant cell lines confirmed a resistance phenotype is seen relative to the chemo-naive counterpart (Figure 3.9). HCC180 demonstrated the most sensitivity to paclitaxel compared to the two other chemo-naive cell lines with a GI_{50} value of 0.0000623 µM. HCC1806^{r}PCL^{20} demonstrated the highest RF, with an RF of 165, whilst MDA-MB-468^{r}PCL^{20} and HCC38^{r}PCL$^{2.5}$ demonstrated RF's of 100 and 5.51 respectively.

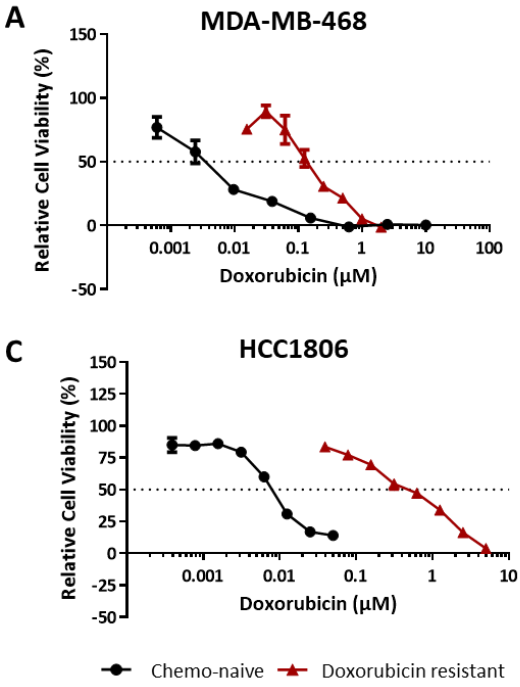
Only one cell line was selected for the panel which is considered to be resistant to 5-Fluorouracil. Analysis of HCC1806^{r}5-F^{1500} confirmed a resistance phenotype is seen relative to HCC1806 (Figure 3.10). HCC1806 had a calculated GI_{50} value of 2.94 µM whilst HCC1806^{r}5-F^{1500} had a GI_{50} value of 318 µM, which calculated to an RF of 108.



D

	GI_{50} + SD (µM)	RF
MDA-MB-468	0.0270 ± 0.00585	-
MDA-MB-468^{r}CDDP^{1000}	8.62 ± 2.87	319
HCC38	1.6 ± 8.89	-
HCC38^{r}CDDP^{3000}	14.2 ± 0.392	8.88
HCC1806	0.745 ± 0.209	-
HCC1806^{r}CDDP^{500}	5.75 ± 0.200	7.71

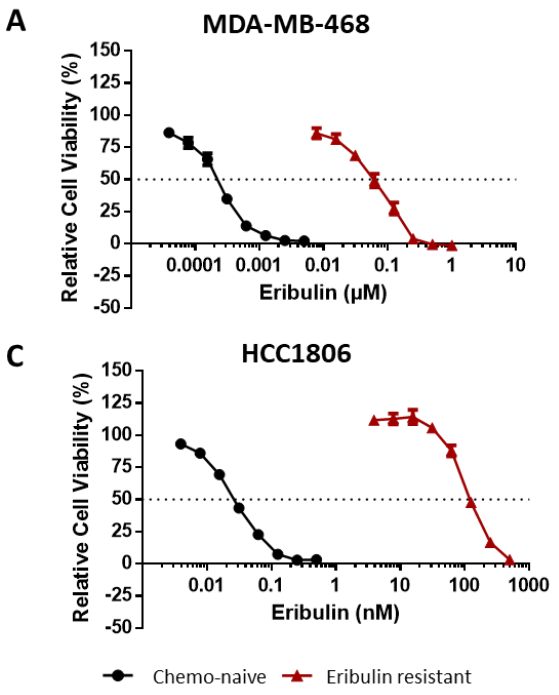
Figure 3.5. GI_{50} value determinations for cisplatin in chemo-naive and cisplatin-resistant TNBC cell lines
Response of chemo-naive and cisplatin-resistant cell lines to cisplatin were determined by 120-hour MTT assay and analysed using GraphPad Prism 6. Dose response curves to A) MDA-MB-468 cell lines, B) HCC38 cell lines, C) HCC1806 cell lines. Data points represent mean ± SD where graphs are a representation of ≥3 independent experiments. Broken line on Y-axis indicates half-maximal growth inhibition. D) Table summarising the mean ± SD GI_{50} values when cells were treated with cisplatin and the calculated RF.



D

	GI_{50} + SD (μM)	RF
MDA-MB-468	0.00518 ± 0.00252	-
MDA-MB-468^{r}DOX^{50}	0.161 ± 0.427	31.029
HCC38	0.0153 ± 0.00556	-
HCC38^{r}DOX^{40}	0.2 ± 0.0724	12.9
HCC1806	0.00669 ± 0.000823	-
HCC1806^{r}DOX$^{12.5}$	0.393 ± 0.0764	58.7

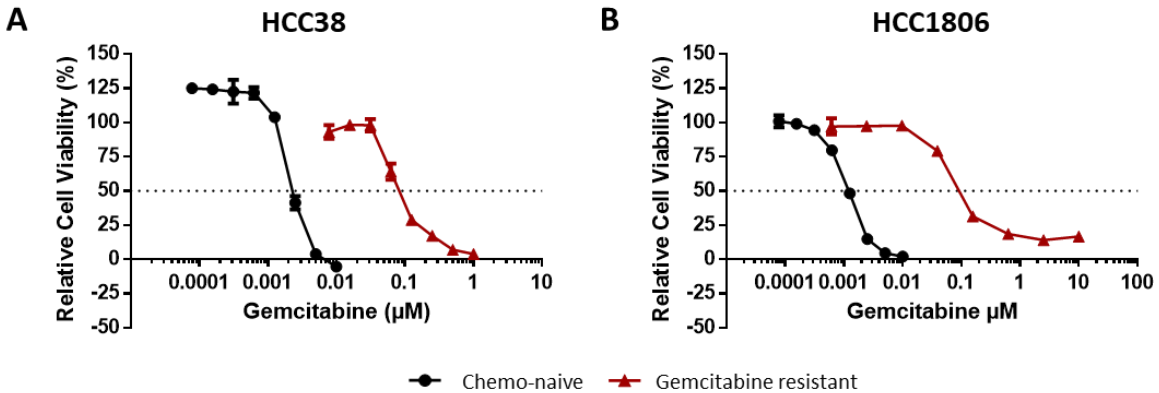
Figure 3.6. GI_{50} value determinations for doxorubicin in chemo-naive and doxorubicin-resistant TNBC cell lines
Response of chemo-naive and doxorubicin-resistant cell lines to doxorubicin were determined by 120-hour MTT assay and analysed using GraphPad Prism 6. Dose-response curves to A) MDA-MB-468 cell lines, B) HCC38 cell lines, C) HCC1806 cell lines. Data points represent mean ± SD where graphs are a representation of ≥3 independent experiments. Broken line on Y-axis indicates half-maximal growth inhibition. D) Table summarising the mean ± SD GI_{50} values when cells were treated with doxorubicin and calculated the RF.



D

	GI_{50} + SD (μM)	RF
MDA-MB-468	0.000167 ± 0.0000358	-
MDA-MB-468^{r}ERI^{100}	0.0526 ± 0.00504	314
HCC38	0.000195 ± 0.0000691	-
HCC38^{r}ERI^{50}	0.00914 ± 0.00242	46.9
HCC1806	0.0000265 ± 0.00000727	-
HCC1806^{r}ERI^{50}	0.157 ± 0.0318	5920

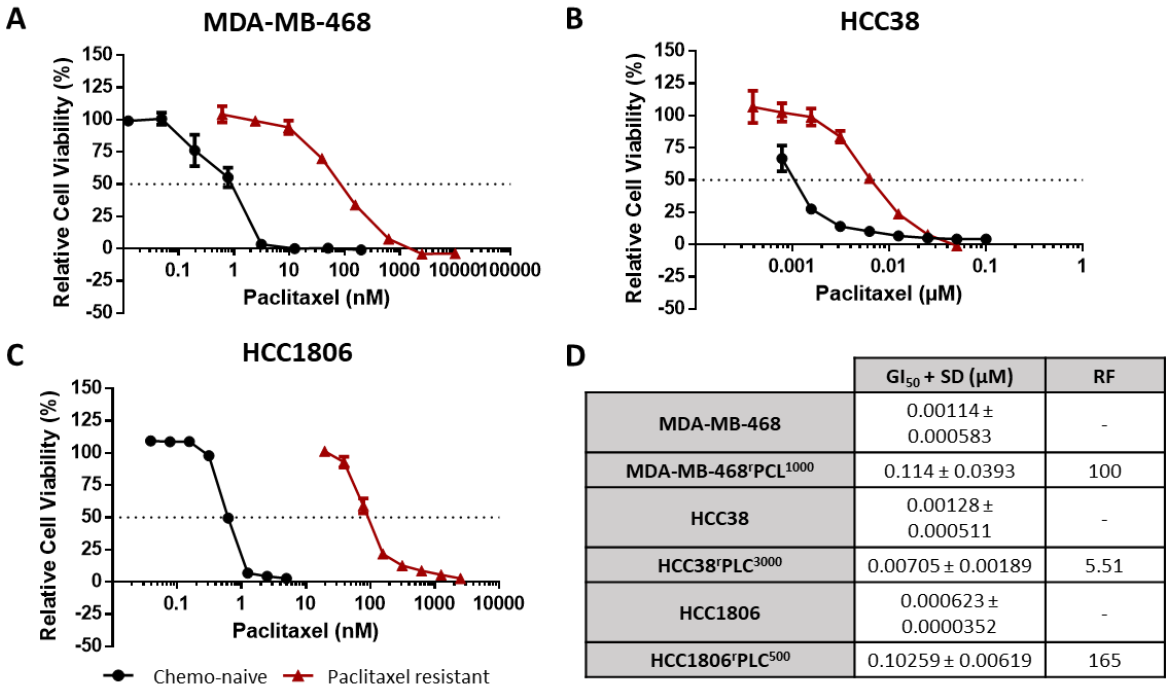
Figure 3.7. GI_{50} value determinations for eribulin in chemo-naive and eribulin-resistant TNBC cell lines
Response of chemo-naive and eribulin-resistant cell lines to eribulin were determined by 120-hour MTT assay and analysed using GraphPad Prism 6. Dose-response curves to A) MDA-MB-468 cell lines, B) HCC38 cell lines, C) HCC1806 cell lines. Data points represent mean ± SD where graphs are a representation of ≥3 independent experiments. Broken line on Y-axis indicates half-maximal growth inhibition. D) Table summarising the mean ± SD GI_{50} values when cells were treated with eribulin and the calculated RF.



C

	GI_{50} + SD (μM)	RF
HCC38	0.00197 ± 0.000147	-
HCC38^{r}GEM^{20}	0.0877 ± 0.00575	44.3
HCC1806	0.00104 ± 0.000156	-
HCC1806^{r}GEM^{20}	0.104 ± 0.0145	99.7

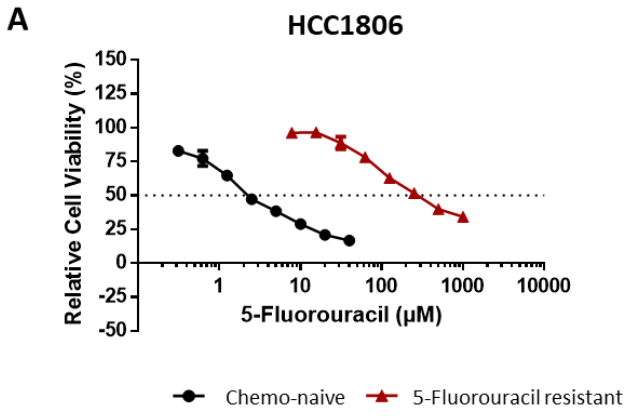
Figure 3.8. GI_{50} value determinations for gemcitabine in chemo-naive and gemcitabine-resistant TNBC cell lines
Response of chemo-naive and gemcitabine-resistant cell lines to gemcitabine were determined by 120-hour MTT assay and analysed using GraphPad Prism 6. Dose-response curves to A) HCC38 cell lines, B) HCC1806 cell lines. Data points represent mean ± SD where graphs are a representation of ≥3 independent experiments. Broken line on Y-axis indicates half-maximal growth inhibition. C) Table summarising the mean ± SD GI_{50} values when cells were treated with gemcitabine and the calculated RF.



D

	GI_{50} + SD (μM)	RF
MDA-MB-468	0.00114 ± 0.000583	-
MDA-MB-468^{r}PCL^{1000}	0.114 ± 0.0393	100
HCC38	0.00128 ± 0.000511	-
HCC38^{r}PLC^{3000}	0.00705 ± 0.00189	5.51
HCC1806	0.000623 ± 0.0000352	-
HCC1806^{r}PLC^{500}	0.10259 ± 0.00619	165

Figure 3.9. GI_{50} value determinations for paclitaxel in chemo-naive and paclitaxel-resistant TNBC cell lines
The response of chemo-naive and paclitaxel-resistant cell lines to paclitaxel were determined by 120-hour MTT assay and analysed using GraphPad Prism 6. Dose-response curves to A) MDA-MB-468 cell lines, B) HCC38 cell lines, C) HCC1806 cell lines. Data points represent mean ± SD where graphs are a representation of ≥3 independent experiments. The broken line on Y-axis indicates half-maximal growth inhibition. D) Table summarising the mean ± SD GI_{50} values when cells were treated with paclitaxel and the calculated RF.



B

	GI_{50} + SD (μM)	RF
HCC1806	2.94 ± 0.251	-
HCC1806^{r}5-F^{1500}	318 ± 38.5	108

Figure 3.10. GI_{50} value determinations for 5-Fluorouracil in chemo-naive and 5-Fluorouracil-resistant TNBC cell lines
The response of chemo-naive and 5-Fluorouracil-resistant cell line to 5-Fluorouracil were determined by 120-hour MTT assay and analysed using GraphPad Prism 6. Dose-response curves to A) HCC1806 cell lines. Data points represent mean ± SD where graphs are a representation of ≥3 independent experiments. The broken line on Y-axis indicates half-maximal growth inhibition. B) Table summarising the mean ± SD GI_{50} values when cells were treated with for 5-Fluorouracil and the calculated RF.

3.2.3 Cross-resistance profiling of chemo-naive and chemo-resistant TNBC cell lines

Having confirmed the resistance phenotype in each of the chemo-resistant cell lines, relative to the respective chemo-naive counterpart, each of the chemo-resistant cell lines were checked for cross-resistance, or acquired vulnerability, to the following chemotherapeutic agents; cisplatin, doxorubicin, eribulin, gemcitabine, paclitaxel and 5-Fluorouracil. Dose-response curves were generated to a serial dilution of each compound using a 120-hour MTT assay allowing for the calculation of the GI_{50} and RF's. Cross-resistance is termed whereby a cell line which has developed resistance to a given drug, demonstrates a resistance phenotype when treated with a different drug. Acquired vulnerability is termed whereby a cell line, which demonstrates resistance to a given drug demonstrates a sensitive phenotype, when treated with a different drug compared to the chemo-naive counterpart. Here cross-resistance is defined when the calculated RF ≥ 2, and acquired vulnerability when the calculated RF ≤ 0.5, and both demonstrated statistical significance using the student's t-test ($p \leq 0.05$). Numerical data presented in Appendix A2.

As already seen, the three cisplatin-resistant cell lines show resistance to treatment with cisplatin (Figure 3.11). MDA-MB-468^{r}PCL^{20} demonstrates cross-resistance to treatment with cisplatin with an RF of 49.76, which is higher than the RF calculated for HCC38^{r}CDDP^{3000} and HCC1806^{r}CDDP^{500}. HCC38^{r}GEM^{20} also demonstrates cross-resistance to treatment with cisplatin with an RF of 2.3 (Figure 3.11). The rest of the chemo-resistant cell lines show no statistically significant change when treated with cisplatin relative to their respective chemo-naive counterparts.

The three doxorubicin-resistant cell lines show resistance to treatment with doxorubicin (Figure 3.12). Many cell lines demonstrated cross-resistance to the treatment of doxorubicin, including; MDA-MB-468^{r}CDDP^{1000}, MDA-MB-468^{r}ERI^{50}, MDA-MB-468^{r}PCL^{20}, HCC38^{r}CDDP^{3000}, HCC38^{r}ERI^{10}, HCC1806^{r}ERI^{50}, HCC1806^{r}GEM^{20}, HCC1806^{r}PCL^{20} with RF's of 5.96, 28.39, 2.45, 3.82, 7.88, 10.79, and 23.85 respectively (Figure 3.12). Two cell lines demonstrated acquired vulnerability to the treatment of doxorubicin; HCC38^{r}GEM^{20} and HCC38^{r}PCL$^{2.5}$ with RF's of 0.36 and 0.382 respectively (Figure 3.12). The rest of the chemo-resistant cell lines show no statistically significant change when treated with doxorubicin relative to their respective chemo-naive counterparts.

As already seen, the three eribulin-resistant cell lines show resistance to treatment with eribulin (Figure 3.13). It was interesting to note that MDA-MB-468^{r}PCL^{20} and HCC1806^{r}PCL^{20} demonstrated resistance upon treatment with eribulin with RF's of 2577.57 and 9280.23 respectively, which are higher than the RF's for any of the eribulin resistant cell lines. MDA-MB-468^{r}CDDP^{1000}, HCC38^{r}ERI^{10}, HCC1806^{r}DOX$^{12.5}$ and HCC1806^{r}GEM^{20} also demonstrated cross-resistance to eribulin with RF's of 49.91, 7.36, 15.33 and 7.59 respectively. Acquired vulnerability was observed to treatment with eribulin in HCC38^{r}CDDP^{3000} and HCC38^{r}GEM^{20} with RF's of 0.12 and 0.39 respectively. The rest of the chemo-resistant cell lines show no statistically significant change when treated with eribulin relative to their respective chemo-naive counterparts.

Both of the two gemcitabine-resistant cell lines show resistance to treatment with gemcitabine (Figure 3.14). Interestingly, HCC38^{r}CDDP^{3000} demonstrated cross-resistance to gemcitabine, with an RF of 2.04. HCC1806^{r}ERI^{50} demonstrated acquired vulnerability to the treatment of gemcitabine with an RF of 0.5. The rest of the chemo-resistant cell lines show no statistically significant change when treated with gemcitabine relative to their respective chemo-naive counterparts.

As already seen, the three paclitaxel-resistant cell lines show resistance to treatment with paclitaxel (Figure 3.15). The only cell lines which were found to demonstrate cross-resistance to paclitaxel, are the three eribulin resistant cell lines; MDA-MB-468^{r}ERI^{50}, HCC38^{r}ERI^{10} and HCC1806^{r}ERI^{50} with RF's of 95.91, 5.67 and 24.95 respectively. The rest of the chemo-resistant cell lines show no statistically significant change when treated with paclitaxel relative to their respective chemo-naive counterparts.

Many cell lines demonstrated cross-resistance to the treatment of 5-Fluorouracil, including; HCC38^{r}CDDP^{3000}, HCC38^{r}DOX^{40}, HCC38^{r}GEM^{20}, HCC38^{r}PCL$^{2.5}$, HCC1806^{r}DOX$^{12.5}$, HCC1806^{r}GEM^{20} and HCC1806^{r}PCL^{20} with RF's of 36.64, 75.86, 7.69, 6.04, 4.67, 2.87 and 4.83 respectively (Figure 3.16).

Both MDA-MB-468^{r}PCL^{20} and HCC1806^{r}CDDP^{500} demonstrated acquired vulnerability to the treatment of 5-Fluorouracil with RF's of 0.061 and 0.13 respectively. The rest of the chemo-resistant cell lines show no statistically significant change when treated with 5-Fluorouracil relative to their respective chemo-naive counterparts.

Taken together, patterns of cross-resistance and acquired vulnerability were observed. No fold change was observed in the MDA-MB-468 chemo-resistant cell lines to 5-Fluorouracil, with exception to MDA-MB-468^{r}PCL^{20} which demonstrated acquired vulnerability, whilst each of the HCC38 chemo-resistant cell lines demonstrated cross-resistance to 5-Fluorouracil. No fold change was observed to gemcitabine in the MDA-MB-468 chemo-resistant cell lines, with the exception of MDA-MB-468^{r}ERI^{50} which demonstrated acquired vulnerability, and yet all demonstrated cross-resistance to doxorubicin. It was noted that each of the cell lines, which demonstrate resistance/cross-resistance to gemcitabine are cross-resistant to 5-Fluorouracil. Each of the eribulin resistant cell lines demonstrated cross-resistance to the treatment of paclitaxel, and each of the paclitaxel-resistant cell lines, with the exception of HCC38^{r}PCL$^{2.5}$, demonstrated cross-resistance to eribulin. Unique to the dataset, HCC1806^{r}5-F^{1500} demonstrated no cross-resistance or acquired vulnerability to any of the other chemotherapeutic agents.

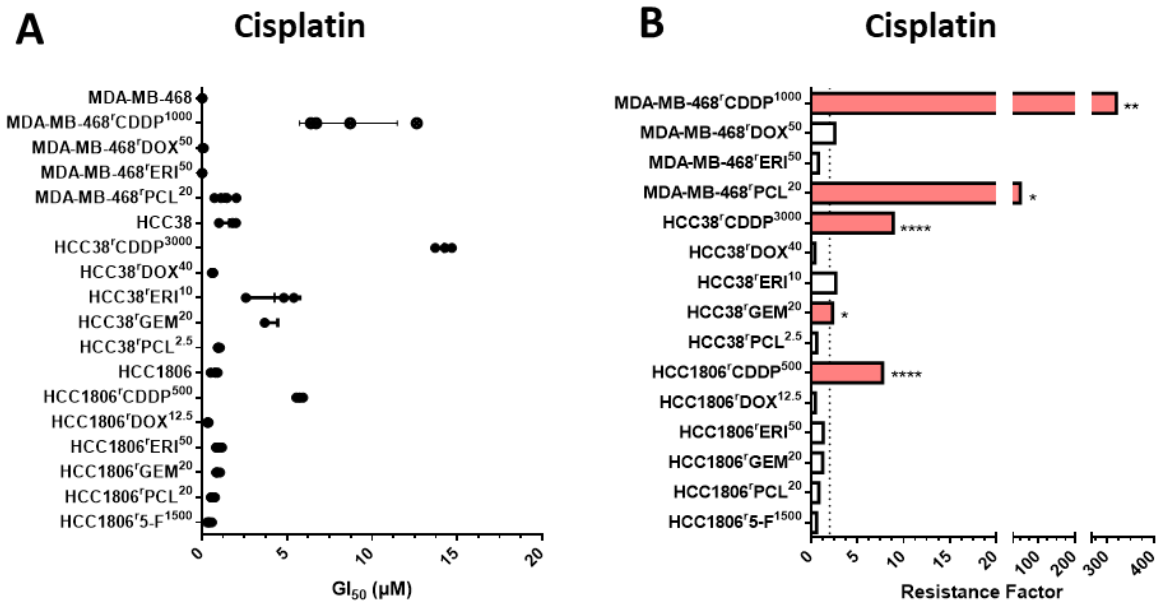


Figure 3.11. GI_{50} value determinations and resistance factors for cisplatin in the TNBC cell lines
The response of TNBC cell lines to cisplatin was determined by 120-hour MTT assay and analysed using GraphPad Prism 6. A) Determined GI_{50} concentrations upon treatment with cisplatin. Each dot represents a single biological repeat from ≥3 independent experiments with horizontal line showing standard deviation and vertical line; the mean. B) Sensitivity of the chemo-resistant cell lines relative to the respective chemo-naive cell line (RF). Dotted line is at two-fold. Statistical significance calculated using student's t-test, * P ≤ 0.05, ** p ≤ 0.01, *** p ≤ 0.001, **** p ≤ 0.00001. Colour indicates resistance status: red = resistant (RF >2, and ≥ * p), blue = acquired vulnerability (RF < 0.5 and ≥ * p), white = no change. Numerical data is presented in Appendix A2.

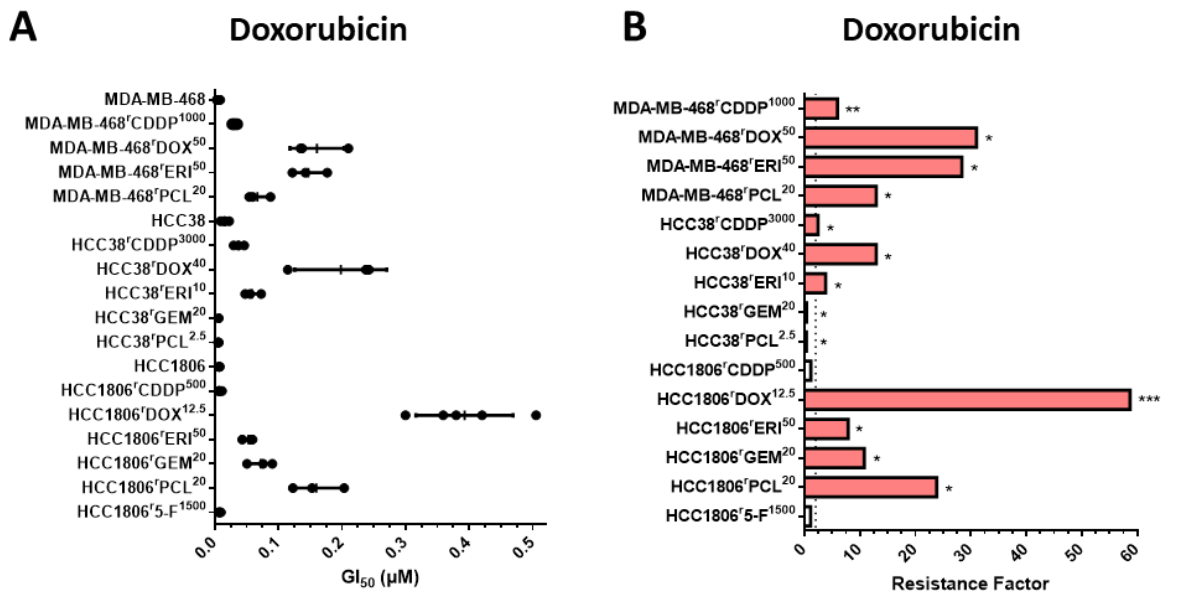


Figure 3.12. GI_{50} value determinations and resistance factors for doxorubicin in the TNBC cell lines
The response of TNBC cell lines to doxorubicin was determined by 120-hour MTT assay and analysed using GraphPad Prism 6. A) Determined GI_{50} concentrations upon treatment with doxorubicin. Each dot represents a single biological repeat from ≥3 independent experiments with horizontal line showing standard deviation and vertical line; the mean. B) Sensitivity of the drug-resistant cell lines relative to the respective chemo-naive line (RF). Dotted line is at two-fold. Statistical significance calculated using student's t-test, * P ≤ 0.05, ** p ≤ 0.01, *** p ≤ 0.001, **** p ≤ 0.00001. Colour indicates resistance status: red = resistant (RF >2, and ≥ * p), blue = acquired vulnerability (RF < 0.5 and ≥ * p), white = no change. Numerical data is presented in Appendix A2.

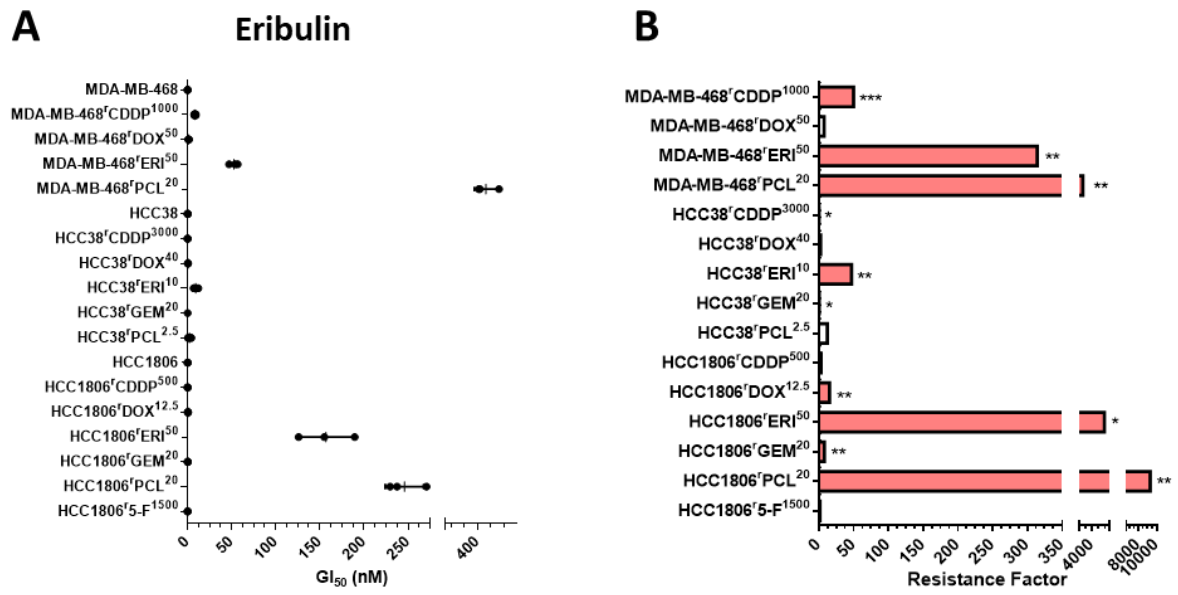


Figure 3.13. GI_{50} value determinations and resistance factors for eribulin in the TNBC cell lines
The response of TNBC cell lines to eribulin was determined by 120-hour MTT assay and analysed using GraphPad Prism 6. A) Determined GI_{50} concentrations upon treatment with eribulin. Each dot represents a single biological repeat from ≥3 independent experiments with horizontal line showing standard deviation and vertical line; the mean. B) Sensitivity of the chemo-resistant cell lines relative to the respective chemo-naive cell line (RF). Dotted line is at two-fold. Statistical significance calculated using student's t-test, * P ≤ 0.05, ** p ≤ 0.01, *** p ≤ 0.001, **** p ≤ 0.00001. Colour indicates resistance status: red = resistant (RF >2, and ≥ * p), blue = acquired vulnerability (RF < 0.5 and ≥ * p), white = no change. Numerical data is presented in Appendix A2.

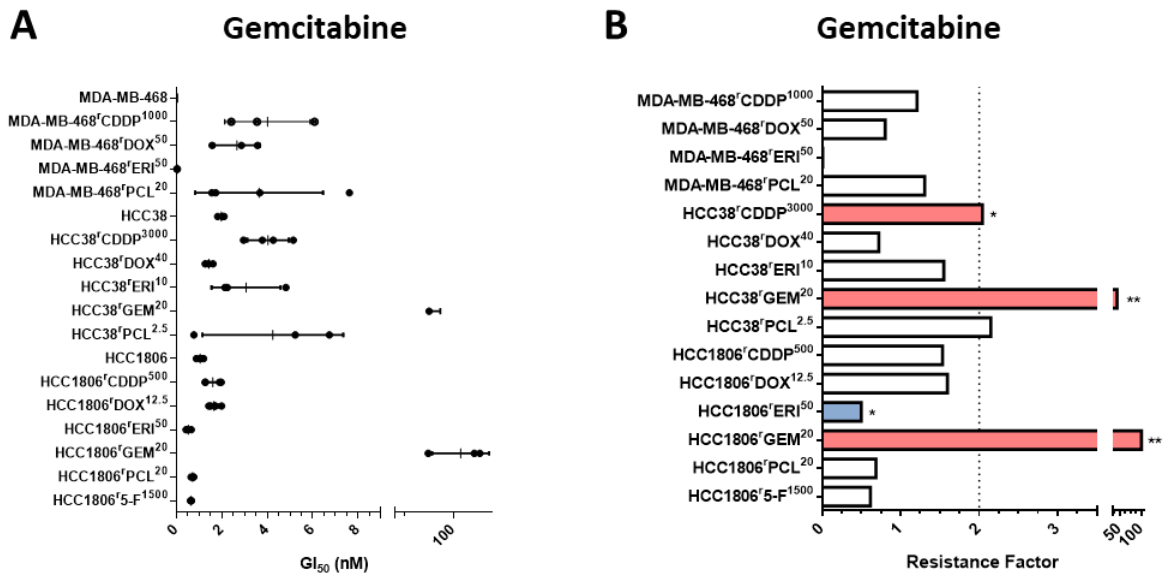


Figure 3.14. GI_{50} value determinations and resistance factors for gemcitabine in the TNBC cell lines
The response of TNBC cell lines to gemcitabine was determined by 120-hour MTT assay and analysed using GraphPad Prism 6. A) Determined GI_{50} concentrations upon treatment with gemcitabine. Each dot represents a single biological repeat from ≥3 independent experiments with horizontal line showing standard deviation and vertical line; the mean. B) Sensitivity of the chemo-resistant cell lines relative to the chemo-naive cell line (RF). Dotted line is at two-fold. Statistical significance calculated using student's t-test, * P ≤ 0.05, ** p ≤ 0.01, *** p ≤ 0.001, **** p ≤ 0.00001. Colour indicates resistance status: red = resistant (RF >2, and ≥ * p), blue = acquired vulnerability (RF < 0.5 and ≥ * p), white = no change. Numerical data is presented in Appendix A2.

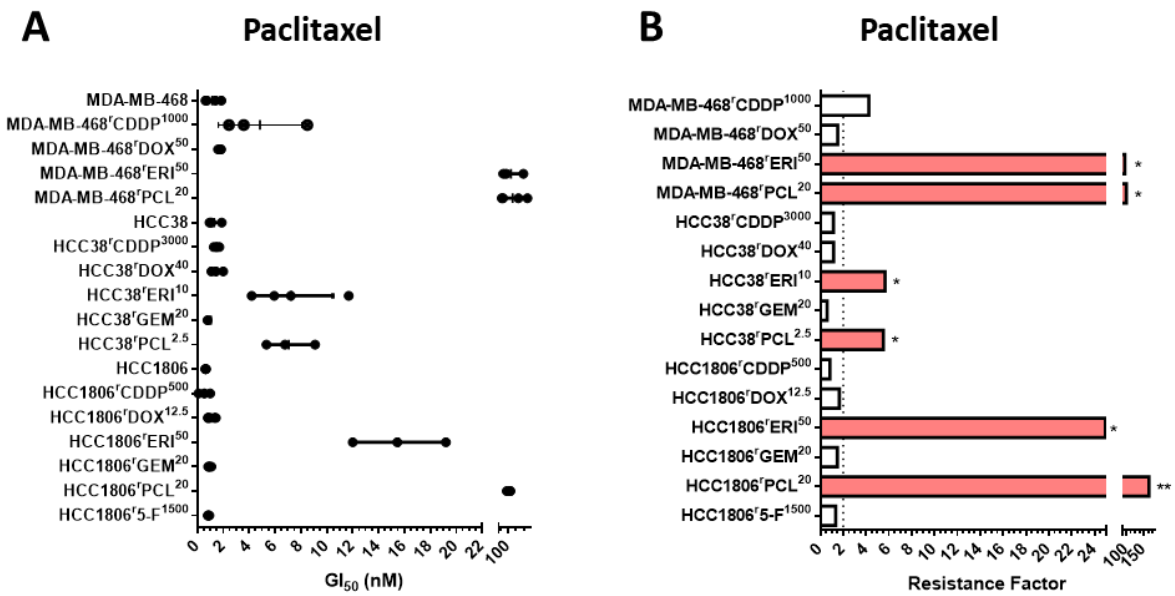


Figure 3.15. GI_{50} value determinations and resistance factors for paclitaxel in the TNBC cell lines
The response of TNBC cell lines to paclitaxel was determined by 120-hour MTT assay and analysed using GraphPad Prism 6. A) Determined GI_{50} concentrations upon treatment with paclitaxel. Each dot represents a single biological repeat from ≥3 independent experiments with horizontal line showing standard deviation and vertical line; the mean. B) Sensitivity of the chemo-resistant cell lines relative to the respective chemo-naive cell line (RF). Dotted line is at two-fold. Statistical significance calculated using student's t-test, * P ≤ 0.05, ** p ≤ 0.01, *** p ≤ 0.001, **** p ≤ 0.00001. Colour indicates resistance status: red = resistant (RF >2, and ≥ * p), blue = acquired vulnerability (RF < 0.5 and ≥ * p), white = no change. Numerical data is presented in Appendix A2.

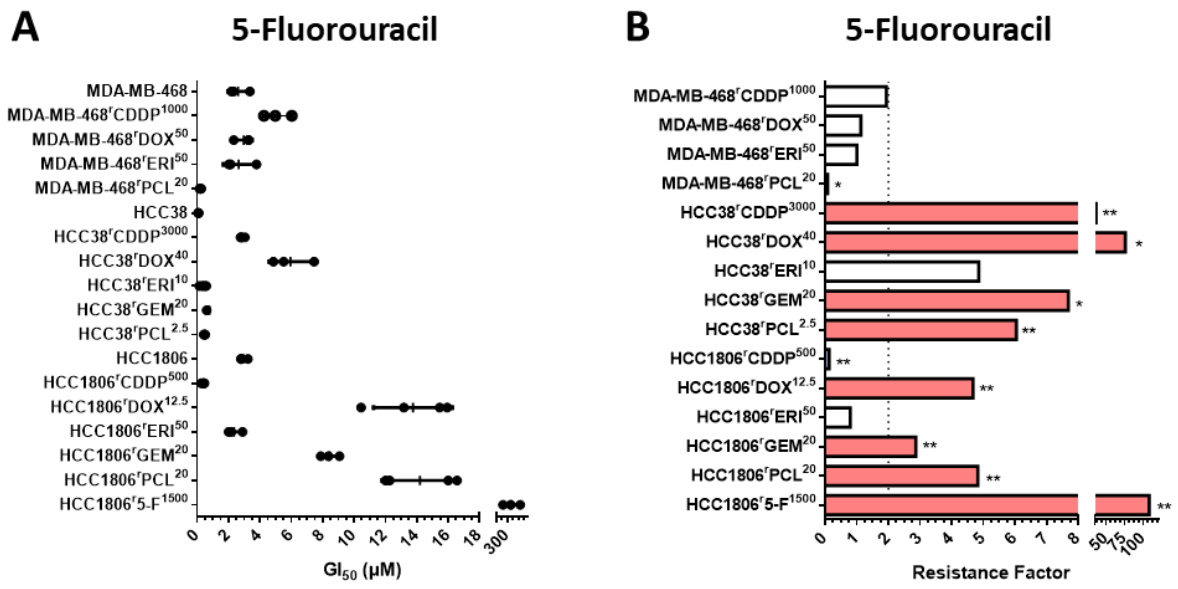


Figure 3.16. GI_{50} value determinations and resistance factors for 5-Fluorouracil in the TNBC cell lines
The response of TNBC cell lines to 5-Fluorouracil was determined by 120-hour MTT assay and analysed using GraphPad Prism 6. A) Determined GI_{50} concentrations upon treatment with 5-Fluorouracil. Each dot represents a single biological repeat from ≥3 independent experiments with horizontal line showing standard deviation and vertical line; the mean. B) Sensitivity of the chemo-resistant cell lines relative to the respective chemo-naive cell line (RF). Dotted line is at two-fold. Statistical significance calculated using student's t-test, * P ≤ 0.05, ** p ≤ 0.01, *** p ≤ 0.001, **** p ≤ 0.00001. Colour indicates resistance status: red = resistant (RF >2, and ≥ * p), blue = acquired vulnerability (RF < 0.5 and ≥ * p), white = no change. Numerical data is presented in Appendix A2.

3.2.4 Investigation of multidrug resistance protein 1 in the TNBC cell lines

Multidrug resistance (MDR) is often a significant contribution to the failure of therapeutic drugs in the clinic. As previously discussed, (section 1.3.2.1) the overexpression of active efflux transporters pumps, such as the ATP-binding cassette (ABC) transporters, contributes to resistance to multiple drugs as the drugs are removed before targeted action has been made (Gottesman, 2002; Türk, Dóra & Szakács, 2009). *ABCB1* encodes for the p-glycoprotein, or multidrug-resistant protein 1 (MDR1), which is often upregulated in cells or patient samples that demonstrate a multidrug resistant phenotype. Upregulation of MDR1 leads to increased drug export, reducing the available intracellular drug, and resulting in a resistance phenotype. Of the tested chemotherapeutic agents in this chapter, doxorubicin, eribulin and paclitaxel have all been shown to be substrates of MDR1 (Chen *et al.*, 2016; *Oba et al.,* 2016). In order to determine if MDR1 is contributing to the observed resistance phenotype, the level of MDR1 protein in each cell line was determined through western blot analysis (Figure 3.17). A vincristine resistant neuroblastoma cell line (UKF-NB-3^{r}VCR^{10}) from the RCCL collection was used as a positive MDR1 control (Michaelis *et al.*, 2009).

The western blot analysis identified MDA-MB-468^{r}ERI^{100}, MDA-MB-468^{r}PCL^{20}, HCC1806^{r}ERI^{50} and HCC1806^{r}PCL^{20} to have elevated expression of MDR1 compared to the chemo-naive counterparts.

As previously seen in this chapter, these cell lines also demonstrated resistance/cross-resistance when treated with doxorubicin, eribulin and paclitaxel, all of which are known substrates of MDR1 (Chen *et al.*, 2016; *Oba et al.,* 2016). This suggests that the cross-resistance of these cell lines to MDR1 substrates is, at least in part, caused by increased ABCB1 levels.

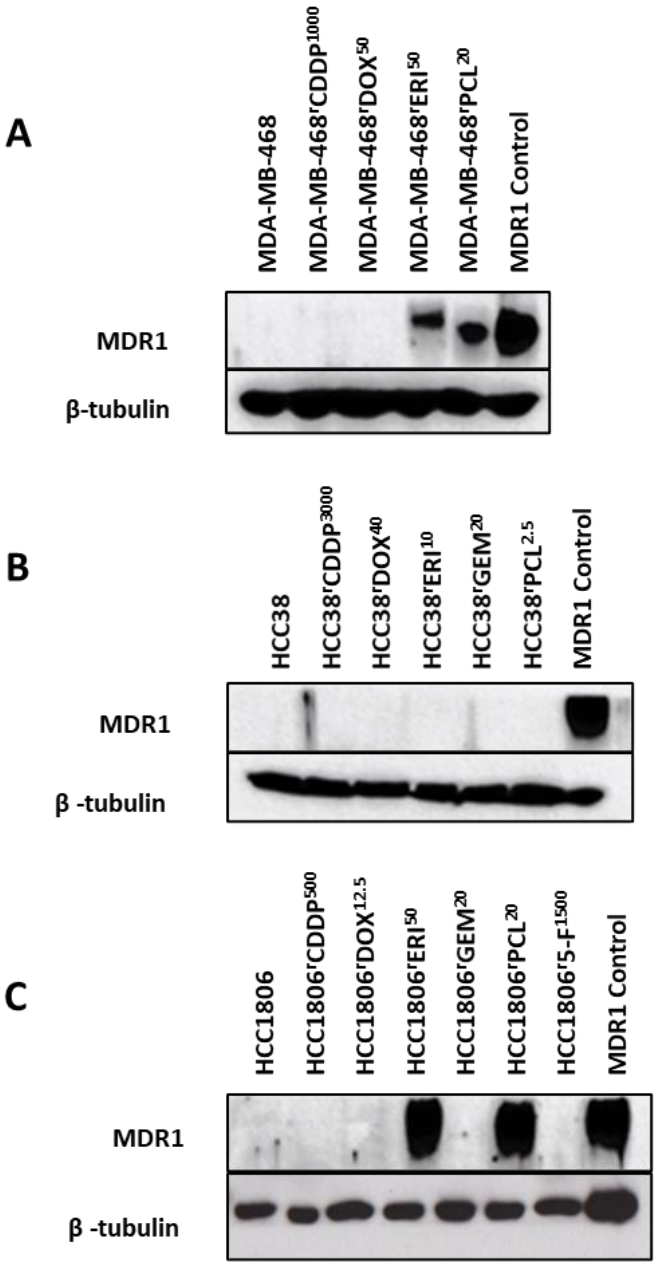


Figure 3.17. Analysis of MDR1 (p-glycoprotein) protein levels in both chemo-naive and chemo-resistant TNBC cell
Cell lysates were prepared and analysed by western blotting to determine the expression levels of MDR1 in A) MDA-MB-468, B) HCC38, C) HCC1806 chemo-naive and chemo-resistant TNBC cell lines. UKF-NB-3^{r}VCR^{10} from RCCL collection, was used as positive control. Blots were probed with antibody against MDR1, with β-tubulin included as a loading control. Blots are representative of ≥3 independent experiments.

3.3 Discussion

This chapter aimed to characterise a panel of chemo-naive and chemo-resistant TNBC cell lines to a panel of chemotherapeutic agents often used for the treatment of TNBC. To this end, eighteen cell lines were selected from the RCCL collection which included; three chemo-naive TNBC cell lines and fifteen chemo-resistant sublines.

Using microscopy, the morphology of each of the cell lines was examined, as this can be a simple technique to identify resistance mechanisms and phenotypes such as epithelial-to-mesenchymal transition (EMT), which has been implicated in drug resistance (Beaufort et al., 2014; Du and Shim, 2016). Each of the chemo-naive TNBC cell lines demonstrated different cell morphology which is reflective of the heterogeneous nature of TNBC. A small number of observations were made when comparing the chemo-naive to the chemo-resistant sub-types. MDA-MB-468^{r}PCL^{20} was found to be mesenchymal in shape compared to MDA-MB-468, and, when considering the growth rates, was also found to grow faster. MDA-MB-468^{r}ERI^{50} was found to have a mix population, with morphology likened to MDA-MB-468 and also MDA-MB-468^{r}PCL^{20}. Cells which demonstrate an EMT morphology are similar to cancer stems cells (CSCs) which have been implicated in the development of drug resistance, and also demonstrate enhanced invasive and migratory phenotypes (Phi *et al.*, 2018). Importantly, both MDA-MB-468^{r}PCL^{20} and MDA-MB-468^{r}ERI^{50} had elevated MDR1 expression. CSCs are known to express ABC transporters, including MDR1, which aids in the EMT cell progression, drug efflux and drug resistance (Moitra, 2015; Sugano et al., 2015). It could be speculated that MDA-MB-468^{r}PCL^{20} and MDA-MB-468^{r}ERI^{50} are displaying a CSCs phenotype, which is contributing to the drug resistant phenotype.

MDR1 was also found to be upregulated in HCC1806^{r}ERI^{50}, and HCC1806^{r}PCL^{20}, although no significant change in morphology or growth rate, compared to HCC1806, was observed. Each of the four cell lines with increased expression of MDR1 also demonstrated resistance/cross-resistance to doxorubicin, eribulin and paclitaxel; known substrates of MDR1 (Chen et al., 2016; Oba, Izumi and Ito, 2016). Whilst this suggests that the cross-resistance in these cell lines to MDR1 substrates is at least in part caused by increased MDR1 levels, further evidence would be required to confirm this. One way in which this could be investigated further, is through combinational assays in which the substrate of MDR1 is used in combination with zosuquidar, a potent inhibitor of MDR1 (Nanayakkara et al., 2018). It was also observed that HCC1806^{r}ERI^{50} demonstrated both an elevated expression of MDR1 as well as an acquired vulnerability to gemcitabine. Bergman et al., 2003 has previously shown that MDR1 overexpression can cause cellular stress, which can result in increased

metabolism and sensitivity to gemcitabine, while treatment with verapamil reverses the sensitivity. This suggests a direct relation between membrane efflux pumps and gemcitabine sensitivity. The observation in the HCC1806^{r}ERI^{50} cell line agrees with the data presented by Bergman et al., 2003, although it is noted that acquired vulnerability is not seen in the three other MDR1 expressing cell lines.

Although increased expression of MDR1 was not observed in the other eleven chemo-resistant cell lines, it cannot be ruled out that increased efflux of the chemotherapeutic agents by another transporter is not occurring. For example, other multi-drug resistant transporters include, multidrug resistance associated protein (MPR1/ABCC1) and the breast cancer resistance protein (BCRP/ABCCG2) which could be driving the resistance phenotype (Natarajan et al., 2012; Lu, Pokharel and Bebawy, 2015).

A pattern of cross-resistance was identified with eribulin and paclitaxel in the eribulin and paclitaxel resistant cell lines. Each of the eribulin resistant cell lines demonstrated cross-resistance to paclitaxel, and each of the paclitaxel resistant cell lines demonstrated cross-resistance to eribulin, with the exception of HCC38^{r}PCL$^{2.5}$, which although had a RF of 11.9, was not found to show statistically significant difference compared to HCC38. Arguably, this cross-resistance pattern could be a result of increased expression of MDR1 in four of the cases, however, despite not showing elevated expression of MDR1, HCC38^{r}ERI^{10} demonstrated cross resistance to paclitaxel. As previously described (section 1.2.4), whilst eribulin and paclitaxel are both tubulin-binding agents targeting β-tubulin, the binding sites and the mechanism of actions differ. However, the result of both of the drugs actions commonly result in mitotic arrest, and if the damage cannot be resolved leads to apoptosis (Weaver, 2014; O'Shaughnessy, Kaklamani and Kalinsky, 2019). It could be speculated that an additional common resistance mechanism exists for eribulin and paclitaxel, such as dysregulation of mitotic arrest, however, none as of yet have been described in literature. Of the four cell lines, which demonstrate increased levels of MDR1, it will be important to investigate if there is still a level of resistance to eribulin and paclitaxel in the presence of an MDR1 inhibitor.

It was also observed that both of the gemcitabine resistant cell lines, HCC38^{r}GEM^{20} and HCC1806^{r}GEM^{20}, demonstrated cross-resistance to 5-Fluorouracil. The only cell line to demonstrate cross-resistance to gemcitabine was HCC38^{r}CDDP^{3000}, which also demonstrated cross-resistance to 5-Fluorouracil. Gemcitabine and 5-fluorouracil often still remain as first-line therapy for the treatment of pancreatic cancer, and as such, many groups have researched cell lines resistant to both of these drugs. In 2017, Ghadban et al., showed that heat shock protein 90 (Hsp90) inhibition

promotes cancer cell apoptosis in pancreatic cell lines that are resistant to both gemcitabine and 5-Fluorouracil. It has been established that Hsp90 stabilises proteins of the PI3K/AKT signalling pathways, and therefore growth factor signalling may be reliant on functional Hsp90 (Solit et al., 2002). Further investigation into dysregulation of Hsp90 or PI3K/AKT pathways could show a common drug resistance mechanism for gemcitabine and 5-Fluorouracil in these cell lines.

Whilst cross-resistance was seen in the majority of the chemo-resistant cell lines, there was no observed cross-resistance or acquired vulnerability to any of the chemotherapeutic agents in HCC1806^{r}5-F^{1500}, making this cell line distinct to the dataset. This could suggest that the resistance mechanism in this cell line is unique to the uptake or action of 5-Fluorouracil.

In summary, this chapter has addressed changes in morphology and growth between the chemo-naive and chemo-resistant cell lines, and confirmed resistance to the chemotherapeutic agents in the chemo-resistant cell lines. Investigation of expression of MDR1 identified increased expression in four cell lines, and drug profiling identified patterns of cross-resistance. The following chapter aims to identify patterns of cross-resistance or acquired vulnerability to inhibitors which target the DNA damage response and consider these as an option for second line therapy after chemo-resistance has emerged.

Chapter 4

Investigating the response of chemo-resistant TNBC to inhibitors of the DNA damage response pathway

4. Investigating the response of chemo-resistant TNBC to inhibitors of the DNA damage response pathway

4.1 Introduction

Cancer cells can acquire resistance to chemotherapeutic agents by altering the binding properties of drugs, detoxification or drug efflux, or by altering the cell survival mechanisms such as cell cycle arrest, apoptosis and DNA damage response and repair (DDR) pathways (see section 1.3) (Cree and Charlton, 2017). Many chemotherapy agents induce DNA damage which requires the activation of the DDR. The DDR is a process by which DNA damage is effectively and accurately repaired to maintain genomic stability, and can be described through several distinct components which ultimately determine the outcome of the cell (see section 1.4)(Blackford and Jackson, 2017). These DDR pathways are often adapted, either through changes of expression or mutations in genes associated with these pathways, which ultimately result in dysregulation and chemo-resistance (Housman *et al.*, 2014).

One mechanism of acquired resistance to chemotherapy agents is to upregulate DDR thereby increasing DNA repair. For example, cisplatin is known to induce interstrand cross-links (ICL) in the DNA, which activates the DDR to remove these bulky lesions. Enhanced capacity to remove cisplatin induced ICLs have been observed in cisplatin resistant ovarian cancer cells (Wynne *et al.*, 2007). Furthermore, increased expression of Aurora A, has been seen to promote upregulation of ATM/CHK2 mediated DNA repair in cisplatin exposed cells. This results in an increase in the rate of cell proliferation and cisplatin resistance (H. Sun *et al.*, 2014).

It has also been shown that mutations of genes in the DDR can lead to chemo-resistance. Many studies have reported that cells deficient in MMR components, particularly MLH1 and MSH2, demonstrate resistance to 5-Fluorouracil in colon cancer. It is thought that induction of MMR dependent recognition of 5-Fluorouracil DNA damage initiates apoptosis (Carethers *et al.*, 1999; M. Meyers *et al.*, 2001; Meyers *et al.*, 2005).Therefore, reduction of MMR activation to 5-Fluorouracil damage subsequently reduces apoptosis activation in the cells. Similar observations have been made in respect to cisplatin resistance (Papouli, Cejka and Jiricny, 2004).

Literature shows that chemo-resistant cells can be reliant on the dysregulation of DDR pathways as a mechanism of resistance, making these pathways an attractive therapeutic target (see section 1.3.2.4) (Goldstein and Kastan, 2015). A new class of drugs have been developed which target the DDR pathways through small molecule inhibition. Many clinical trials are underway to investigate

these compounds both as monotherapies, and also as combinational therapies (Brandsma *et al.*, 2017).

The DDR inhibitor monotherapies have been found to work particularly well with cancers that already demonstrate a mutational defect in a DDR gene. These pre-existing defects in the DDR can make the cells reliant on a secondary repair pathway, of which therapeutic inhibition results in synthetic lethality (see section 1.4.2). One of the most recent successful monotherapy treatments to be registered as an approved drug is olaparib, a PARP inhibitor, for the treatment of ovarian cancer. Inhibition of PARP in cell lines with a loss-of-function germline mutation in *BRCA1* or *BRCA2*, proteins that are critical for the repair of double strand DNA (dsDNA) breaks, result in a synthetic lethal relationship (Lord and Ashworth, 2017). It is evident that more biomarkers of sensitivity need to be identified for such monotherapies to be successful.

Research into combined use of DDR inhibitors and chemotherapeutic agents for the treatment of cancer are underway. The combinational therapy had been found to not only enhance and augment the chemotherapeutic agents, but also overcome drug resistance. One such example, is the combination of cisplatin with inhibition of CHK1. A recent publication has found that the CHK1 inhibitors, Prexasertib and AZD6672, in combination with cisplatin, enhanced cisplatin antitumor activity and overcame cisplatin resistance in small cell lung cancer (SCLC) cells (Hsu *et al.*, 2019). Another example, is the phase II clinical study underway in advanced solid tumours, with the combination of the CHK1 inhibitor, SRA737, with gemcitabine (NCT02797977).

Defects in genes associated to DDR pathways, that may have arisen in the process of acquiring drug resistance, may result in a synthetic lethal relationship when a second component on the same or alternate DDR pathway is inhibited with targeted drugs. This highlights the potential of DDR inhibitors as a second line therapy option. Screening of pre-clinical chemo-resistant cell models with inhibitors of the DDR may identify patterns of cross-resistance or acquired vulnerability, which can be used to predict the use of these drugs as a second line therapy.

In Chapter 3, the morphology and cross-resistance to chemotherapy agents in both the chemo-naive and chemo-resistant TNBC cell lines were investigated to identify patterns of cross- resistance and vulnerability. This chapter aims to assess the option of DDR inhibitors as a second-line therapy, in these chemo-resistant settings. To achieve this, both the chemo-naive and chemo-resistant TNBC cell lines will be screened against a panel of DDR inhibitors to identify patterns of cross-resistance and acquired vulnerability.

4.2 Results

4.2.1 Drug profiling of DNA damage response inhibitors in TNBC cell lines

In order to determine if any of the chemo-resistant TNBC cell lines demonstrated patterns of cross-resistance or acquired vulnerability to inhibitors of the DDR, each of the cell lines was treated with the inhibitors in Table 4.1. Dose response curves were generated to a serial dilution of the inhibitors using a 120-hour MTT assay, and the GI_{50} and the RF determined, as previously described in Chapter 3. As stated in Chapter 3, cross-resistance is defined when the calculated RF ≥ 2, and acquired vulnerability when the calculated RF is ≤ 0.5, and both demonstrated statistical significance using the student's t-test ($p \leq 0.05$). Calculated GI_{50} and RF's for each cell line and drug combination are in Appendix A3. Cells were plated at their optimal seeding density as established in Chapter 3.

Table 4.1. Panel of inhibitors targeting the DNA damage response and repair pathways

Drug name	Alternative names	Drug target
MK-8776	SCH900776	CHK1
SRA37	CCT245737	CHK1
Rabusertib	LY2603618	CHK1
Prexasertib	LY2606368	CHK1/2
Ceralasertib	AZD6738	ATR
Berzosertib	VE-822, VX-970	ATR
AZD0156	-	ATM
CCT241533	-	CHK2
Adavosertib	MK-1775, AZD1775	Wee1
Olaparib	AZD2281	PARP
Rucaparib	-	PARP
BO2	-	RAD51
BI2536	-	PLK1
SBE13	SBE 13 HCL	PLK1/2
Alisertib	MLN8237	AURKA
Tozasertib	VX-680	AURK(pan)

4.2.1.1 Drug profiling of CHK1 inhibitors; Rabusertib, MK-8776, SRA737, Prexasertib

The response of each of the TNBC cell lines was determined, when treated with the four structurally different CHK1 inhibitors; rabusertib, MK-8776, SRA737 and prexasertib (Figure 4.1). Each of the CHK1 inhibitors have demonstrated to be potent inhibitors of CHK1, and are all ATP-competitive (Montano *et al.*, 2012; King *et al.*, 2014, 2015; Walton *et al.*, 2015). It was interesting to note that the same patterns of cross-resistance in the cell lines were not seen with each of the four CHK1 inhibitors. This could be a result of differences in the drugs specificity, efficacy, and drug influx or efflux.

It was noted that both HCC38 and HCC1806 are sensitive to each of the CHK1 inhibitors, with GI_{50} values in the range of and 0.000834 µM – 0.122 µM and 0.0043 µM – 0.465 µM respectively, compared to the MDA-MB-468 which demonstrated higher GI_{50} values in the range of 0.0233 µM – 4.3 µM. This suggests intrinsic resistance to the CHK1 inhibitors in the MDA-MB-468, compared to the two other chemo-naive cell lines.

Only HCC38^{r}CDDP^{3000} and HCC1806^{r}DOX$^{12.5}$ demonstrated cross-resistance to all four of the CHK1 inhibitors with RF's in the range of 8.42 – 15.34 and 2.15 – 3.67 respectively. HCC38^{r}CDDP^{3000} had a higher RF of 15.34 when treated with rabusertib, compared to the other cell lines which demonstrated an RF < 4. MDA-MB-468^{r}PCL^{20} demonstrated acquired vulnerability to rabusertib (RF = 0.5), but cross-resistance to MK-8776 (RF = 3.84). Acquired vulnerability was also seen in MDA-MB-468^{r}CDDP^{1000} to prexasertib (RF = 0.23). Both MDA-MB-468^{r}DOX^{50} and MDA-MB-468^{r}ERI^{50} demonstrated cross-resistance to SRA737 (RF = 2.57, 2.32 respectively), but only MDA-MB-468^{r}ERI^{50} demonstrated cross-resistance to prexasertib (RF = 8.99). HCC38^{r}GEM^{20} demonstrated cross-resistance to each of the CHK1 inhibitors except MK-8776. HCC1806^{r}ERI^{50} and HCC1806^{r}PCL^{20}, demonstrated cross-resistance to both rabusertib (RF = 3.4, 2.87 respectively) and MK-8776 (RF = 3.13, 4.62 respectively), but not to the other two CHK1 inhibitors. HCC1806^{r}GEM^{20} and HCC1806^{r}PCL^{20} also demonstrated cross-resistance to MK-8776 (RF = 3.17, 4.62 respectively). The rest of the chemo-resistant cell lines showed no statistically significant fold change to the CHK1 inhibitors.

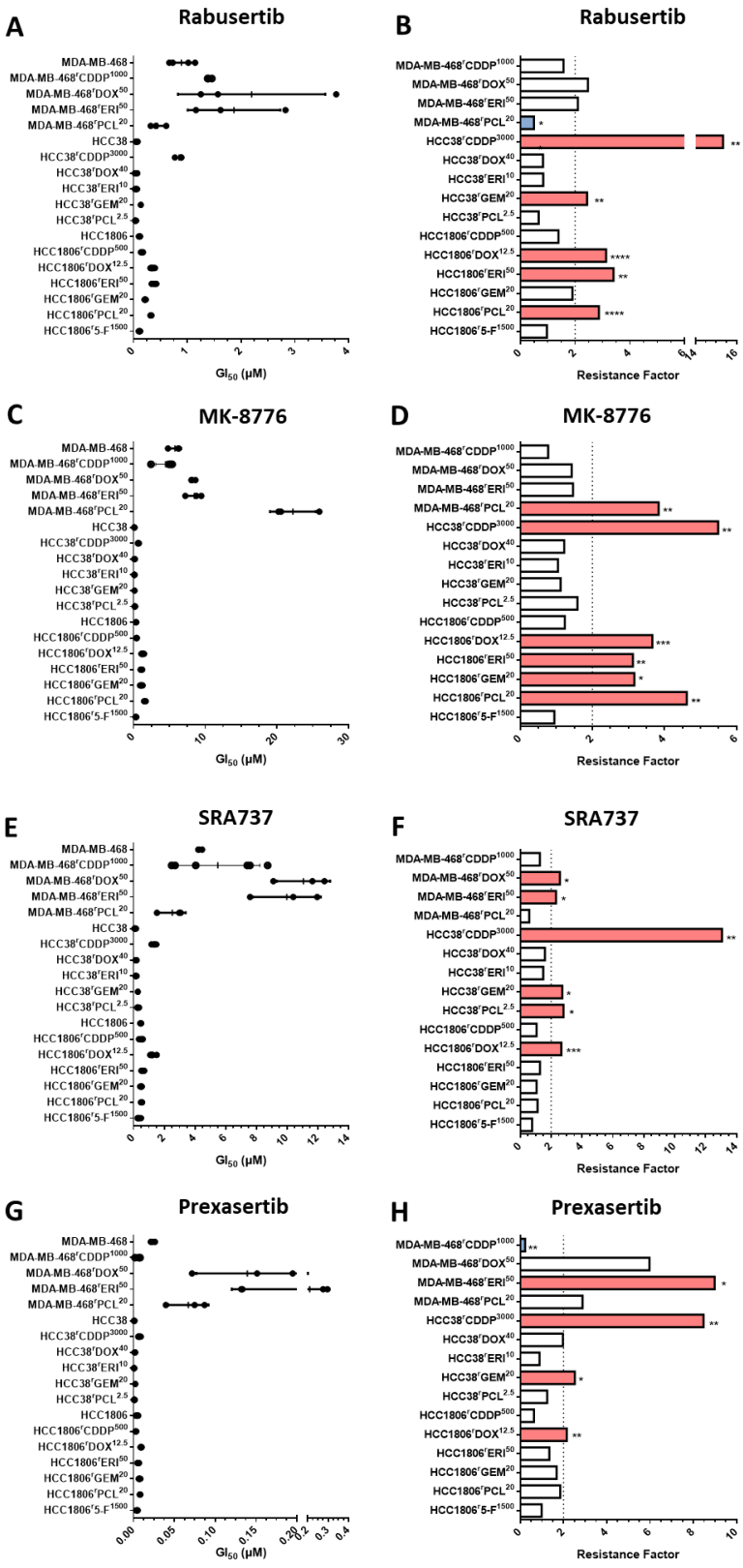


Figure 4.1 GI_{50} value determinations and resistance factors upon treatment with CHK1 inhibitors in the TNBC cell lines
The response of TNBC cell lines to stated CHK1 inhibitors was determined by 120-hour MTT assay and analysed using GraphPad Prism 6. A, C, E, G) Determined GI_{50} concentrations upon treatment with stated CHK1 inhibitors. Dot represents a single biological repeat from ≥3 independent experiments, horizontal line shows standard deviation and vertical line shows the mean. B, D, F, H) Calculated resistance factor (RF) to stated CHK1 inhibitors. Dotted line indicates two-fold RF. Statistical significance calculated using student's t-test, * $P \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$, **** $p \leq 0.00001$. Colour indicates resistance status: red = resistant (RF >2, and ≥ * p), blue = acquired vulnerability (RF < 0.5 and ≥ * p), white = no change. Numerical data is presented in Appendix A3.

4.2.1.2 Drug profiling of ATR inhibitors; Ceralasertib, Berzosertib, and ATM inhibitor; AZD0156

The response of each of the TNBC cell lines was determined when treated with two structurally different ATR inhibitors; ceralasertib, berzosertib, as well as the ATM inhibitor; AZD0156 (Figure 4.2). ATR is found directly upstream of CHK1, and its activating phosphorylation on CHK1 occurs in response to DNA damage. Ceralasertib is a sulfoximine morpholinopyrimidine ATR inhibitor, and berzosertib is an aminopyrazine ATR inhibitor. Both are ATP-competitive and selective for ATR inhibition (Vendetti *et al.*, 2015; Foote *et al.*, 2018; Gorecki *et al.*, 2020). ATM is found directly upstream of CHK2, but has also been found to be implicated in phosphorylation and activation of CHK1. AZD0156 is a potent, selective inhibitor of ATM (Pike *et al.*, 2018).

It was observed that HCC38 and HCC1806 are more sensitive to both of the ATR inhibitors (Figure 4.2 A-D), with GI_{50} values in the range of and 0.0631 µM – 0.216 µM and 0.0.888 µM – 0.34 µM respectively, compared to MDA-MB-468, which has higher GI_{50} values in the range of 0.1793 µM – 0.651 µM. This pattern of sensitivity is the same as the response to CHK1 inhibitors (section 4.2.1.1). Each of the chemo-naive cell lines are more sensitive to berzosertib than to ceralasertib. The cell lines; HCC38^{r}CDDP^{3000}, HCC1806^{r}DOX$^{12.5}$, HCC1806^{r}GEM^{20} and HCC1806^{r}PCL^{20} all show cross-resistance to both ceralasertib (RF = 7.32, 3.31, 3.70 and 6.96 respectively) and berzosertib (RF = 4.56, 2.61, 2.43, 2.78). MDA-MB-468^{r}DOX^{50}, HCC38^{r}GEM^{20}, HCC1806^{r}CDDP^{500}, HCC1806^{r}ERI^{50} all demonstrate cross-resistance to ceralasertib (RF = 2.17, 3.99, 1.99, and 2.91 respectively), but not to berzosertib. It was noted that HCC38^{r}CDDP^{3000} and HCC1806^{r}DOX$^{12.5}$, which demonstrated cross-resistance to all four of the CHK1 inhibitors, both demonstrated cross-resistance to the ATR inhibitors (section 4.2.1.1). The rest of the chemo-resistant cell lines showed no statistically significant fold change to the ATR inhibitors.

When the TNBC cells were treated with the ATM inhibitor; AZD0156 (Figure 4.2 E-F), MDA-MB-468 and HCC1806 had GI_{50} values of 1.6 µM and 0.637 µM respectively, whilst HCC38 demonstrated a higher GI_{50} value of 4.01 µM. MDA-MB-468^{r}CDDP^{1000}, MDA-MB-468^{r}DOX^{50}, MDA-MB-468^{r}ERI^{50} and HCC38^{r}CDDP^{3000} all showed acquired vulnerability to AZD0156 (RF = 0.40, 0.43, 0.18, 0.35 respectively). HCC1806^{r}DOX$^{12.5}$, HCC1806^{r}ERI^{50}, HCC1806^{r}GEM^{20} and HCC1806^{r}PCL^{20} all showed cross-resistance to treatment with AZD0156 (RF = 3.66, 2.16, 2.93, 7.70 respectively). The rest of the chemo-resistant cell lines showed no statistically significant fold change to the ATM inhibitors.

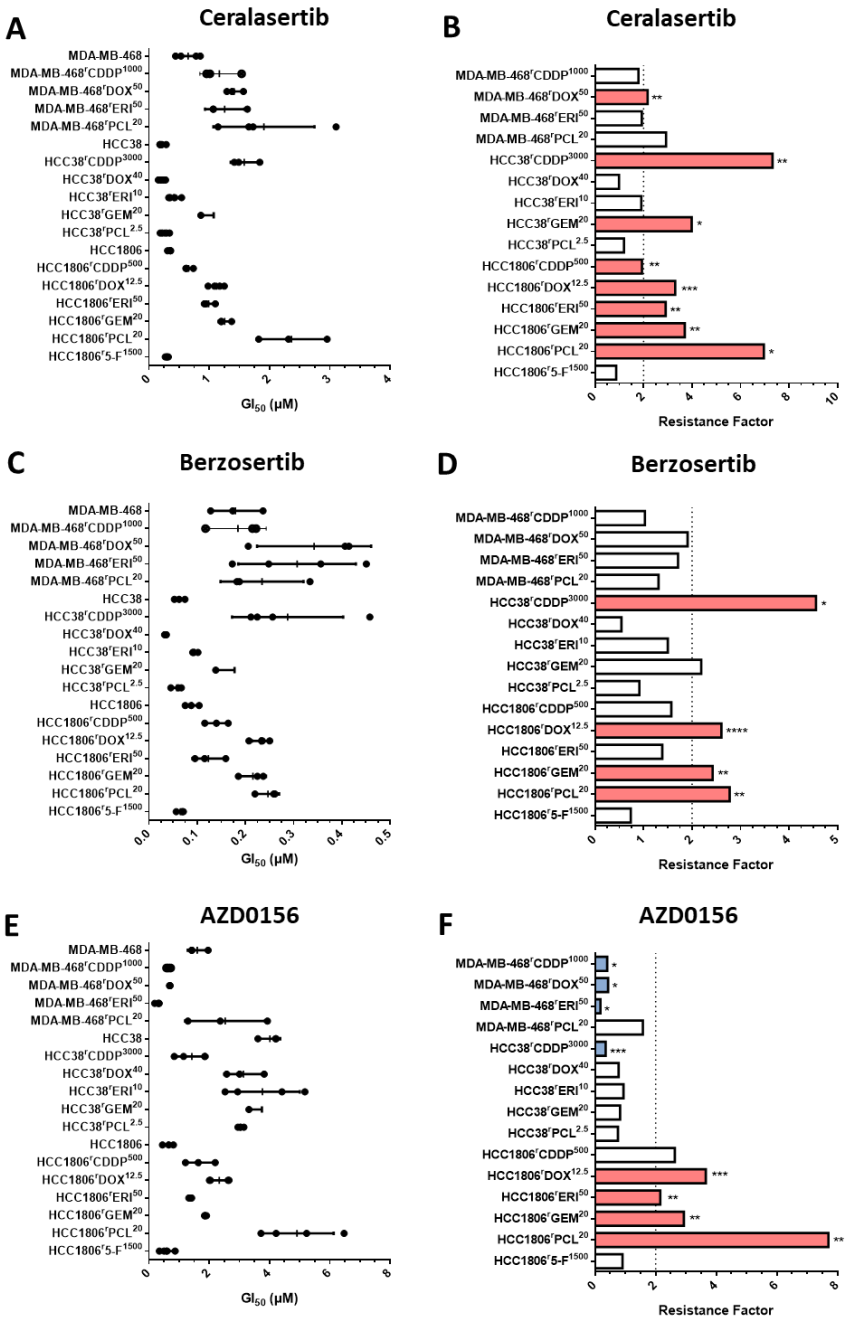


Figure 4.2 GI_{50} value determinations and resistance factors upon treatment with ATR/ATM inhibitors in the TNBC cell lines
The response of TNBC cell lines to stated ATR/ATM inhibitors was determined by 120-hour MTT assay and analysed using GraphPad Prism 6. A, C, E) Determined GI_{50} concentrations upon treatment with stated ATR/ATM inhibitors. Dot represents a single biological repeat from ≥3 independent experiments, horizontal line shows standard deviation and vertical line shows the mean. B, D, F) Calculated resistance factor (RF) to stated ATR/ATM inhibitors. Dotted line indicates two-fold RF. Statistical significance calculated using student's t-test, * P ≤ 0.05, ** p ≤ 0.01, *** p ≤ 0.001, **** p ≤ 0.00001. Colour indicates resistance status: red = resistant (RF >2, and ≥ * p), blue = acquired vulnerability (RF < 0.5 and ≥ * p), white = no change. Numerical data is presented in Appendix A3.

4.2.1.3 Drug profiling of WEE1 inhibitors; Adavosertib, and CHK2 inhibitor; CCT241533

WEE1 is a serine/threonine kinase known to be downstream of CHK1, with important roles in cell cycle regulation, such as its key role in inhibitory phosphorylation of CDK1 (Russell and Nurse, 1987). Adavosertib has been shown to be a potent ATP-competitive selective inhibitor of WEE1 (Guertin *et al.*, 2013; Yuan *et al.*, 2018). Upon treatment with adavosertib, HCC38 and HCC1806 show GI_{50} values of 0.0489 µM and 0.0651 µM, whilst MDA-MB-468 has a much higher GI_{50} value of 0.207 µM, approximately 3-4-fold greater (Figure 4.3 A-B). This pattern of sensitivity is the same as the response to CHK1 inhibitors (section 4.2.1.1), and the ATR inhibitors (section 4.2.1.2). MDA-MB-468^{r}ERI^{50}, HCC38^{r}CDDP^{3000}, HCC1806^{r}DOX$^{12.5}$, HCC1806^{r}ERI^{50} and HCC1806^{r}PCL^{20} all demonstrate cross-resistance upon treatment with adavosertib (RF = 5.40, 4.88, 2.16, 3.30, 2.92 respectively). MDA-MB-468^{r}CDDP^{1000} demonstrated acquired vulnerability to adavosertib with an RF = 0.35. It was noted that also MDA-MB-468^{r}CDDP^{1000} also demonstrated acquired vulnerability to prexasertib and AZD0156, and showed no cross-resistance to any of the CHK1 or ATR inhibitors (section 4.2.1.1, 4.2.1.2). HCC38^{r}CDDP^{3000} and HCC1806^{r}DOX$^{12.5}$, which demonstrated cross-resistance to all four of the CHK1 inhibitors, and both ATR inhibitors, also demonstrated cross-resistance to Adavosertib (section 4.2.1.1, 4.2.1.2). The rest of the chemo-resistant cell lines showed no statistically significant fold change to adavosertib.

CHK2, is a serine/threonine kinase known to be phosphorylated by ATM upon dsDNA damage. CCT241533 is an ATP-competitive, selective inhibitor of CHK2 (Anderson *et al.*, 2011b). Upon treatment with CCT241533, it was observed that the chemo-naive cell lines; MDA-MB-468, HCC38, HCC1806, had very similar GI_{50} values of 1.39µM, 1.03 µM and 2.27 µM respectively (Figure 4.3 C-D). The chemo-resistant cell lines demonstrated no statistically significant fold change in response to CCT241533, with an exception of MDA-MB-468^{r}PCL^{20}, which demonstrated cross-resistance (RF = 2.42), and HCC1806^{r}5-F^{1500}, which demonstrated acquired vulnerability (RF = 0.422).

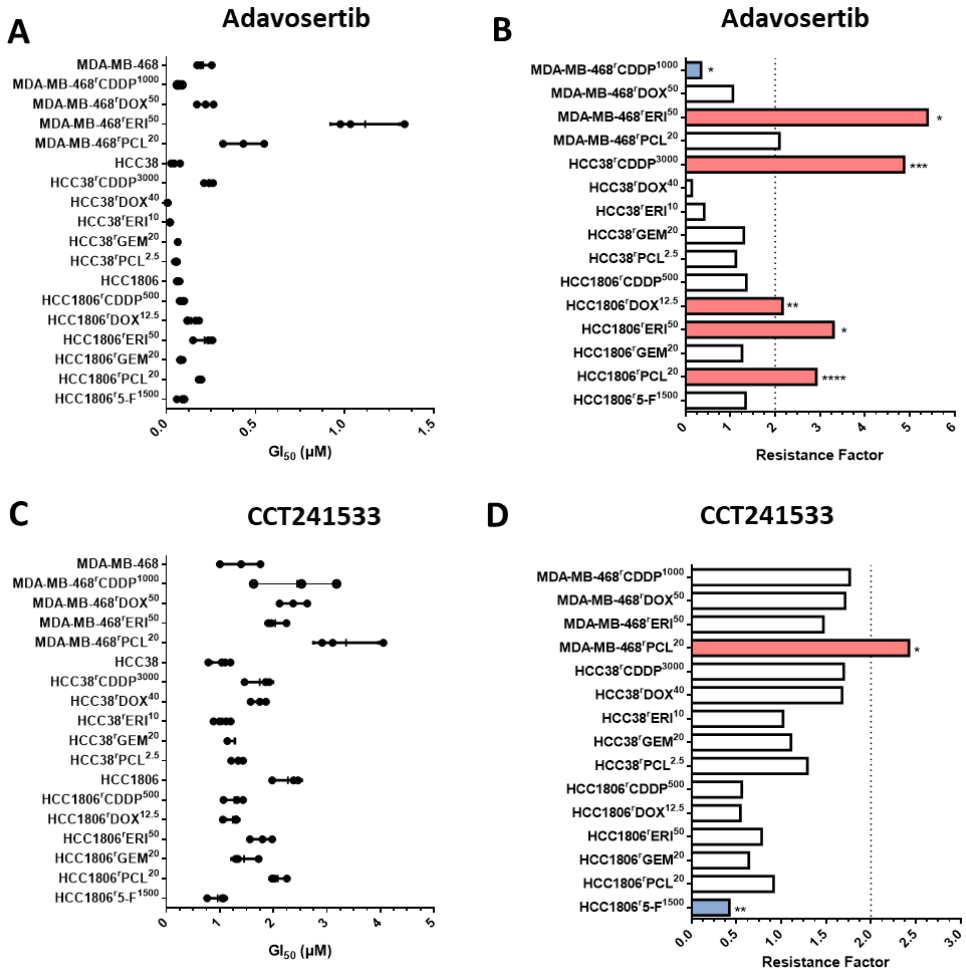


Figure 4.3 GI_{50} value determinations and resistance factors upon treatment with WEE1/CHK2 inhibitors in the TNBC cell lines
The response of TNBC cell lines to stated WEE1/CHK2 inhibitors was determined by 120-hour MTT assay and analysed using GraphPad Prism 6. A, C, E, G) Determined GI_{50} concentrations upon treatment with stated WEE1/CHK2 inhibitors. Dot represents a single biological repeat from ≥3 independent experiments, horizontal line shows standard deviation and vertical line shows the mean. B, D, F, H) Calculated resistance factor (RF) to stated WEE1/CHK2 inhibitors. Dotted line indicates two-fold RF. Statistical significance calculated using student's t-test, * P ≤ 0.05, ** p ≤ 0.01, *** p ≤ 0.001, **** p ≤ 0.00001. Colour indicates resistance status: red = resistant (RF >2, and ≥ * p), blue = acquired vulnerability (RF < 0.5 and ≥ * p), white = no change. Numerical data is presented in Appendix A3.

4.2.1.4 Drug profiling of Rad51 inhibitor; B02 and PARP inhibitors; Olaparib, Rucaparib

RAD51 is a key enzyme which plays an important role in homologous strand exchange, a crucial step in HRR (section 1.4.1.2). B02 has been identified as a RAD51 specific inhibitor, inhibiting the development of RAD51 foci formation (Huang and Mazin, 2011; Huang *et al.*, 2012). Upon treatment with B02, MDA-MB-468 and HCC1806 show a similar sensitivity with GI_{50} values of 4.5 µM and 5.56 µM respectively. A higher GI_{50} was observed in HCC38 with a value of 11.5 µM, approximately two-fold greater than MDA-MB-468 and HCC1806 (Figure 4.4A). Only HCC1806^{r}PCL^{20} demonstrated cross-resistance to B02 with an RF value of 2.64 (Figure 4.4A, B). No other statistically significant fold changes were identified when the chemo-resistant cell lines were treated with B02.

PARP proteins are associated with the repair of dsDNA break, and the polymerised chains can act as a signal for other DNA repair proteins (section 1.4.1.2). Both PARP inhibitors target the PARP proteins but are structurally different, with olaparib inhibiting PARP 1/2/3, whilst rucaparib is more promiscuous in its binding to the PARP proteins (Thomas *et al.*, 2007; Wahlberg *et al.*, 2012). The response of each of the TNBC cell lines was determined, when treated with the two structurally different PARP inhibitors; olaparib and rucaparib (Figure 4.4C-F). MDA-MB-468 is the most sensitive out of the chemo-naive cell line with a GI_{50} value of 1.02 µM and 1.77 µM to olaparib and rucaparib respectively. Higher GI_{50} values are seen in HCC38 and HCC1806 which demonstrate 4.36 µM and 3.7 µM to olaparib respectively, and 5.39 µM and 5.36 µM to rucaparib respectively. Each cell line that demonstrates cross-resistance to rucaparib (MDA-MB-468^{r}CDDP^{1000}; RF = 4.27, MDA-MB-468^{r}ERI^{50}; RF = 5.18 , MDA-MB-468^{r}PCL^{20}; RF = 7.97, HCC38^{r}CDDP^{3000}; RF = 3.19, HCC38^{r}DOX^{40}; RF = 4.10, HCC38^{r}PCL$^{2.5}$; RF =2.57 , HCC1806^{r}CDDP^{500}; RF = 2.36, HCC1806^{r}DOX$^{12.5}$; RF =3.72, HCC1806^{r}ERI^{50}; RF =2.74, HCC1806^{r}PCL^{20}; RF = 4.40) also demonstrates cross-resistance to olaparib (MDA-MB-468^{r}CDDP^{1000}; RF = 7.14, MDA-MB-468^{r}ERI^{50}; RF = 21.4, MDA-MB-468^{r}PCL^{20}; RF = 21.50, HCC38^{r}CDDP^{3000}; RF = 3.04, HCC38^{r}DOX^{40}; RF = 3.812, HCC38^{r}PCL$^{2.5}$; RF = 3.7, HCC1806^{r}CDDP^{500}; RF = 6.42, HCC1806^{r}DOX$^{12.5}$; RF = 4.06, HCC1806^{r}ERI^{50}; RF = 3.51, HCC1806^{r}PCL^{20}; RF = 9.35). HCC1806^{r}GEM^{20} and MDA-MB-468^{r}DOX^{50} only demonstrate cross-resistance to olaparib and not to rucaparib (RF = 2.11, 4.82 respectively). HCC38^{r}ERI^{10}, HCC38^{r}GEM^{20} and HCC1806^{r}5-F^{1500} show no statistically significant change to both of the PARP inhibitors.

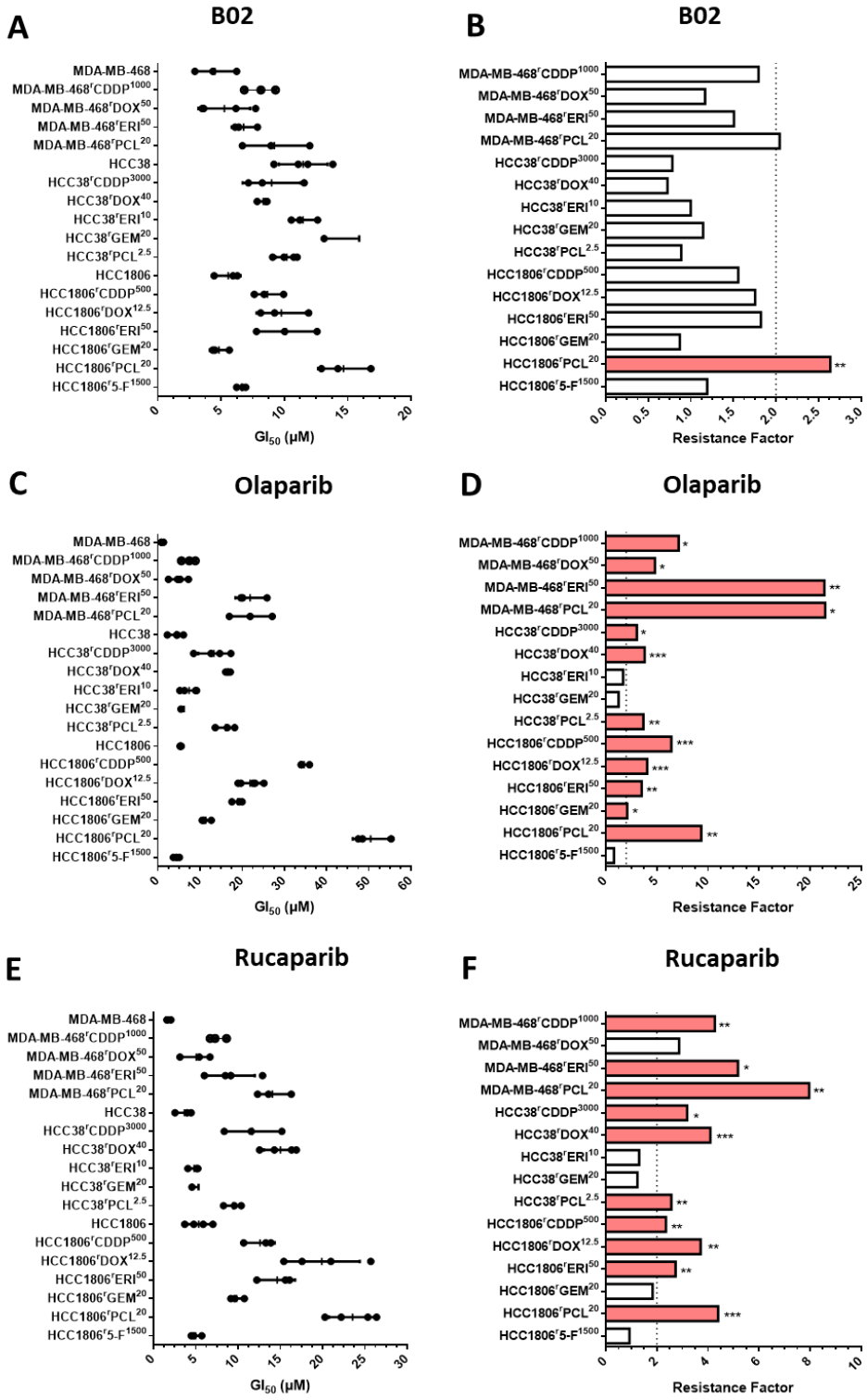


Figure 4.4 GI_{50} value determinations and resistance factors upon treatment with RAD51/PARP inhibitors in the TNBC cell lines
The response of TNBC cell lines to stated RAD51/PARP inhibitors was determined by 120-hour MTT assay and analysed using GraphPad Prism 6. A, C, E, G) Determined GI_{50} concentrations upon treatment with stated RAD51/PARP inhibitors. Dot represents a single biological repeat from ≥3 independent experiments, horizontal line shows standard deviation and vertical line shows the mean. B, D, F, H) Calculated resistance factor (RF) to stated RAD51/PARP inhibitors. Dotted line indicates two-fold RF. Statistical significance calculated using student's t-test, * P ≤ 0.05, ** p ≤ 0.01, *** p ≤ 0.001, **** p ≤ 0.00001. Colour indicates resistance status: red = resistant (RF >2, and ≥ * p), blue = acquired vulnerability (RF < 0.5 and ≥ * p), white = no change. Numerical data is presented in Appendix A3.

4.2.1.5 Drug profiling of Polo kinase inhibitors; BI2536, SBE13 and Aurora Kinase inhibitors; Alisertib, Tozasertib

The final four drugs analysed in this drug profiling, are second-generation mitotic drugs, targeting polo kinases (PLK) and aurora kinases (AURK). Both PLK, and AURK have integral roles in cell cycle progression, mitotic spindle assembly and are activated during the DDR (section 1.4.2). BI2536 is a potent, selective PLK1 inhibitor, but has some weaker activity against PLK2, PLK3, and bromodomain 4 (BRD4) (Steegmaier *et al.*, 2007; Gohda *et al.*, 2018). SBE13 inhibits both PLK1 and PLK2, with weaker activity against PLK3 and tyrosine-protein kinase BTK (Kumar and Kim, 2015).

When treated with BI2536, MDA-MB-468, HCC38 and HCC1806 showed little difference in sensitivity with GI_{50} values of 2.01 nM, 3.2 nM and 1.16 nM. A greater difference was seen when treated with SBE13. MDA-MB-468 was more sensitive, with a GI_{50} value of 4.97 µM, compared to HCC38 and HCC1806 which showed similar GI_{50} values of 47.6 µM and 31.6 µM respectively. Overall, the chemo-naive cell lines demonstrated more sensitivity to BI2536 compared to SBE13 (Figure 4.5A, C). Cross-resistance was seen in MDA-MB-468^{r}ERI^{50}, MDA-MB-468^{r}PCL^{20}, HCC38^{r}ERI^{10}, HCC1806^{r}DOX$^{12.5}$, HCC1806^{r}ERI^{50}, HCC1806^{r}GEM^{20} and HCC1806^{r}PCL^{20} upon treatment with BI2536 (RF = 13.68, 24.17, 9665.57, 3.24, 1.99, 15.03 respectively). Only MDA-MB-468^{r}PCL^{20} demonstrated cross-resistance to SBE13 (RF = 4.48), and HCC38^{r}CDDP^{3000} and HCC1806^{r}5-F^{1500} demonstrated an acquired vulnerability to SBE 13 (RF = 0.366, 0.45 respectively) (Figure 4.5 B, D). The rest of the chemo-resistant cell lines showed no statistically significant fold change to BI2536 and SBE13.

Alisertib is a selective, potent inhibitor of AURKA, and is structurally different from tozasertib which inhibits both AURKA and AURKB (Bebbington *et al.*, 2009; Manfredi *et al.*, 2011). When treated with alisertib both MDA-MB-468 and HCC1806 showed similar GI_{50} of 5.05 µM and 6.41 µM, whilst HCC38 had a higher GI_{50} of 10.9 µM (Figure 4.5E). This was not seen when the chemo-naive cell lines were treated with tozasertib, as both MDA-MB-468 and HCC38 have GI_{50} values of 0.0139 µM and 0.0329 µM, whilst HCC1806 had a higher GI_{50} value of 0.268 µM (Figure 4.5G). The chemo-naive cell lines demonstrated a greater sensitivity to tozasertib than alisertib. MDA-MB-468^{r}PCL^{20} demonstrated cross-resistant to both alisertib and tozasertib (RF = 8.74 and 29.28 respectively). HCC38^{r}ERI^{10} only demonstrated cross-resistance to alisertib (RF = 3.56), and MDA-MB-468^{r}CDDP^{1000}, MDA-MB-468^{r}DOX^{50}, MDA-MB-468^{r}ERI^{50} and HCC1806^{r}DOX$^{12.5}$ demonstrated cross-resistance only to tozasertib (RF = 3.42, 2.64, 21.16, 49.63). It was noted that MDA-MB-468^{r}PCL^{20} had demonstrated a strong cross-resistance to each of the second-generation mitotic drugs tested here (Figure 4.5 F, H).

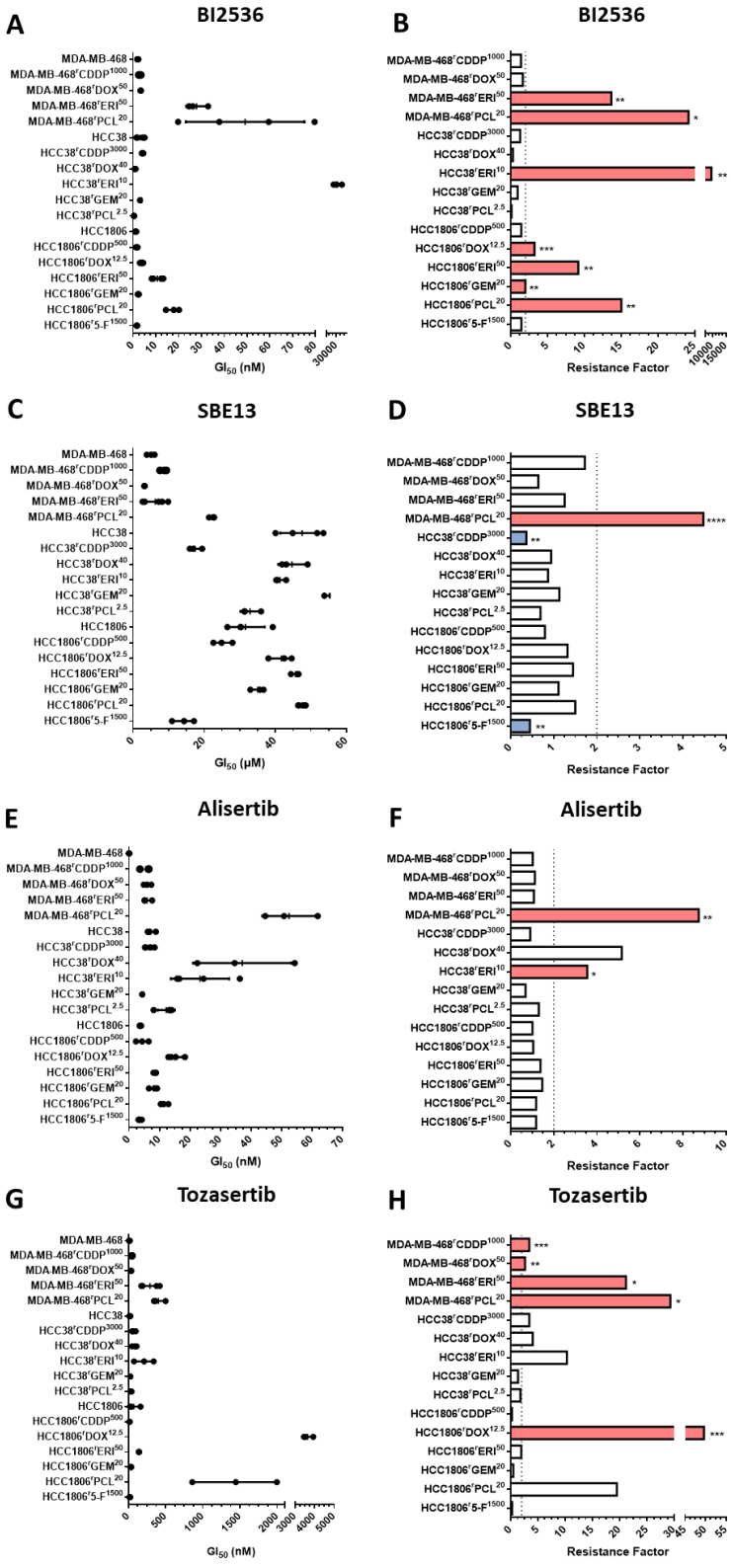


Figure 4.5 GI_{50} value determinations and resistance factors upon treatment with Polo/Aurora kinase inhibitors in the TNBC cell lines
The response of TNBC cell lines to stated Polo/Aurora kinase inhibitors was determined by 120-hour MTT assay and analysed using GraphPad Prism 6. A, C, E, G) Determined GI_{50} concentrations upon treatment with stated Polo/Aurora kinase inhibitors. Dot represents a single biological repeat from ≥3 independent experiments, horizontal line shows standard deviation and vertical line shows the mean. B, D, F, H) Calculated resistance factor (RF) to stated Polo/Aurora kinase inhibitors. Dotted line indicates two-fold RF. Statistical significance calculated using student's t-test, * P ≤ 0.05, ** p ≤ 0.01, *** p ≤ 0.001, **** p ≤ 0.00001. Colour indicates resistance status: red = resistant (RF >2, and ≥ * p), blue = acquired vulnerability (RF < 0.5 and ≥ * p), white = no change. Numerical data is presented in Appendix A3.

4.2.2 Intrinsic resistance and cross-resistance patterns

TNBC is highly heterogenic, harbouring different genetic mutations, which can result in cell lines demonstrating a different predisposition to certain drugs. The chemo-naive cell lines in this panel were selected to reflect this heterogenic nature (Table 3.1). The measurement of cross-resistance in the chemo-resistant TNBC cell lines, are relative to the respective chemo-naive, and does not reflect an absolute measure of drug activity. Consideration of patterns of intrinsic drug resistance was conducted in the chemo-naive cell lines, to understand if this had an outcome on the cross-resistance patterns seen in the chemo-resistant cell lines (Figure 4.6).

MDA-MB-468 showed intrinsic resistance to three of the CHK1 inhibitors; rabusertib, SRA737, MK-8776 (Figure 4.6). It was noted that very few MDA-MB-468 chemo-resistant cell lines demonstrated cross-resistance to these drugs, all be it, that the GI_{50} values were higher than any of the HCC38 and HCC1806 chemo-naive and chemo-resistant cell lines (Figure 4.1). HCC38 demonstrated intrinsic resistance to both AZD0156 and B02 inhibitors, compared to MDA-MB-468 and HCC1806. MDA-MB-468 showed acquired vulnerability to AZD0156 compared to the other chemo-naive cell lines, although the GI_{50} values were similar to HCC1806 chemo-resistant cell lines, which demonstrated cross-resistance (Figure 4.2). No statistically significant fold changes were seen with inhibition of B02, with an exception to HCC1806^{r}PCL^{20}, which demonstrated cross-resistance. Finally, both HCC38 and HCC1806 demonstrated intrinsic resistance to cisplatin and SBE13 compared to MDA-MB-468 (Figure 4.6). HCC38^{r}CDDP^{3000} demonstrated acquired vulnerability to SBE13, although the GI_{50} value was higher than the MDA-MB-468 cell lines, with an exception to MDA-MB-468^{r}PCL^{20} (Figure 4.5). Although MDA-MB-468^{r}PCL^{20} demonstrated cross-resistance to SBE13 the GI_{50} value was lower than the HCC38 and HCC1806 cell lines (Figure 4.5).

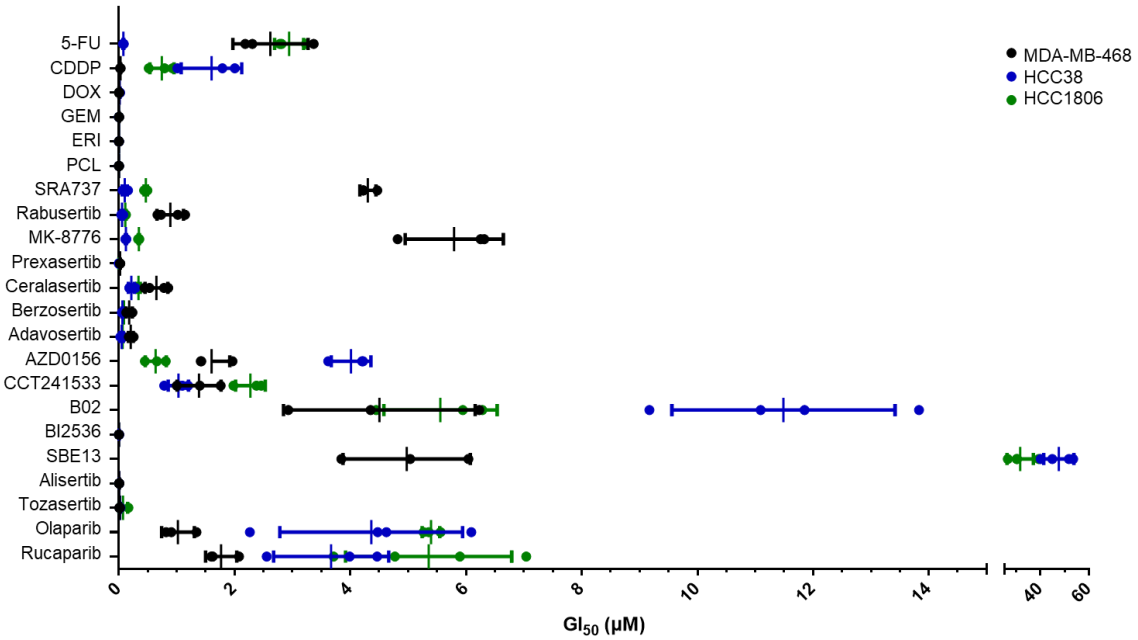


Figure 4.6 GI_{50} value determinations of chemo-naive cell lines to panel of drugs
The response of chemo-naive TNBC cell lines to stated drugs was determined by 120-hour MTT assay and analysed using GraphPad Prism 6. Each dot represents a single biological repeat from ≥3 independent experiments, horizontal line shows standard deviation and vertical line shows the mean. Numerical data is presented in Appendix A1, A2.

Heatmaps were generated to visualise patterns of cross-resistance and acquired vulnerability in the chemo-resistant cell lines, in relation to the chemo-naive counterpart (Figure 4.7). Chemo-resistant MDA-MB-468 all demonstrated cross-resistance to doxorubicin, tozasertib, olaparib, and the majority to eribulin and rucaparib. All showed either acquired vulnerability, or no change, in response to AZD0156, rabusertib, berzosertib and B02, and most to CCT241533 and adavosertib (Figure 4.7A). Chemo-resistant HCC38 cell lines showed almost all to be cross-resistant to 5-Fluorouracil, but no other significant cross-resistant patterns. All chemo-resistant HCC38 cell lines showed acquired vulnerability or no change to AZD0156, CCT241533, SBE13, and almost all to BI2536, alisertib and tozasertib (Figure 4.7B). Most of the chemo-resistant HCC1806 cell lines showed cross-resistance to 5-Fluorouacil, eribulin, doxorubicin, olaparib, rucaparib, alisertib, MK-8776, rabusertib, ceralasertib and AZD015, whilst most showed acquired vulnerability or no change to CCT241533, SRA737, cisplatin, SBE13, B02 and tozasertib (Figure 4.7C).

Taken together, these data suggests that patterns of cross-resistance or acquired vulnerability seen in the chemo-resistant cell lines can be a result of intrinsic resistance or vulnerability in the chemo-naive cell line. This could be due to genetic mutations present in the chemo-naive cell line which can pre-dispose the cell lines to certain drugs.

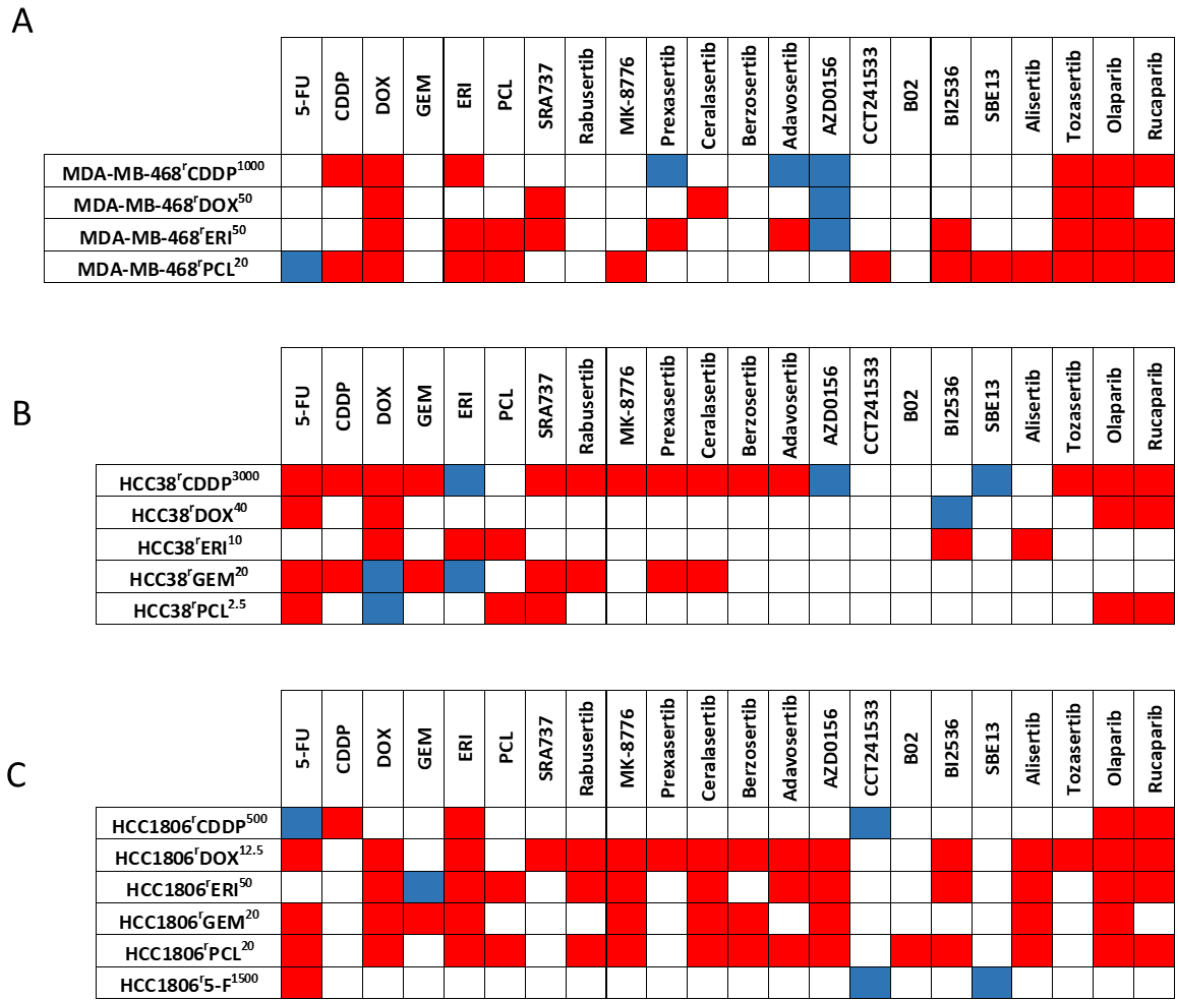


Figure 4.7 Heatmap showing cross-resistance and acquired vulnerability after drug profiling analysis in chemo-resistant TNBC cell lines
Cross-resistance and acquired vulnerability were determined from the calculated resistance factor (RF) in the chemo-resistant cells lines derived from A) MDA-MB-468, B) HCC38, C) HCC1806. Statistical significance and 2-fold RF indicate resistance status whereby; red = resistant (RF >2, and ≥ * p), blue = acquired vulnerability (RF < 0.5 and ≥ * p), white = no change relative to the respective chemo-naïve cell line. Numerical data is presented in Appendix A1, A2.

4.2.3 Patterns of cross-resistance in cell lines adapted to the same drug

Next, chemo-resistant TNBC cell lines adapted to have chemo-resistance to the same drug, were compared to each other, in order to identify any similar patterns of cross-resistance or acquired vulnerability. Volcano plots were generated to consider the GI_{50} fold change in relation to the p value calculated using the students t-test generated earlier in this chapter.

All three cisplatin resistant cell lines showed cross-resistant to olaparib and rucaparib, whilst all showed no change or acquired vulnerability to paclitaxel, AZD0156, CCT241533, B02, BI2536, SBE13 and alisertib (Figure 4.8). It was noted that HCC38^{r}CDDP^{3000} showed cross-resistance to all CHK1, ATR and WEE1 inhibitors, and this was seen at a greater RF than to cisplatin itself (Figure 4.8). This pattern of cross-resistance was not seen in the other two cisplatin resistant cell lines. Both MDA-MB-468^{r}CDDP^{1000} and HCC1806^{r}CDDP^{500} demonstrated cross-resistance to eribulin, but HCC38^{r}CDDP^{3000} demonstrated acquired vulnerability (Figure 4.8).

All three doxorubicin resistant cell lines showed cross-resistance to olaparib, whilst all showed either no change or acquired vulnerability to cisplatin, gemcitabine, paclitaxel, CCT241533, B02 and SBE13. HCC1806^{r}DOX$^{12.5}$ showed cross-resistance to all CHK1, ATR, WEE1 and ATM inhibitors. MDA-MB-468^{r}DOX^{50} demonstrated cross-resistance to only SRA737 and ceralasertib and acquired vulnerability to AZD0156, whilst HCC38^{r}DOX^{40} demonstrated no change to the CHK1, ATR, WEE1 and ATM inhibitors (Figure 4.9).

Each of the eribulin resistant cell lines demonstrated cross-resistance to paclitaxel, doxorubicin and BI2536, whilst all showed no change or acquired vulnerability to 5-Fluorouracil, cisplatin, berzosertib, CCT241533, B02, SB13 (Figure 4.10). HCC38^{r}ERI^{10} demonstrated a greater RF to BI2536, than to eribulin itself.

All three paclitaxel resistant cell lines demonstrated cross-resistance to olaparib and rucaparib, and showed no change or acquired vulnerability to gemcitabine and prexasertib (Figure 4.11). MDA-MB-468^{r}PCL^{20} and HCC1806^{r}PCL^{20} demonstrated cross-resistance to eribulin, and with a greater RF than to paclitaxel itself. Both of these cell lines also exhibited cross-resistance to BI2536 and doxorubicin. This pattern of cross-resistance to BI2536, doxorubicin and eribulin was not seen in HCC38^{r}PCL$^{2.5}$ (Figure 4.11).

Both of the gemcitabine resistant cell lines only demonstrated cross-resistance to 5-Fluorouracil, and showed no change or acquired vulnerability to paclitaxel, adavosertib, CCT241533, B02, BI2536, SBE13, tozasertib and rucaparib (Figure 4.12). HCC38^{r}GEM^{20} demonstrated acquired vulnerability to doxorubicin and eribulin with an RF of 0.36 and 0.39 respectively, whilst HCC1806^{r}GEM^{20} demonstrated cross-resistance to doxorubicin and eribulin with an RF of 10.79 and 7.59 respectively.

There was only one 5-Fluorouracil resistant cell line in this data set; HCC1806^{r}5-F^{1500}. HCC1806^{r}5-F^{1500} did not demonstrate any cross-resistance to any of the drugs in the panel (Figure 4.13). However, other cells lines had demonstrated cross-resistance to 5-Fluorouracil (Figure 3.16). Acquired vulnerability was seen in HCC1806^{r}5-F^{1500} to two drugs; CCT241533 (RF = 0.42) and SBE13 (RF = 0.45).

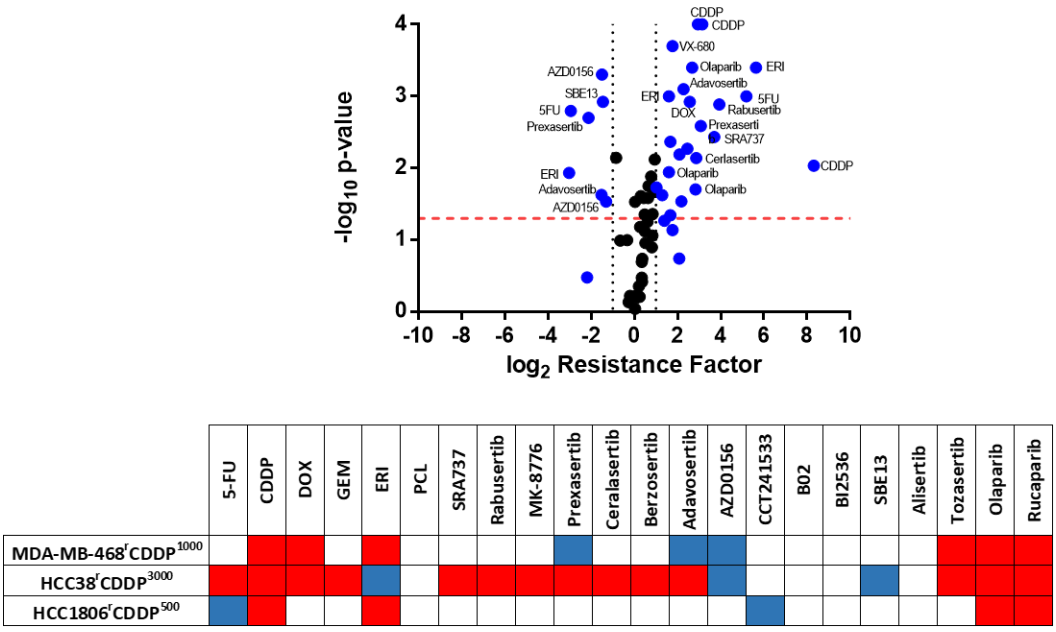


Figure 4.8 Volcano plot and heatmap of drug profiling analysis of cisplatin resistant TNBC cell lines
Analysis of cross-resistance and acquired vulnerability in the cisplatin resistant cell lines. Volcano plot shows calculated resistance factor (RF) against calculated p value from student t-test. Each dot represents a drug from the panel tested against a cisplatin resistant cell line. Blue dots have a ≥ 2-fold RF, black dots are < 2-fold RF. Dotted line indicates * $p < 0.05$. Heat map shows cross-resistance and acquired vulnerability in cisplatin resistant cell lines. Statistical significance and 2-fold RF indicate resistance status whereby; red = resistant (RF >2, and ≥ * p), blue = acquired vulnerability (RF < 0.5 and ≥ * p), white = no change relative to the respective chemo-naive. Numerical data is presented in Appendix A1, A2.

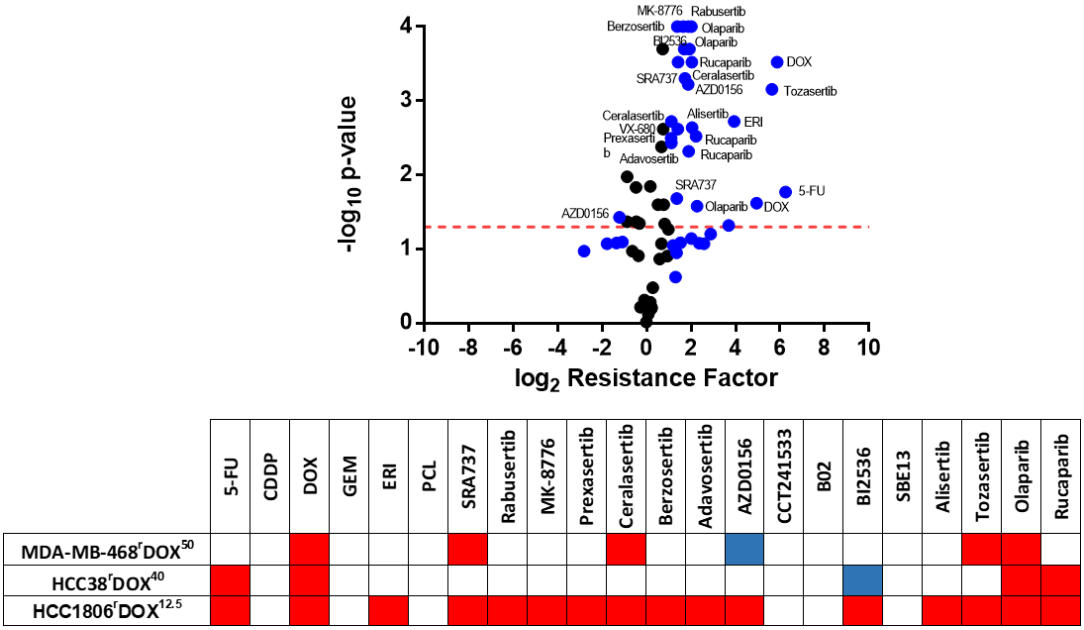


Figure 4.9 Volcano plot and heatmap of drug profiling analysis of doxorubicin resistant TNBC cell lines
Analysis of cross-resistance and acquired vulnerability in the doxorubicin resistant cell lines. Volcano plot shows calculated resistance factor (RF) against calculated p value from student t-test. Each dot represents a drug from the panel tested against a doxorubicin resistant cell line. Blue dots have a ≥ 2-fold RF, black dots are < 2-fold RF. Dotted line indicates * $p < 0.05$. Heat map shows cross-resistance and acquired vulnerability in doxorubicin resistant cell lines. Statistical significance and 2-fold RF indicate resistance status whereby; red = resistant (RF >2, and ≥ * p), blue = acquired vulnerability (RF < 0.5 and ≥ * p), white = no change relative to the respective chemo-naive. Numerical data is presented in Appendix A1, A2.

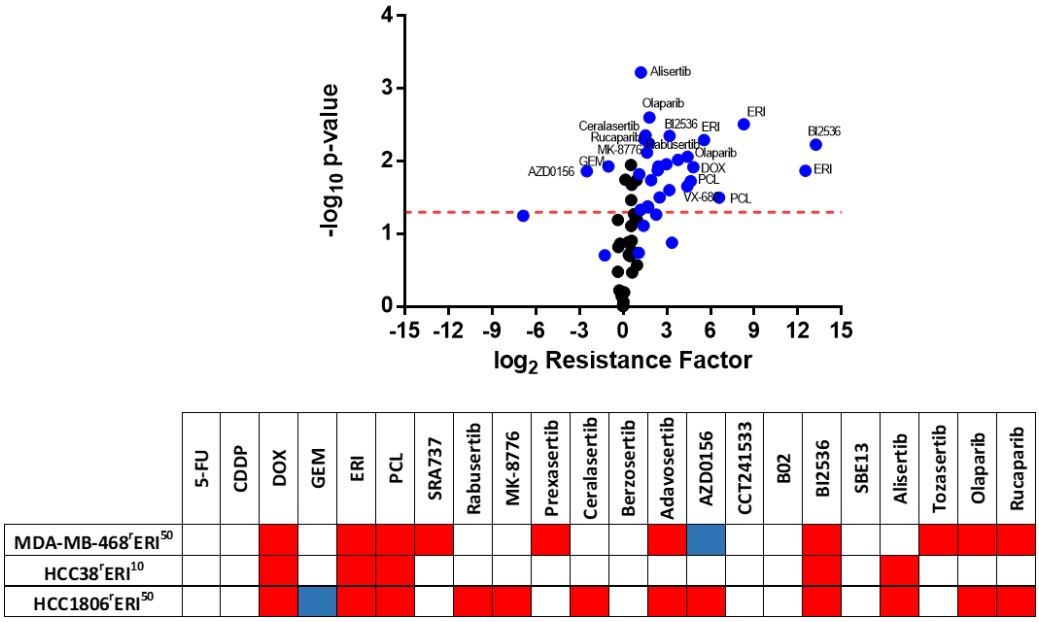


Figure 4.10 Volcano plot and heatmap of drug profiling analysis of eribulin resistant TNBC cell lines
Analysis of cross-resistance and acquired vulnerability in the eribulin resistant cell lines. Volcano plot shows calculated resistance factor (RF) against calculated p value from student t-test. Each dot represents a drug from the panel tested against an eribulin resistant cell line. Blue dots have a ≥ 2-fold RF, black dots are < 2-fold RF. Dotted line indicates * $p < 0.05$. Heat map shows cross-resistance and acquired vulnerability in eribulin resistant cell lines. Statistical significance and 2-fold RF indicate resistance status whereby; red = resistant (RF >2, and ≥ * p), blue = acquired vulnerability (RF < 0.5 and ≥ * p), white = no change relative to the respective chemo-naive. Numerical data is presented in Appendix A1, A2.

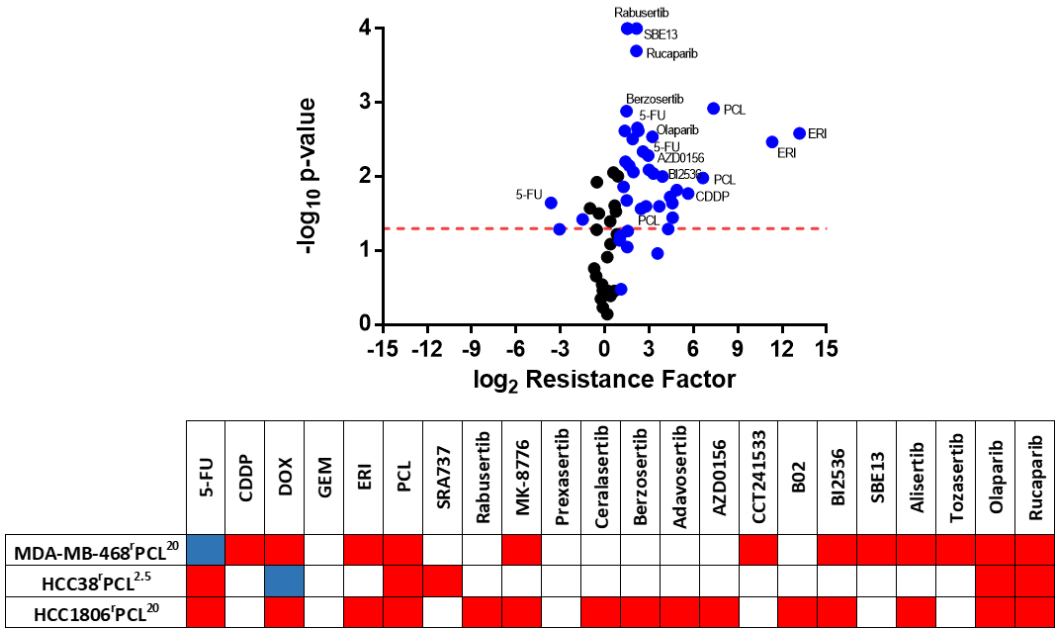


Figure 4.11 Volcano plot and heatmap of drug profiling analysis of paclitaxel resistant TNBC cell lines
Analysis of cross-resistance and acquired vulnerability in the paclitaxel resistant cell lines. Volcano plot shows calculated resistance factor (RF) against calculated p value from student t-test. Each dot represents a drug from the panel tested against a paclitaxel resistant cell line. Blue dots have a ≥ 2-fold RF, black dots are < 2-fold RF. Dotted line indicates * $p < 0.05$. Heat map shows cross-resistance and acquired vulnerability in paclitaxel resistant cell lines. Statistical significance and 2-fold RF indicate resistance status whereby; red = resistant (RF >2, and ≥ * p), blue = acquired vulnerability (RF < 0.5 and ≥ * p), white = no change relative to the respective chemo-naive. Numerical data is presented in Appendix A1, A2.

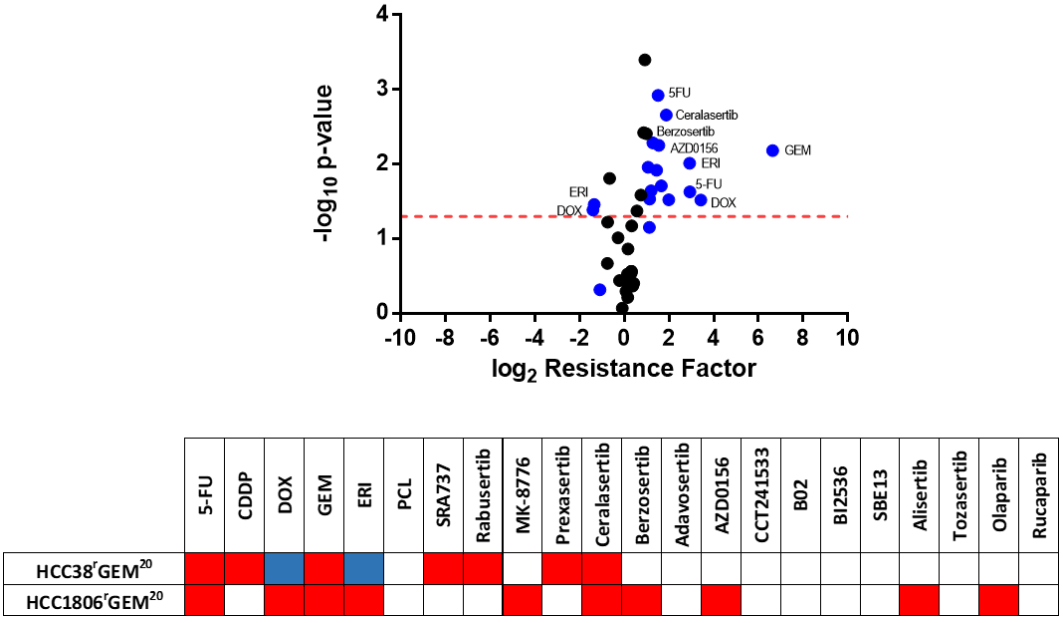


Figure 4.12 Volcano plot and heatmap of drug profiling analysis of gemcitabine resistant TNBC cell lines
Analysis of cross-resistance and acquired vulnerability in the gemcitabine resistant cell lines. Volcano plot shows calculated resistance factor (RF) against calculated p value from student t-test. Each dot represents a drug from the panel tested against a gemcitabine resistant cell line. Blue dots have a ≥ 2-fold RF, black dots are < 2-fold RF. Dotted line indicates * $p < 0.05$. Heat map shows cross-resistance and acquired vulnerability in gemcitabine resistant cell lines. Statistical significance and 2-fold RF indicate resistance status whereby; red = resistant (RF >2, and ≥ * p), blue = acquired vulnerability (RF < 0.5 and ≥ * p), white = no change relative to the respective chemo-naive. Numerical data is presented in Appendix A1, A2.

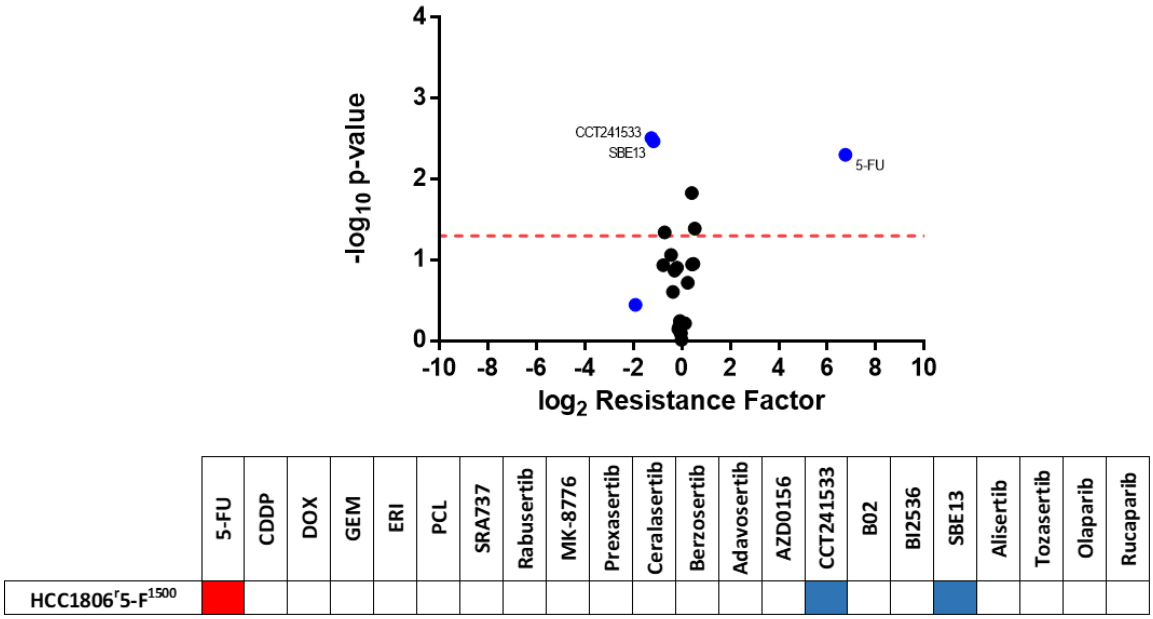


Figure 4.13 Volcano plot and heatmap of drug profiling analysis of 5-fluorouracil resistant TNBC cell line
Analysis of cross-resistance and acquired vulnerability in the 5-flurouracil resistant cell line. Volcano plot shows calculated resistance factor (RF) against calculated p value from student t-test. Each dot represents a drug from the panel tested against the 5-flurouracil resistant cell line. Blue dots have a ≥ 2-fold RF, black dots are < 2-fold RF. Dotted line indicates * $p < 0.05$. Heat map shows cross-resistance and acquired vulnerability in 5-flurouracil resistant cell line. Statistical significance and 2-fold RF indicate resistance status whereby; red = resistant (RF >2, and ≥ * p), blue = acquired vulnerability (RF < 0.5 and ≥ * p), white = no change relative to the respective chemo-naive. Numerical data is presented in Appendix A1, A2.

4.2.4 Overall cross-resistance

Given the need to determine viable second-line therapies after acquired drug resistance has occurred, it was important to divide the drugs based on their response in the cell lines. The chemo-resistant cell lines were divided into two groups based on the calculated RF for each drug; 1. resistant/ cross-resistant, 2. acquired vulnerability/ no statistically significant change (Figure 4.14).

Out of the fifteen chemo-resistant cell lines; twelve and ten cell lines demonstrated cross-resistance to olaparib and rucaparib respectively. Eleven demonstrated resistance/cross-resistance to doxorubicin. One cell line showed cross-resistance to CCT214533, B02 and SBE13, with the rest demonstrating either acquired vulnerability/ no statistically significant change. Also, only one cell line demonstrated cross-resistance to gemcitabine, with the other two counted having been generated to have acquired chemo-resistance. Taken together, these data suggest that it is common to develop cross-resistance to olaparib, rucaparib and doxorubicin in cell lines with resistance to chemotherapy agents, whilst it is uncommon for these cell lines to develop cross-resistance to gemcitabine CCT241533, B02, alisertib and SBE13.

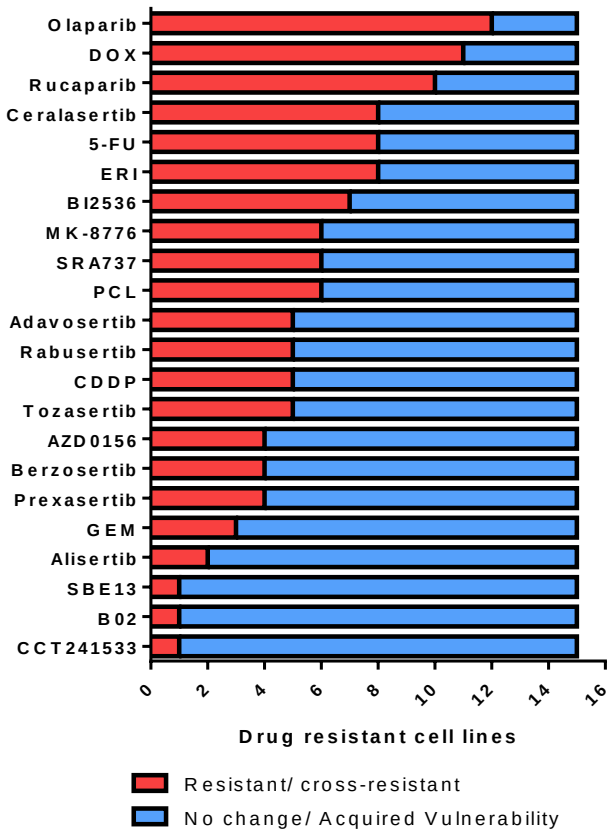


Figure 4.14 Summary of cross-resistance and acquired vulnerability of each drug across the drug-resistant TNBC cell lines
Each of the fifteen chemo-resistant cell lines were considered for cross-resistance or acquired vulnerability/no change to each drug tested in the panel (as stated). Red indicates the number of cell lines showing resistance/cross-resistance to stated drug, and blue indicates the number of cell lines showing acquired vulnerability/ no change to stated drug.

4.2.5 Δ method to identify differential effects of drug activity profiles in TNBC cell lines

Earlier in this chapter, patterns of cross-resistance were defined in the chemo-resistant cell lines, which considers the development of cross-resistance relative to the chemo-naive counterpart, but does not reflect an absolute measure of drug activity in the cell lines. The next analysis aimed to consider the drug activity profiles across the cell line panel. The Δ (delta) method was developed by the National Cancer Institute (NCI), and can be used in order to emphasise differential effects of the panel of drugs against the chemo-naive and chemo-resistant TNBC cell lines (Bracht *et al.*, 2006). The GI_{50} values previously obtained in this chapter and Chapter 3 were transformed to ΔGI_{50} values using the following equation:

ΔGI_{50} = log (average GI_{50} in drug over all cell lines) – log (individual GI_{50} in drug for each cell line)

Linear regression analysis of ΔGI_{50X} versus ΔGI_{50Y} where X and Y represent two different drugs from the panel, were 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 $p \leq 0.05$ was considered significant. This analysis was performed on each combination of drugs from the drug panel. Table 4.2 shows the summary of the analysis, whereby the value in the table are the (*r*) values and the heatmap indicates the level of statistical significance (p values).

Strong positive correlation between the activities of the CHK1 inhibitors was found where $r = 0.8495 – 0.9423$, $p \leq 0.0001$. A strong positive correlation was also found between the two ATR inhibitors where $r = 0.9282$, $p \leq 0.0001$, and between the two PARP inhibitors where $r = 0.9433$, $p \leq 0.0001$. Given that the drugs here target the same protein, this was not unexpected.

A strong positive correlation was found between eribulin and paclitaxel, where $r = 0.9316$, $p \leq 0.0001$. Doxorubicin positively correlated with eribulin ($r = 0.491$, $p \leq 0.05$), and both doxorubicin and eribulin positively correlated with alisertib ($r = 0.548$, $p \leq 0.05$ and $r = 0.5302$, $p \leq 0.05$ respectively) and tozasertib ($r = 0.7.374$, $p \leq 0.001$ and $r = 0.5732$, $p \leq 0.05$ respectively). Paclitaxel only positively correlated with the tozasertib ($r = 0.575$, $p \leq 0.05$). Eribulin was also found to positively correlate with BI2536 ($r = 0.484$, $p \leq 0.05$), although this was not seen with paclitaxel. Further to this, both eribulin and paclitaxel positively correlate with olaparib, alisertib ($r = 0.488$, $p \leq 0.05$ and $r = 0.5048$, $p \leq 0.05$ respectively). Doxorubicin was also found to positively correlate with both PARP inhibitors; olaparib and rucaparib ($r = 0.5101$ $p \leq 0.05$ and $r = 0.6071$ $p \leq 0.01$ respectively).

It was also interesting to note a positive correlation in the drug activity profiles of eribulin and paclitaxel to the CHK1 inhibitor; MK-8776 ($r = 0.5057$, $p \leq 0.05$ and $r = 0.5044$, $p \leq 0.05$ respectively). Figure 4.1D shows that HCC1806^{r}ERI^{50}, MDA-MB-468^{r}PCL^{20} and HCC1806^{r}PCL^{20} each demonstrate cross-resistance to MK-8776. The activity to other CHK1 inhibitors differed whereby these cell lines showed no change in response to SRA737 and Prexasertib compared to the respective chemo-naive, and treatment with rabusertib showed acquired vulnerability in MDA-MB-468^{r}PCL^{20} (RF = 0.50) and cross-resistance in HCC1806^{r}ERI^{50} and HCC1806^{r}PCL^{20} (RF = 3.4 and 2.87 respectively) (Figure 4.1).

A positive correlation was seen between the drug activity profiles of CHK1, ATR, WEE1 and CHK2 inhibitors, whilst all demonstrated a negative correlation with the PLK1/2 inhibitor SBE13 ($r = -0.4767$ to -0.8533 $p \leq 0.0001$ to 0.05). The CHK1 inhibitors, with the exception to Prexasertib, all demonstrated a negative correlation with ATM inhibitor; AZD0156 ($r = -0.5751$ to -0.6483 $p \leq 0.01$-0.05), and SRA737 and Prexasertib demonstrated a negative correlation with the RAD51 inhibitor;

B02 (r = -0.4938 $p \leq 0.05$ and r = -0.481 $p \leq 0.05$ respectively). Opposite to this pattern, cisplatin demonstrated a negative correlation with the CHK1 inhibitors (r = -0.4702 to -0.6388 $p \leq 0.05$ – 0.01), but a positive correlation with both B02 and SBE13 (r = 0.5292 $p \leq 0.05$ and r = 0.5312 $p \leq 0.05$ respectively). AZD0156 demonstrated a positive correlation with both B02 and SBE13 (r = 0.6437 $p \leq 0.05$ and r = 0.7096 $p \leq 0.001$ respectively). Taken together, an inverse relationship is seen between the drug activity of CHK1, ATR, WEE1 and CHK2 inhibitors, to cisplatin, ATM, RAD51 and PLK1/2 inhibitors.

Table 4.2: Summary of Pearson correlation coefficients and two-tailed test obtained from Δ method analysis

					Microtubule		Chk1			Chk1/Chk2	ATR		Wee1	ATM	CHK2	RAD51	PLK1	PLK1,2	Aurora A	Aurora (pan)	PARP	
	CDDP	DOX	GEM	5-FU	ERI	PCL	Rabusertib	MK-8776	SRA737	Prexasertib	Ceralasertib	Berzosertib	Adavosertib	AZD0156	CCT241533	B02	BI2536	SBE13	Alisertib	Tozasertib	Olaparib	Rucaparib
CDDP	1	-0.2163	0.4467	-0.2803	-0.1521	-0.1925	-0.3598	-0.4702	-0.4803	-0.6388	-0.04013	-0.1964	-0.4139	0.3788	-0.1229	0.5292	0.1183	0.5312	0.0758	-0.1047	0.2464	0.2409
DOX		1	-0.2533	0.218	0.4911	0.4247	0.3333	0.367	0.2748	0.4056	0.4732	0.3907	0.1615	0.03466	0.3736	0.102	0.3238	0.0006027	0.548	0.7374	0.5101	0.6071
GEM			1	-0.1957	-0.3474	-0.4456	-0.2164	-0.2231	-0.2579	-0.3277	0.06139	0.004831	-0.3523	0.4624	-0.1891	0.003953	-0.1257	0.265	0.03104	-0.334	-0.2629	-0.1892
5-FU				1	-0.1574	-0.1287	0.2138	0.1163	0.1678	0.1923	0.1211	0.07552	0.0637	-0.4094	-0.03066	-0.3363	-0.2507	-0.2026	-0.2481	0.1578	-0.01432	0.1692
ERI					1	0.9316	0.2646	0.5057	0.2759	0.3515	0.4576	0.28	0.3736	0.09283	0.5255	0.3497	0.4849	0.003682	0.5302	0.5732	0.4877	0.432
PCL						1	0.2828	0.5044	0.3011	0.4191	0.4632	0.3191	0.4968	0.0151	0.536	0.3396	0.461	-0.04814	0.4678	0.5755	0.5048	0.4141
Rabusertib							1	0.8622	0.9423	0.8495	0.7718	0.8569	0.7768	-0.587	0.5798	-0.383	-0.02318	-0.7504	-0.2507	0.1571	-0.03476	0.02465
MK-8776								1	0.9066	0.9075	0.725	0.7601	0.7748	-0.4483	0.7199	-0.4066	0.04952	-0.6823	0.0809	0.2715	0.05423	0.1019
SRA737									1	0.8908	0.6405	0.7653	0.7375	-0.6483	0.6167	-0.4938	-0.04883	-0.8533	-0.1694	0.1244	-0.08551	-0.04682
Prexasertib										1	0.6641	0.7574	0.821	-0.5751	0.6338	-0.481	0.006958	-0.7168	-0.03754	0.25	0.04151	0.06158
Ceralasertib											1	0.9282	0.7152	-0.1329	0.5623	0.02164	0.1873	-0.266	0.03747	0.4642	0.3707	0.4091
Berzosertib												1	0.7955	-0.2974	0.4671	-0.1797	0.1749	-0.4767	-0.1419	0.3467	0.1682	0.1727
Adavosertib													1	-0.4665	0.4403	-0.1911	0.02962	-0.5047	-0.217	0.2699	0.1571	0.1153
AZD0156														1	-0.3187	0.6437	0.1868	0.7096	0.5452	0.1791	0.1962	0.1908
CCT241533															1	-0.1923	-0.02043	-0.3443	0.2498	0.3135	0.2793	0.3418
B02																1	0.318	0.6656	0.3407	0.4032	0.4972	0.4147
BI2536																	1	0.1033	0.413	0.4083	0.104	-0.01209
SBE13																		1	0.3437	0.2763	0.4127	0.4117
Alisertib																			1	0.5358	0.4058	0.4593
Tozasertib																				1	0.6121	0.6714
Olaparib																					1	0.9433
Rucaparib																						1

p ≤ 0.0001	p ≤ 0.001	p ≤ 0.01	p ≤ 0.05

4.2.6 Validation of $p21^{CIP1/WAF1}$ as a mechanism of drug resistance in $HCC38^{r}CDDP^{3000}$

Whilst the drug activity profiles of cisplatin negatively correlated with the CHK1 inhibitors (section 4.2.5), cross-resistance profiling identified that $HCC38^{r}CDDP^{3000}$ demonstrated a very strong cross-resistance to each of the CHK1 inhibitors (4.2.1.1), the ATR inhibitors (4.2.1.2) and the WEE1 inhibitor (4.2.1.3). This contradiction between the drug activity profiles and cross-resistance profiling warranted further investigation.

Briefly, cisplatin is known to induce S-phase cell cycle arrest, which is reliant on the activation of the DDR pathways through CHK1 (Zhang *et al.*, 2008; Wagner and Karnitz, 2009; Barr *et al.*, 2017). CHK1 activity leads to the inhibition of CDC25 and prevents activation of CDK1/2, required for the cell to enter G2 (Thompson and Eastman, 2013). It was found that a subset of cancer cell lines were sensitive to CHK1 inhibitor, MK-8776, as a monotherapy due to CHK1 requirement for CDK2 activation in S phase, and those that showed resistance had increased activation of CDK2 in S phase (Sakurikar *et al.*, 2016). Further to this, it has also been seen that $p21^{CIP1/WAF1}$ induction inhibited endoreduplication through direct cyclin E/CDK2, and high $p21^{CIP1/WAF1}$ levels mediate G1 arrest via CDK inhibition (Stewart and Leach, 1999). Here it was hypothesised that by-pass of S-G2/M checkpoint resulted in increased $p21^{CIP1/WAF1}$, due to accumulation of DNA damage in S phase, leading to inhibition of cyclin E/CDK2 in G1 to allow for cisplatin induced DNA repair. To this end, $p21^{CIP1/WAF1}$ expression was investigated as a potential biomarker of cisplatin/ CHK1 resistance.

4.2.6.1 Protein and gene expression of $p21^{CIP1/WAF1}$ /CDKN1A in $HCC38^{r}CDDP^{3000}$

In order to compare the level of $p21^{CIP1/WAF1}$ protein between HCC38 and $HCC38^{r}CDDP^{3000}$, western blot analysis was conducted. Heavily implicated in the cell cycle and growth arrest, $p21^{CIP1/WAF1}$ has also been shown to be involved in cell contact inhibition (Gartel, 2006). Cells were seeded at different stated densities and grown for 72 hours, before lysate proteins obtained for western blot analysis. $p21^{CIP1/WAF1}$ protein expression is elevated, and cyclin E1 is decreased in $HCC38^{r}CDDP^{3000}$ compared to HCC38 (Figure 14B). Further to this, no change was observed in $p21^{CIP1/WAF1}$ expression at different seeding densities which suggests in these cell lines that $p21^{CIP1/WAF1}$ expression is not affected by contact inhibition (Figure 14A-B). The seeding density of two million cells was used for the next experiment, which aimed to determine if dosing with increasing concentration of cisplatin induced $p21^{CIP1/WAF1}$ expression. The highest dose of cisplatin (40 µM) showed a $p21^{CIP1/WAF1}$ induction in HCC38, but a decrease was observed in $HCC38^{r}CDDP^{3000}$ (Figure 4.15C).

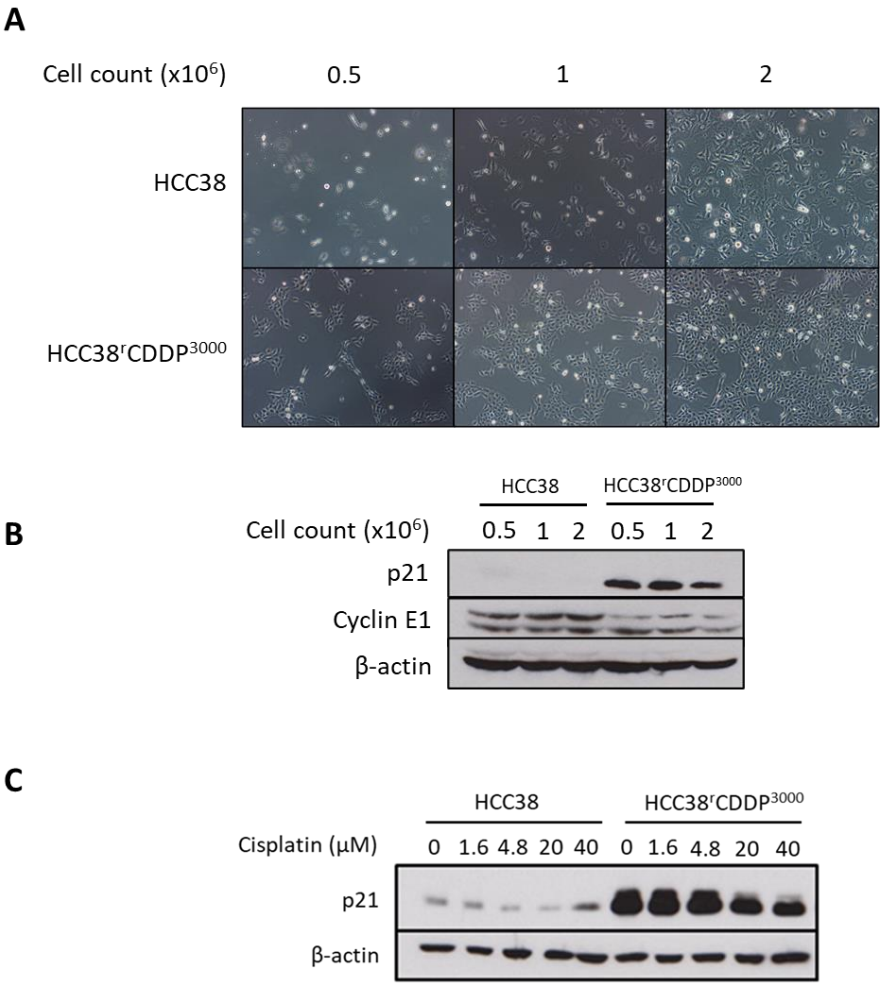


Figure 4.15 p21$^{CIP1/WAF1}$ and cyclin E1 protein levels in HCC38 and HCC38^{r}CDDP^{3000} cell lines
Cells were plated at 0.5, 1 or 2x10^{6} in 10cm tissue culture dishes and incubated at normal growth conditions for 72 hours. A) Cells were visualised at x40 magnification, and images taken. B) Cell lysates were prepared and analysed by western blotting to determine the expression levels of p21$^{CIP1/WAF1}$ and cyclin E1. C) Cells plated at 2x10^{6} in 10cm tissue culture dishes and incubated at normal growth conditions for 24 hours before dosing with stated concentration of cisplatin and incubating for a further 48 hours. Cell lysates were prepared and analysed by western blotting to determine the expression levels of p21$^{CIP1/WAF1}$. B and C) β-actin was used as a loading control. All data are representative of ≥3 independent experiments.

As protein expression of p21$^{CIP1/WAF1}$ was increased at basal levels in HCC38^{r}CDDP^{3000} compared to HCC38, investigation into the levels of *CDKN1A* gene expression was conducted in order to determine if increased protein expression is the result of elevated gene expression. Briefly, RNA was extracted, and purified from both HCC38 and HCC38^{r}CDDP^{3000}, and normalised to an equal concentration. Using the enzyme reverse transcriptase, cDNA was generated through reverse transcription. Qualitative polymerase chain reaction was then performed with the cDNA and two sets of primers; which have previously demonstrated successful amplification of *CDKN1A* in past publications (Al-Haj, Blackshear and Khabar, 2012; He *et al.*, 2017). The primers are termed Primer set 1 and Primer set 2, and their targets are indicated on Figure 4.16A. The amplification of *CDKN1A* was measured in each cell line, and normalised to the housekeeping gene *GAPDH*. From this, the

fold change in *CDKN1A* amplification for the two primers can be calculated in HCC38^{r}CDDP^{3000} relative to HCC38 (Figure 4.16B). Both primer sets demonstrated increased amplification of *CDKN1A* in HCC38^{r}CDDP^{3000} versus HCC38, with a mean fold change of 34.62 and 25.89 in Primer set 1 and Primer set 2 respectively. This suggests that there is a higher expression of *CDKN1A* in HCC38^{r}CDDP^{3000} compared to HCC38, which could result in the higher observed p21$^{CIP1/WAF1}$ levels (Figure 4.15). Taken together, these data (Figure 4.15 and Figure 4.16) show that both gene and protein expression of *CDKN1A*/p21$^{CIP1/WAF1}$ are elevated in HCC38^{r}CDDP^{3000} compared to HCC38.

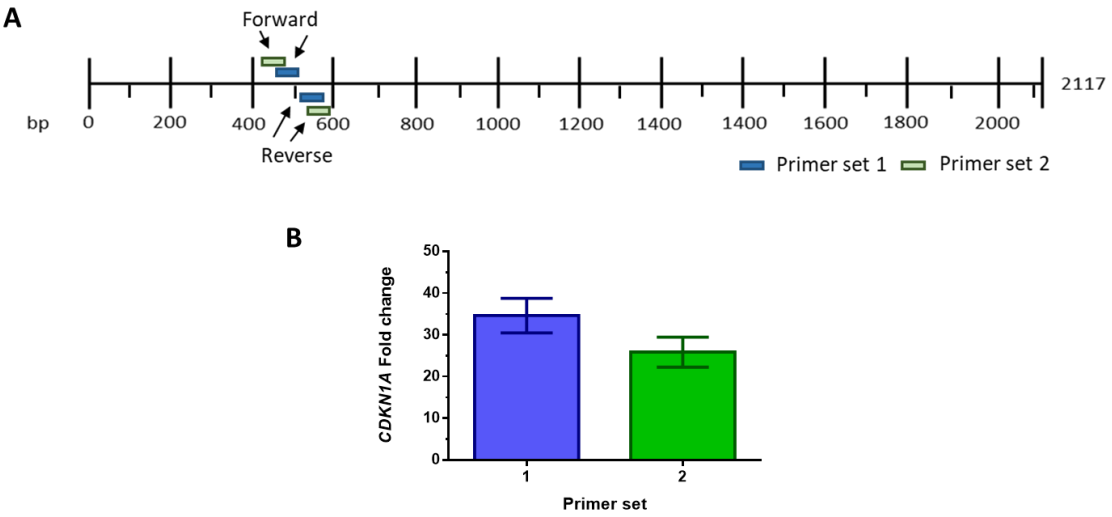


Figure 4.16 RT-qPCR analysis of *CDKN1A* in HCC38 and HCC38^{r}CDDP^{3000} cell lines
Analysis of *CDKN1A* expression conducted using reverse transcription - quantitative polymerase chain reaction (qPCR). A) Schematic of the regions of CDKN1A mRNA where the primers; Primer set 1 and Primer set 2, target. RNA extracted from HCC38 and HCC38^{r}CDDP^{3000} and reverse transcribed to obtain cDNA. Normalised cDNA from HCC38 and HCC38^{r}CDDP^{3000} underwent RT-qPCR with Primer set 1 and Primer Set 2. GAPDH used as housekeeping gene. B) Calculated fold change of *CDKN1A* amplification in HCC38^{r}CDDP^{3000} compared to HCC38 when using each primer set. Data is of 3 biological independent experiments.

4.2.6.2 Investigation of elevated expression of CDKN1A as a mechanism of resistance using RNA interference

With HCC38^{r}CDDP^{3000} demonstrating both increased gene and protein expression of *CDKN1A*/p21$^{CIP1/WAF1}$, and cross-resistance to CHK1 inhibitors compared to HCC38, it was hypothesised that knockdown of p21$^{CIP1/WAF1}$, through small interfering RNA (siRNA), will sensitise the cells to both cisplatin and CHK1 inhibitors. In order for efficient protein knockdown by siRNA, the transfection conditions must be optimised for oligonucleotide delivery. Lipofectamine 2000, a common cationic lipid-based transfection reagent, was selected and optimisation experiments was conducted to determine the conditions, which enabled the greatest transfection efficiency and the lowest transfection reagent toxicity. HCC38^{r}CDDP^{3000} cells were plated in a 96-well plate and reverse

transfected with 0.025 - 0.1 % Lipofectamine 2000, which was either combined with non-targeting siRNA (NT, nonspecific siRNA to measure knockdown levels versus background) or death control siRNA (a pooled selection of siRNA's which are known to target genes required for cell survival).The death siRNA was used at two different concentrations; 5 nM or 25 nM. A mock transfection control was also used, which did not contain siRNA. Both the mock and NT transfections were used to assess transfection reagent toxicity, whilst the death siRNA was used to assess the transfection efficiency. This was carried out over two seeding densities of the HCC38^{r}CDDP^{3000} cells; 12800 19200 cells per well. After reverse transfection, the cells were left to grow in standard growth conditions for four-days, after which cell viability was measured by SRB assay.

Both death controls (5 nM, 25 nM) reduced cell viability at both cell densities (12800, 19200), but demonstrated a greater loss of cell viability, and therefore transfection effciency, at 12,800 cells/well (Figure 4.17). Both mock and NT controls showed a concentration dependent reduction in cell viability, demonstrating transfection reagent toxicity at both cell densities. It was also noted that a higher concentration of Lipofectamine 2000 was required for efficient transfection in the higher cell density (Figure 4.17B). The greatest window of transfection efficiency and minimal transfection regent toxicity was observed with the conditions; 0.05% Lipofectamine 2000 at 12800 cells/well.

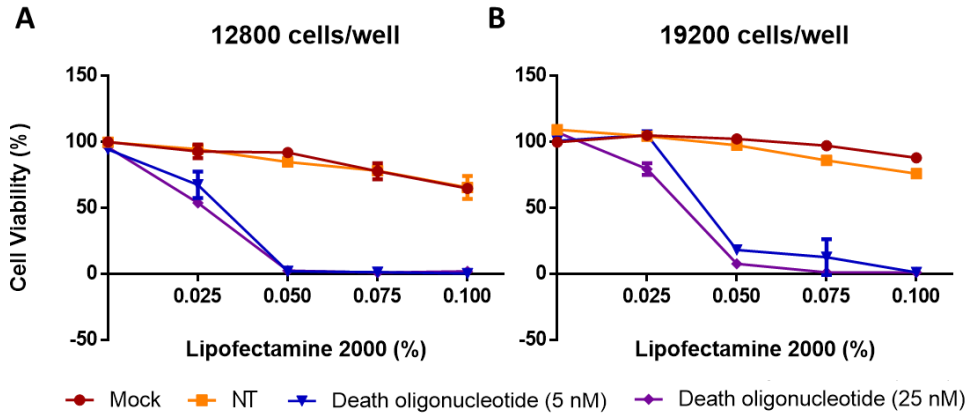


Figure 4.17 Determination of optimal conditions for siRNA knockdown studies in HCC38^{r}CDDP^{3000}
HCC38^{r}CDDP^{3000} cells plated at A) 12800 or B) 19200 cells per well and reverse transfected with Lipofectamine 2000 at the concentrations indicated, with RNAse free water (mock), 25 nM non-targeting Allstar negative control oligonucleotide (NT), or 5 or 25 nM Allstar positive control oligonucleotide (Death). Incubated at normal growth conditions for 96 hours before cells were fixed and analysed using the SRB cell viability assay. Data normalised to untreated cells and analysed using GraphPad Prism 6.

With the optimal siRNA transfection conditions identified, p21$^{CIP1/WAF1}$ siRNA knockdown was performed in HCC38^{r}CDDP^{3000} to investigate if the loss of p21$^{CIP1/WAF1}$ alone can sensitise the cells to cisplatin and/or an inhibitor of CHK1. Four p21$^{CIP1/WAF1}$ targeting oligonucleotides were tested for their ability to knockdown p21$^{CIP1/WAF1}$ levels after 48 hours. The oligonucleotides from Dharmacon

were 5, 6, 7, and 9 and targeted the mRNA (Figure 4.18A). The oligonucleotides were selected as successful p21$^{CIP1/WAF1}$ siRNA knockdown was previously observed (Di Stefano *et al.*, 2011). The optimised reverse transfection conditions were scaled up from 96 well, to 6 well plates in order to obtain sufficient protein for western blotting. Mock and NT did not reduce levels of p21$^{CIP1/WAF1}$, which shows that Lipofectamine 2000 and transfected NT does not reduce p21$^{CIP1/WAF1}$ levels (Figure 4.18). Both 5 and 9 showed some reduction of p21$^{CIP1/WAF1}$ levels, although the effciency was not as good as 7 and 6, which both demonstrated the greatest reduction (Figure 4.18).

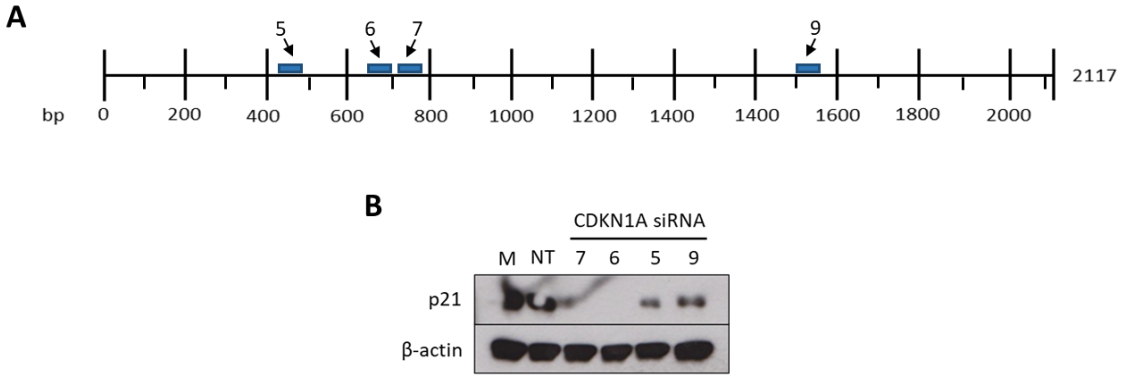
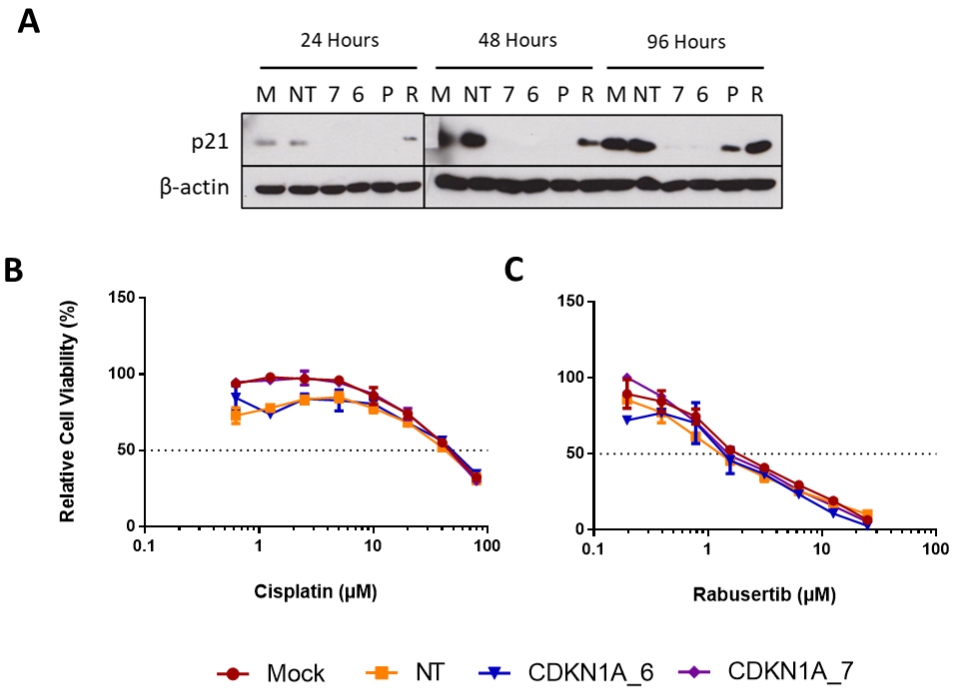


Figure 4.18 Optimisation of p21$^{CIP1/WAF1}$ **oligonucleotide knockdown**
A) Schematic of the regions of p21 mRNA where the p21$^{CIP1/WAF1}$ oligonucleotides; 5, 6, 7, and 9 targets. B) Western blot analysis of p21$^{CIP1/WAF1}$ knockdown with p21$^{CIP1/WAF1}$ oligonucleotides**.** HCC38^{r}CDDP^{3000} were reverse transfected with RNAse free water (mock; M), 25 nM non-targeting Allstar negative control oligonucleotide (NT), or 10 nM of p21 oligonucleotides (5, 6, 7, 9). Reverse transfection was performed using 0.05% Lipofectamine 2000 in 6-well plates and incubated for 48 hours. Cells were lysed and analysed by western blotting. Lysate proteins were separated by SDS-PAGE, transferred to PVDF membranes and probed with antibodies (indicated) against proteins and developed. β-actin was used as a loading control.

To determine if loss of p21$^{CIP1/WAF1}$ expression can induce sensitivity to cisplatin and a CHK1 inhibitor (rabusertib), a knockdown and dose response was performed. HCC38^{r}CDDP^{3000} demonstrated the highest RF to Rabusertib, compared to the four CHK1 inhibitors, and was therefore selected for this assay (Figure 4.1). p21$^{CIP1/WAF1}$ knockdown was performed in 96 well plates with HCC38^{r}CDDP^{3000} cells, as per the optimal conditions previously established, and treated with either cisplatin or rabusertib, in a dose response manner, and incubated for an additional 72 hours, before cell viability was assessed with an SRB assay (Figure 4.19B-C). Simultaneously, HCC38^{r}CDDP^{3000} cells were plated into 6 well plates, without drug treatment, in order to observe knockdown at several time points throughout the assay (Figure 4.19A).

NT transfection was not observed to reduce p21$^{CIP1/WAF1}$ levels across all time points, and both the mock and NT demonstrated consistent p21$^{CIP1/WAF1}$ levels compared to HCC38^{r}CDDP^{3000} cells used as a positive control (Figure 4.19A). Successful knockdown of p21$^{CIP1/WAF1}$ was observed with transfection of 6 and 7 oligonucleotides across the full 96-hour period, demonstrating that

knockdown occurred during the 72-hour drug incubation period (Figure4.19A). Despite the successful knockdown of p21$^{CIP1/WAF1}$, no significant change was observed between the mock and the transfected oligonucleotides when treated with either cisplatin or rabusertib (Figure B-D). Knockdown of p21$^{CIP1/WAF1}$ did not induce sensitivity in HCC38^{r}CDDP^{3000} to either cisplatin or rabusertib. This suggests that p21$^{CIP1/WAF1}$ alone is not a sufficient biomarker of resistance to either cisplatin or the CHK1 inhibitor rabusertib.



D

		M	NT	CDKN1A_6	CDKN1A_7
Cisplatin	GI_{50} (μM)	45.73	47.26	44.28	55.5
	RF	-	1.0334572	0.9682921	1.2136453
Rabusertib	GI_{50} (μM)	2.161	1.501	1.423	1.974
	RF	-	0.6945858	0.6584914	0.913466

Figure 4.19 p21$^{CIP1/WAF1}$ siRNA knockdown in HCC38^{r}CDDP^{3000} and response to Cisplatin or Rabusertib
HCC38^{r}CDDP^{3000} cells were reverse-transfected in 96 well plates with RNAse free water (mock; M), 25 nM non-targeting Allstar negative control oligonucleotide (NT), or 10 nM of p21$^{CIP1/WAF1}$ oligonucleotides (6; CDKN1A_6, 7; CDKN1A_7) using 0.05% Lipofectamine 2000. 24 hours after transfection, cells were treated with either a serial dilution of Cisplatin or Rabusertib. A) HCC38^{r}CDDP^{3000} cells were reverse-transfected as above in 6-well dishes simultaneously with 96 well plate transfection as knockdown effciency controls. Plates were incubated for 24, 48, 96 hours as indicated. Cells were used and analysed by western blotting as per Figure4.18. β-actin was used as a loading control. Dose response curves generated and half-maximal growth inhibition concentrations (GI_{50}) were calculated using GraphPad Prism 6 for treatment with B) Cisplatin or C) Rabusertib. Growth curves were normalised to untreated control, for each transfection condition. Dotted line indicates the GI_{50}. D) Summary of GI_{50} values for Cisplatin and Rabusertib transfected HCC38^{r}CDDP^{3000} cells. Data representative of two independent experiments. RF = resistance factor.

4.2.6.3 Clinical relevance of CDKN1A expression and drug response

Given the aim of this thesis is to identify biomarkers of resistance and potential second-line therapies, it was imperative to obtain a clinical relevance when investigating these. A computational analysis using data from the Sanger Genomics of Drug Sensitivity in Cancer (GDSC), was run simultaneously alongside the wet-lab validation of $p21^{CIP1/WAF1}$ as a biomarker for cisplatin and CHK1 drug resistance. The GDSC is a resource in which >1000 genetically characterised human cancer cell lines were screened with anti-cancer therapeutics (Yang *et al.*, 2013). Sensitivity patterns of cell lines to drugs can be correlated with regards to genomic and expression data to identify cancer biomarkers.

CDKN1A expression data was extracted from GDSC alongside the response to the drugs; cisplatin, AZD7762 (CHK1 inhibitor), 681640 (WEE1/CHK1 inhibitor), QL-VIII-58 (ATR/MTOR inhibitor), and KU-55933, CP466722 (ATM inhibitors). The data extracted was pan-cancer cell lines, as restricting the analysis to breast cancer cell lines result in a small data set which resulted in unattainable statistical analysis. The number of cell lines which had both drug data and *CDKN1A* expression data differed in the GDSC dataset. The same 88 cell lines were analysed upon treatment with cisplatin and AZD77632. For the ATM inhibitors, KU-55933 and CP466722, 93 and 103 cell lines were analysed respectively, and for the ATR inhibitor (QL-V111-58), and the WEE1/CHK1 inhibitor (681640), 55 and 86 cell lines were analysed respectively. The response of the drugs was measured as area under the curve, a method which takes into account percentage of cells killed, which is not reflected in GI_{50} values. The *CDKN1A* expression data was divided into two categories based on the mean, whereby higher than the mean is considered high expression, and lower than the mean is considered low expression. Box plots were created of the distribution of response to the drugs, vs high and low expression, and a student t-test conducted (Figure 4.20).

It was observed that the response of patients in this data set to cisplatin is not perturbed in relation to *CDKN1A* expression (Figure 4.20A). It also shows that the response of the patients to inhibition of CHK1, WEE1, ATR or ATM is not divergent when *CDKN1A* expression is high or low (Figure 4.20B-F). Taken together this data shows that no significant difference was seen in the response to each of the drugs considered when *CDKN1A* expression was high or low.

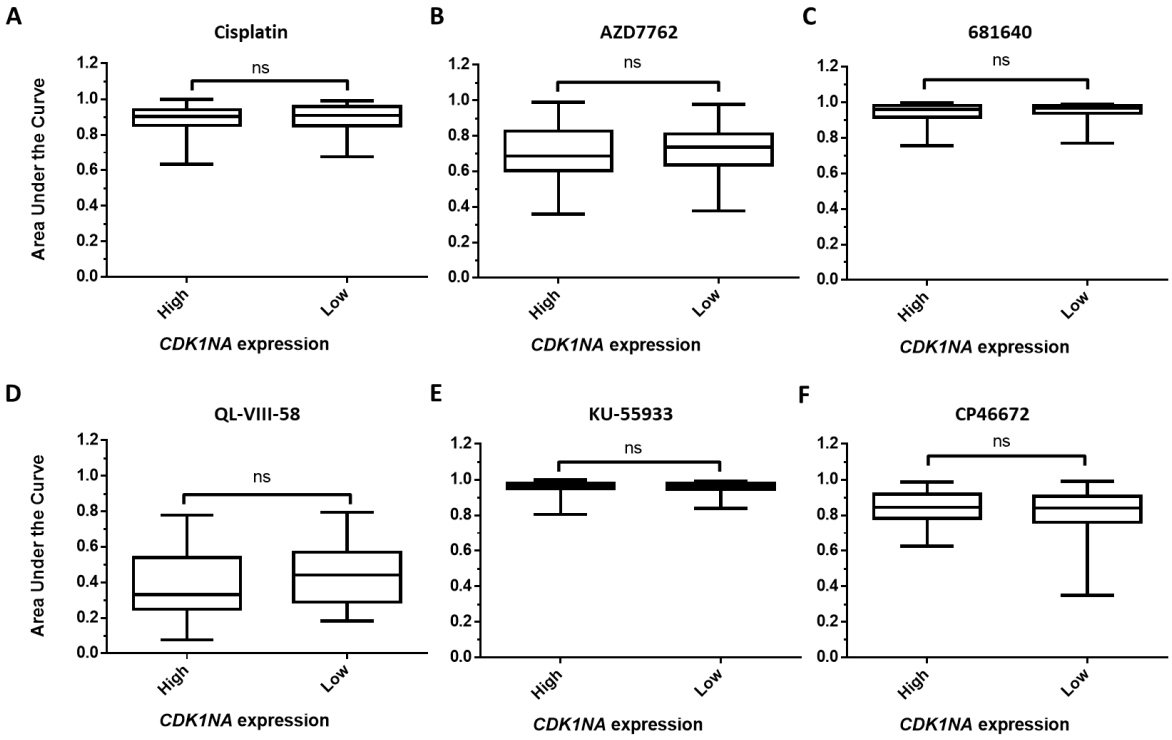


Figure 4.20 Drug response of patients with high or low CDKN1A expression using Genomics of drug sensitivity in cancer (GDSC) database

Pan-cancer patient data extracted from GDSC to identify correlation between high or low expression of CDKN1A and response to stated drugs. High and low *CDKN1A* expression determined by dividing the mean. Area under the curve determined by dose-response to stated drug, Box plots show area under the curve when patient are treated with drug; A) Cisplatin, B) AZD7762 (CHK1 inhibitor), C) 681640 (WEE1/CHK1 inhibitor), D) QL-VIII-58 (ATR/mTOR inhibitor), E) KU-55933 (ATM inhibitor), F) CP46672 (ATM inhibitor). Statistical significance calculated using Students t-test. ns = not significant.

4.3 Discussion

The primary aim of this chapter was to investigate the effect of a set of DDR inhibitors on the cell viability of a panel of chemo-naive and chemo-resistant TNBC cell lines. The output of this study would allow the identification of patterns of cross-resistance or acquired vulnerability and assess the options of the DDR inhibitors as a second line therapy. Identification of a pattern of cross-resistance to inhibitors of the ATR-CHK1-WEE1 axis in a cisplatin resistant cell line, warranted an investigation of $p21^{CIP1/WAF1}$ as a mechanism of drug resistance, establishing a secondary aim of this chapter.

4.3.1 Drug profiling to determine cross-resistance and acquired vulnerability patterns to DNA damage response inhibitors

Drug profiling provided a broad overview of the variability in the response of each of the cell lines to the panel of drugs, and identified patterns of cross-resistance and acquired vulnerability in the chemo-resistant cell lines. It was observed that cross-resistance profiles differed between drugs that target the same protein. These differences in drug profiles could be a result of different drug uptake mechanisms, activity of efflux transporters, or the specificity and efficacy of the drug.

For example, the Δ method identified a positive correlation between eribulin and paclitaxel to MK-8776 (Table 4.2). Cross-resistance to MK-8776 was observed in HCC1806^{r}ERI^{50}, MDA-MB-468^{r}PCL^{20} and HCC1806^{r}PCL^{20}, which differed to the observed response to the other CHK1 inhibitors, and each of these cell lines demonstrated increased MDR1 expression (Figure 4.1D, Figure 3.17). It has previously been shown that MK-8776 can restore sensitivity of chemotherapeutics that are substrates of MDR1 in cells overexpressing MDR1 (Cui *et al.*, 2019). It does this by binding to high expressing MDR1, stimulating the ATPase activity of the transporter, which has been hypothesised to competitively limit the uptake of substrates of MDR1 leading to the inhibition of efflux function (Cui *et al.*, 2019). Literature also shows that substrates of MDR1 can stimulate the ATPase activity of MDR1 transporter, therefore, MK-8776 could be predicted to be a substrate of MDR1.

Alternatively, tariquidar, a third-generation MDR1 inhibitor, has been found to stimulate the MDR1 ATPase activity, but not be transported (Loo *et al.*, 2012). Further investigation as to whether MK-8776 is effluxed by MDR1 would need to be conducted to elucidate the role of MDR1 in MK-8776 resistance. This can be investigated through a combinational assay of MK-8776 and a MDR1 inhibitor, such as zosuquidar, to determine if the cells are sensitised to MK-8776 activity, or through

the use of a UIC2 shift assay (Park *et al.*, 2003). If sensitisation is observed, increased expression of MDR1 could be a candidate biomarker for MK-8776 resistance. If true, this could explain the difference in cross-resistance patterns seen between MK-8776 and the other CHK1 inhibitors investigated here.

It was found that intrinsic resistance to DDR inhibitors in the chemo-naive cell lines had an impact on whether the chemo-resistant cell lines was considered drug-resistant. For example, MDA-MB-468 demonstrated intrinsic resistance to the CHK1 inhibitors, and although the GI_{50} values, determined in the MDA-MB-468 chemo-resistant cell lines to the CHK1 inhibitors, were higher than the other cell lines in the panel, these were not classed as cross-resistant (Figure 4.6, 4.1). These differences can be explained as a result of the heterogenetic nature of TNBC, as certain genetic phenotypes can result in different levels of intrinsic drug resistance. To address this, the absolute drug activity profiles of the cell lines were investigated using the Δ method (section 4.2.5). Within context of the chemo-naive cell lines they were derived from, these pre-clinical models can be used to determine the emergence of resistance through consideration of changes of variants within the cell populations, a method addressed in chapter 5.

Consideration of the absolute drug activity profiles using the Δ method, identified correlations that could indicate whether a drug could be considered as an appropriate second line therapy. A strong positive correlation was identified between the drug activity profiles of eribulin and paclitaxel (Table 4.2). Whilst cross-resistance was seen between the eribulin and paclitaxel resistant cell lines to these drugs, the strong positive correlation indicates that this trend goes beyond these cell lines (Figure 3.13, 3.15, Table 4.2). These data suggest that as resistance in the cell lines increases to treatment with eribulin, there is an observed increase in resistance to the treatment of paclitaxel, and vice versa, indicating that these would not be an appropriate second line therapy following the development of resistance to each other.

It must be noted that some significant positive correlations were also seen between the drugs that are considered to be substrates of MDR1 which included; BI2536, eribulin, doxorubicin, olaparib, paclitaxel, rucaparib and tozasertib (Wu, 2013; Michaelis *et al.*, 2014; Vaidyanathan *et al.*, 2016). Whilst high expression of MDR1 may be responsible for cross-resistance patterns seen in some of the cell lines, this correlation may go beyond MDR1 efflux. For example, HCC38^{r}ERI^{10} demonstrated cross-resistance to BI2536, with a higher RF than to eribulin itself, despite not demonstrating increased expression of MDR1 compared to HCC38. The statistically significant positive correlations identified, in combination with the cross-resistance patterns, show that as resistance increases to

one, resistance increases to the other, suggesting that each of these would be poor choice as a second line therapy after resistance has developed to the aforementioned.

The Δ method analysis also identified an inverse relationship between the drug activity profiles of CHK1 (rabusertib, MK-8776, SRA737, Prexasertib), ATR (ceralasertib, berzosertib), WEE1 (adavosertib) and CHK2 (CCT241533) inhibitors versus cisplatin and the ATM (AZD0156), RAD51(B02) and PLK1/2 (SBE13) inhibitors (Table 4.2). It indicates that as resistance increases to CHK1, ATR, WEE1 or CHK2 inhibitors, resistance also increases to each other, but decreases to cisplatin, and the ATM, RAD51 and PLK1/2 inhibitors. These data suggest that after resistance has occurred to either inhibitors of CHK1, ATR, WEE1 and CHK2, sequential treatment with the aforementioned would not be an appropriate second line therapy. However, after resistance has developed to these inhibitors, it suggests that treatment with either cisplatin or the inhibitors of ATM, RAD51 or PLK1/2 may prove to beneficial. It was odd to note that the activity of the CHK2 inhibitor negatively correlated with the activity of the ATM inhibitor, as ATM is known to directly phosphorylate CHK2, and is required for CHK2 activation. This polarity in drug activity could be explained through the differences in the known downstream targets of ATM and CHK2.

Additionally, it was noted that cisplatin demonstrated a positive correlation with the activity profile of RAD51 (B02) and PLK1/2 (SBE13) inhibitors, indicating that as resistance occurred to one, resistance was observed in the other (Table 4.2). This suggests that RAD51 and PLK1/2 inhibitors would not be an appropriate second line therapy following cisplatin resistance, and vice versa. Furthermore, the data shows an inverse relationship between the drug activity profiles of cisplatin and the CHK1 inhibitors (Table 4.2). It is interesting to note that cisplatin resistance positively correlates with resistance to the RAD51 inhibitor and negatively correlates with CHK1 inhibitors. As previously mentioned, RAD51 is required for the repair of dsDNA breaks, specifically through HRR, whilst CHK1 is predominantly required for the repair of ssDNA breaks and at stalled replication forks (section 1.4). However, repair of dsDNA using ssDNA as a template, a method commonly used in CRISPR/Cas9-mediated gene editing, is RAD51-independent (Gallagher *et al.*, 2020). It could suggest that resistant mechanisms induced during cisplatin resistance can make cell more reliant on additional repair pathways, other than those that repair dsDNA. Furthermore, cisplatin induces ICLs which can trigger stalling of replication in S-phase. CHK1 is known to have a role at stalled replication forks during S-Phase, and can induce the DDR. This data is consistent with literature in which CHK1 inhibitors have been shown to enhance cisplatin antitumor activity, and overcame cisplatin resistance in SCLC preclinical models (Hsu *et al.*, 2019).

	Cisplatin	CHK1i	SBE13	B02
Cisplatin	+	-	+	+
CHK1i	-	+	-	-

Figure 4.21 Drug activity profiles of Cisplatin and CHK1 inhibitors (CHK1i)
Pearson correlation of statistical significance calculated via the Δ method for cisplatin, CHK1i, SBE13 and B02. Plus (+) indicates a positive correlation and minus (–) indicates negative correlation between the drug activity profiles of stated drugs. Red indicates poor choice; green indicates good choice prediction of second-line therapy following drug resistance to the drugs on the left of the table.

Some of the most interesting observations were made when considering the development of cross-resistance to the DDR inhibitors within context of the chemo-naive cell line the chemo-resistant cell lines were derived from. It was found that the majority of the chemo-resistant cell lines demonstrated cross-resistance to the PARP inhibitors; olaparib and rucaparib, as well as to the chemotherapeutic agent; doxorubicin (4.14). Whilst resistance to these three drugs have been linked to high expression of MDR1, only four cell lines in the panel demonstrated an increased expression of MDR1 (Figure 3.17) (Vaidyanathan *et al.*, 2016). Furthermore, as mentioned, the Δ method identified a positive correlation between each of the drug activity profiles for doxorubicin, olaparib and rucaparib (Table 4.2). Taken together, these data suggest that olaparib, rucaparib and doxorubicin would be a poor choice as a second-line therapy either after the emergence of chemo-resistance, or sequentially following each other, regardless of the patients MDR1 status. This is an important observation because, as previously mentioned, anthracyclines such as doxorubicin are still often routinely used in the clinic as a first line treatment for TNBC (section 1.6)(Bergin and Loi, 2019). Furthermore, with PARP inhibitors in clinical trials for the treatment of TNBC, and approved for other cancer types (ovarian), it is important to understand both how resistance occurs, and an appropriate second line therapy to these agents (Mehanna *et al.*, 2019; Robson *et al.*, 2019). Here we have shown in pre-clinical cell line models that doxorubicin, olaparib and rucaparib will not be appropriate secondary therapies following the development of chemo-resistance.

The majority of the chemo-resistant cell lines demonstrated either acquired vulnerability, or no statistically significant fold change to the treatment with CHK2 (CCT241533), RAD51 (B02) or a PLK1/2 (SBE13). However, each of these drugs are currently not in clinical trials. As previously mentioned CCT241533 demonstrated a good selectively profile for CHK2, but required further medicinal chemistry development to obtain a compound with appropriate pharmacological properties (section 1.4.2) (Anderson *et al.*, 2011; Caldwell *et al.*, 2011; *Chk2 Inhibitor Programme | Commercial Partnerships | Cancer Research UK*, 2020). Whilst B02 has been shown to effectively inhibit DSB induced HRR in human cells, and supress ionizing radiation induced RAD51 foci formation, it has not progressed into clinical trials (Budke *et al.*, 2016). Finally, SBE13 is still considered to be in the preclinical development stage and has also not progressed to clinical trials

(Keppner *et al.*, 2009; Eckerdt, 2011). Whilst these drugs are not in clinical trials, the inhibition of these targets could provide some therapeutic benefit as an appropriate second-line therapy after chemoresistance has emerged.

Whilst the observations in this chapter hypothesise as to whether a drug may be an appropriate second line therapy after chemo-resistance has occurred, this has only been observed in a small set of TNBC pre-clinical cell lines models. These must be validated if they are to become useful for application in the clinic. One such method of validation is to obtain TNBC patient tumour samples pre-and post-relapse after chemotherapy treatment. From this, the response to the drug in question can be determined to assess if the drug is suitable as a second line therapy treatment.

4.3.2 Validation of $p21^{CIP1/WAF1}$ as a mechanism of drug resistance to CHK1 inhibitors and cisplatin in $HCC38^{r}CDDP^{3000}$.

The cisplatin resistant cell line, $HCC38^{r}CDDP^{3000}$, had been found to demonstrate strong cross-resistance to each of the CHK1 inhibitors (rabusertib, MK-8776, SRA737 and prexasertib), as well as to both the ATR inhibitors (ceralasertib and berzosertib), and the WEE1 inhibitor (adavosertib), but acquired vulnerability to the ATM inhibitor (AZD0156) (Figure 4.1 -4.3). In contrast to this, the Pearson correlations generated from the Δ method, identified that cisplatin and CHK1, ATR, WEE1 inhibitors to have a negative correlation. Given this juxtaposition, further investigation was warranted.

Cisplatin is known to induce S-phase cell cycle arrest, which is reliant on the activation of the DDR pathways through CHK1 (Zhang *et al.*, 2008; Wagner and Karnitz, 2009; Barr *et al.*, 2017). CHK1 activity leads to the inhibition of CDC25 and prevents activation of CDK1/2, required for the cell to enter G2 (Thompson and Eastman, 2013). It was found that a subset of cancer cell lines was sensitive to CHK1 inhibitor, MK-8776, as a monotherapy due to CHK1 requirement for CDK2 activation in S phase, and those that showed resistance had increased activation of CDK2 in S phase (Sakurikar *et al.*, 2016). Further to this, it has also been seen that $p21^{CIP1/WAF1}$ induction inhibited endoreduplication through direct cyclin E/CDK2, and high $p21^{CIP1/WAF1}$ levels mediate G1 arrest via CDK inhibition (Stewart and Leach, 1999). Here it was hypothesised that by-pass of S-G2/M checkpoint resulted in increased $p21^{CIP1/WAF1}$, due to accumulation of DNA damage in S phase, leading to inhibition of cyclin E/CDK2 in G1 to allow for cisplatin induced DNA repair. To this end, $p21^{CIP1/WAF1}$ expression was validated as a mechanism of cisplatin/ CHK1 resistance.

Initial studies showed that both $p21^{CIP1/WAF1}$ protein expression and amplification of $p21^{CIP1/WAF1}$ through western blot and RT-qPCR respectively, were seen in $HCC38^{r}CDDP^{3000}$ compared to HCC38 (figure 4.15, 4.16). This suggested that increased transcription of the *CDKN1A* gene resulted in higher protein expression found in $HCC38^{r}CDDP^{3000}$ compared to HCC38. To validate the importance of elevated expression of $p21^{CIP1/WAF1}$ as a resistance mechanism, $p21^{CIP1/WAF1}$ activity was reduced in $HCC38^{r}CDDP^{3000}$ via siRNA knockdown to determine if this reduced resistance to the treatment of cisplatin or a CHK1 inhibitor; rabusertib. Knockdown of $p21^{CIP1/WAF1}$ in $HCC38^{r}CDDP^{3000}$ cells did not reduce resistance to cisplatin or rabusertib (Figure 4.19). This may suggest that $p21^{CIP1/WAF1}$ knockdown alone was not sufficient alone to reduce resistance to cisplatin or to rabusertib, but additional resistance drivers may also play a part in the mechanism of resistance. Genomic and transcriptome characterisation of $HCC38^{r}CDDP^{3000}$ through exome sequencing or RNA sequencing (RNA-seq) analysis; could identify additional resistance drivers and provide an insight which may be context dependent and multi-factorial.

There were several limitations to the experiment which may have influenced the result. Simultaneous siRNA knockdown in cells in both 6 well and 96 well formats were conducted to visualise if sufficient knockdown had occurred in the dose response. Scaling up the siRNA knockdown into 6-well plates was required in order to yield sufficient detectable protein. This may have been problematic, as knockdown may not have necessarily occurred in both 6-well and 96-well plates as per the assumption taken. One way in which this could be overcome would be to develop a $p21^{CIP1/WAF1}$ -specific cell-based enzyme-linked immunosorbent assay (ELISA) to confirm knockdown in a 96-well plate format. Further to this, the use of siRNA is only a short-term knockdown of a target protein, and can be dependent on protein turnover. The turnover of the protein may affect the duration and efficiency of the knockdown. The turnover of $p21^{CIP1/WAF1}$ could have been measured using a cycloheximide chase assay in both the HCC38 and $HCC38^{r}CDDP^{3000}$ cell lines to 1) to determine $p21^{CIP1/WAF1}$ stability in each of the cell lines and 2) and compare differences in the rate of protein degradation. Stable knockouts may have been a more appropriate method for this experiment, either using short hairpin RNA (shRNA), or a complete knockout using the clustered regularly interspaced short palindromic repeats – associated protein 9 (CRISPR-cas9) system.

To investigate clinical relevance of increased $p21^{CIP1/WAF1}$ expression, the response of cancer patients cell lines to cisplatin, and inhibitors of CHK1, ATR, WEE1, when *CDKN1A* expression was high or low, was investigated using the publicly available data in the GDSC database. No significant difference was seen with the response of the patient derived cell lines to any of the drugs when

CDKN1A expression was high or low (Figure 4.20). Nevertheless, there were many limitations in the analysis performed here. Firstly, what was considered high or low *CDKN1A* expression was arbitrary and based only on the data available for cell lines which also had drug response data. Further to this, not every cancer cell line had a combination of both *CDKN1A* expression data and each of the drugs, so linking the data of a cell line across each of the drugs was not possible. Additionally, the cancer cell lines do not have acquired drug resistance. Whilst some may have demonstrated intrinsic resistance to the drug, the window of response to the drug was small, and cannot be likened to cell lines with acquired drug resistance. Lastly, the data set of cell lines with *CDKN1A* expression and a matching response to the drugs was too small to only consider human breast cancer cell lines, and the analysis was performed using pan-cancer patient derived cell lines. This analysis may not have been appropriate as cancers, and tumour types differ in the development of drug resistance, as this can be context specific.

In summary, the data presented in this chapter has addressed whether a panel of DDR inhibitors are an appropriate second line therapy after acquired resistance has occurred to chemotherapy agents in these cell line models. Both patterns of cross-resistance and absolute drug activity profiles were considered in this analysis. The chapter also validated that $p21^{CIP1/WAF1}$ alone is not involved in a mechanism of drug resistance to cisplatin or rabusertib. Pre-clinical models can be used to determine genomic changes that occur in the emergence of drug resistance. Chapter 5 will analyse the exomic landscape of the chemo-resistant cell lines to identify variants that could be driving drug resistance.

Chapter 5

Identification of candidate drivers of drug resistance using whole exome sequencing

5. Identification of candidate drivers of drug resistance using whole exome sequencing

5.1 Introduction

Resistance to chemotherapy agents can occur through a number of mechanisms such as; changes in drug target or drug target levels, changes in drug metabolism, drug compartmentalisation, increased drug efflux, decrease in drug uptake, increased resistance to apoptosis and increased DNA damage repair (section 1.3). Often, these mechanisms are a result of genetic aberrations, such as structural alterations of the genome, or variants in genes which impact the encoded protein sequence structure and function. It was therefore reasoned that drug resistance in the chemo-resistant TNBC cell lines may be a result of the introduction of variants in the cell line models.

Next-generation sequencing techniques, such as whole genome and whole exome sequencing, have allowed for the identification of changes in the genomic landscape of cancer cells, including cancer cells demonstrating resistance to drugs. Whilst whole genome sequencing (WGS) captures the full genome, a more cost-effective approach lies with whole exome sequencing (WES). The exons make up approximately 1% of the entire human genome, and WES captures these coding sequences, which includes approximately 22,000 genes and their flanking regions (Meyerson, Gabriel and Getz, 2010; Suwinski *et al.*, 2019). Analysis of WES covers actionable areas of the genome whereby variants can be identified, which could be predicted to affect protein function and demonstrate a disease-causing phenotype.

WES has successfully been used to identify mediators of acquired resistance to cancer drugs in patients. For example, WES was used to detect clonal selection and outgrowth of patient melanoma tumour samples with acquired resistance to anti-programmed death 1 (PD-1) blockade immunotherapy (Zaretsky *et al.*, 2016). This identified resistance associated loss of function variants in the genes which encode for interferon-receptor-associated Janus kinase 1 (JAK1) or Janus kinase 2 (JAK2). It was found that these truncating variants resulted in a lack of response to interferon gamma, including an insensitivity to its antiproliferative effects on cancer cells (Zaretsky *et al.*, 2016). Clinical trials have also adopted WES as a technique to understand biological context for the use of a candidate drug. For example, in a phase 2 clinical trial, it was questioned in 50 patients with metastatic, castration-resistant prostate cancer, if they would respond to the PARP inhibitor Olaparib (NCT01682772). Exome sequencing identified that patients which had defects in DNA-repair genes led to a higher response rate to PARP inhibition (Weinstein, 2013). WES has also been used to further understand heterogeneity of patient tumour samples and the contribution this has

to a drug resistant phenotype. For example, a single cell DNA exome sequencing of 20 TNBC patients, during neoadjuvant chemotherapy (NAC), identified that resistant genotypes were pre-existing and adaptively selected for by NAC (Kim *et al.*, 2018).

Moreover, WES has been used to identify resistance mechanisms in cell line models which have been developed to have acquired drug resistance. Beauchamp *et al.*, 2014, used WES to analyse 645 genes in lung squamous cell carcinoma models resistant to dasatinib. Here they identified that a variant in the *discoidin domain containing receptor-2* (*DDR2*), as well as a loss of function splice site variant in *neurofibromin* (*NF1*) gene as drivers of a resistance mechanism to dasatinib.

WES has also developed to become a powerful tool for identifying genomic changes, or unique molecular signatures, termed biomarkers, which can be used to stratify patient populations and inform therapeutic decisions regarding candidate drugs for treatment (section 1.5). WES analysis can allow for molecular characterisation of cancer cells, prognostics and monitoring (Johannessen and Boehm, 2017).

Now routinely used in the clinic, biomarkers can be used in a diagnostic and prognostic fashion as well as predictive for response to drugs. For example, Osteopontin, an extracellular structural protein, has been found to be overexpressed in tumours and serum of ovarian cancer patients and acts as a diagnostic biomarker for the progression of ovarian cancer (Brakora *et al.*, 2004; Song *et al.*, 2008). Recently, the drug Larotrectinib, was approved by the Food and Drug Administration (FDA), as the first molecularly targeted therapeutic drug based on the tumour biomarker of a chromosomal translocation of the *NTRK* genes. This is the first drug to be based on an array of cancer types, not just the tumour origin in the body (*FDA Approves First Targeted Therapeutic Based on Tumor Biomarker, Not Tumor Origin - American Association for Cancer Research (AACR)*; Drilon *et al.*, 2018). Variants in *phosphatase and tensin homolog* (*PTEN*), a tumour suppressor gene, are routinely identified as a biomarker of poor patient prognosis (McCabe, Kennedy and Prise, 2016). There are also many drug predictive biomarkers such as; BRCA1/BRCA2 mutants as good indicators of response to PARP inhibitors (Olaparib), and ALK mutants as good indicators of response to ALK targeted drugs (Crizotinib, Cetitinib, Alectinib and Brigatinib) (Michels *et al.*, 2014; Ahmadzada *et al.*, 2018).

Large scale efforts to analyse patient clinical tumour samples for molecular aberration's in DNA, RNA, protein and epigenetic levels have been curated in open-source databases. One of the first to do so were The Cancer Genome Atlas (TCGA) pan-cancer analysis project, whilst more recently in combination with this project, and the International Cancer Genome Consortium (ICGC), was the Pan-Cancer Analysis of Whole Genomes (PCWAG) (Group, 2010; Weinstein, 2013; Campbell *et al.*,

2020). These databases can be used to evaluate drivers of cancer and also mediate the investigation of drug resistance in patient tumour samples. Using these databases, experimentally identified biomarkers can be considered in the context of known clinically obtained patient variants and gene expression. Therefore, WES is a powerful tool for identifying resistance mechanisms and biomarkers in both preclinical models and in patients which have developed acquired drug resistance.

In the previous chapters, cross-resistance to both chemotherapeutic agents and DDR inhibitors in the chemo-resistant cell lines have been discussed. The aim of this chapter was to identify exomic alterations in the chemo-resistant TNBC cell lines which could be driving the resistant phenotype. Here, it was aimed to further characterise each chemo-resistant cell line in respect to the chemo-naive cell line it was derived from, but also to compare between cell lines which have been developed to have resistance to the same chemotherapeutic agent. Identified candidate biomarkers were compared to the TCGA to identify those that have clinical relevance. This analysis using WES data led to the identification of clinically relevant candidate resistance biomarkers, or drivers of a drug resistance phenotype, which could be taken forward for further experimental validation.

5.2 Results

5.2.1 Identification and characterisation of variants in chemo-naive and chemo-resistant TNBC cell lines

5.2.1.1 Identification of variants using exome sequencing in the chemo-naive and chem-resistant TNBC cell lines

To identify variants in genes which could be driving resistance in the chemo-resistant population, each cell line in the panel underwent WES using the transposase-based method by Illumina. Illumina HisSeq2000 was used with an input of nucleotide paired end sequences with an output of nucleotide paired ends reads in FASTQ format. This step was performed by the Genomics Core Facility, Philipps-University, Marburg, Germany. Each cell line was sequenced in two separate lanes, and the data combined. The FASTQ output from the sequencing is a FASTA file which includes the sequence of bases, termed reads, for overlapping loci as well as quality control for the sequencing performed. The total number of reads identified for each cell line is outlined in Figure 5.1 (numerical data in Appendix A4). Both MDA-MB-468 and HCC1806 chemo-naive and chemo-resistant cell lines had a similar number of reads across the groups, whilst the HCC38 cell lines showed some variability. Both HCC38^{r}CDDP^{3000} and HCC38^{r}DOX^{40} had a lower number of reads compared to both HCC38 and the MDA-MB-468 and HCC1806 cell lines, whilst HCC38^{r}GEM^{20} had a higher number of reads.

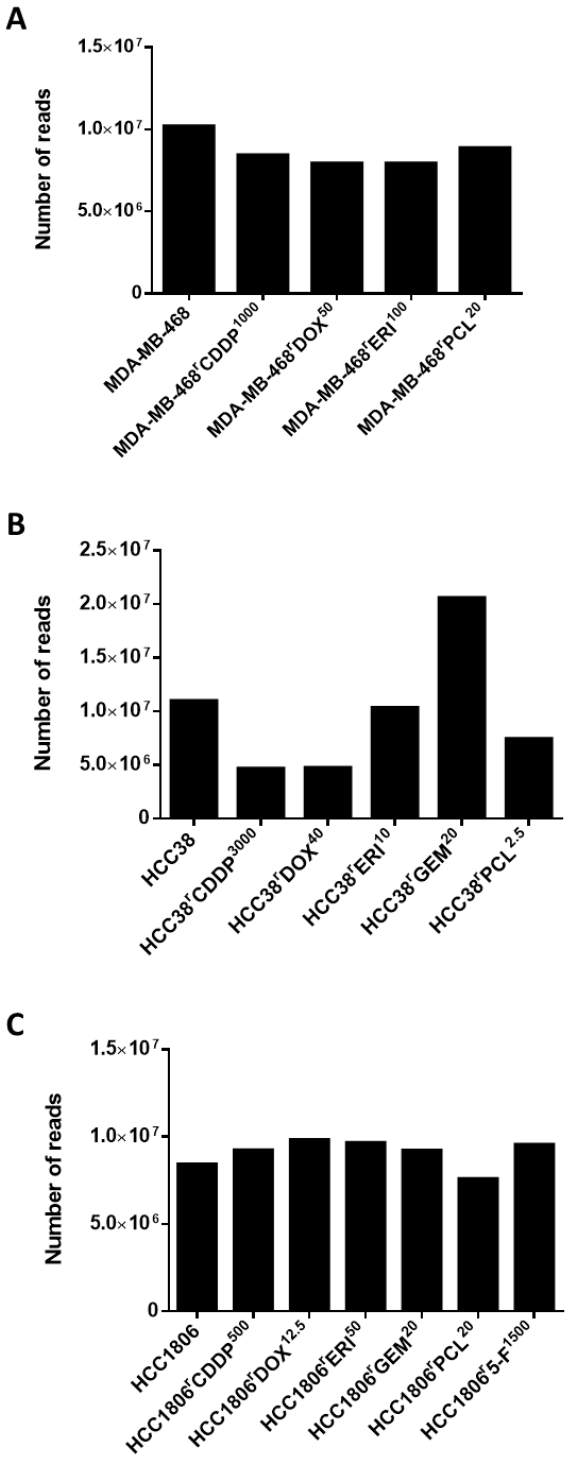


Figure 5.1 Number of reads identified from exome sequencing for chemo-naive and chemo-resistant TNBC cell lines
Number of paired-end reads identified in FASTQ file after exome sequencing was performed on each of the chemo-naive and chemo-resistant cell lines A) MDA-MB-468, B) HCC38, C) HCC1806. Numerical data in Appendix A4.

An in-house variant calling pipeline was partly written and designed by a previous PhD student; Miguel Julia, a current PhD student; Magdalena Antczak, as well as contribution from the work in this chapter. The steps for the calling of variants are outlined in Figure 5.2. Briefly, the FASTQ files were analysed using FASTQC for quality control. The reads length for each cell line ranged from 35-101 bases with a percentage GC content of 46-47%, and none were flagged as poor quality (Appendix A4). The reads then underwent trimming to remove any present sequencing adaptors using Trimmomatic, before the reads were mapped and aligned to the genome reference consortium human build 37 (GRCh37) using the Burrows-Wheeler Alignment Tool (v.0.7.12) and Picard tools SortSam (v.1.112)(Burrows and Wheeler, 1994; Li *et al.*, 2009; Church *et al.*, 2011; Bolger, Lohse and Usadel, 2014). Reads that had identical chromosome start and end, termed PCR duplicates, were marked and removed to reduce biased variant calling using Picard Tools MarkDuplicates. The two files from the two separate lanes for each cell lines were then merged, and PCR duplication removal repeated again. The reads then under-went realignment and base score recalibration to ensure insertions and deletions (INDELs) were correctly positioned using Picard Tools BuildBamIndex. From this, the single nucleotides variants (SNVs) and INDELs were identified and annotated with a Variant Effect Predictor (VEP)(McLaren *et al.*, 2016). Using the coordinates of the variants and nucleotide changes, VEP was used to determine the genes and transcripts affected, the location of the variant, the consequence of a variant, if the variant is already known in databases, as well as a variant severity prediction score from programmes including SIFT and PolyPhen-2 (Ng and Henikoff, 2003; Adzhubei, Jordan and Sunyaev, 2015). The final file contained the identified annotated variants in a Variant Calling Format (VCF) file.

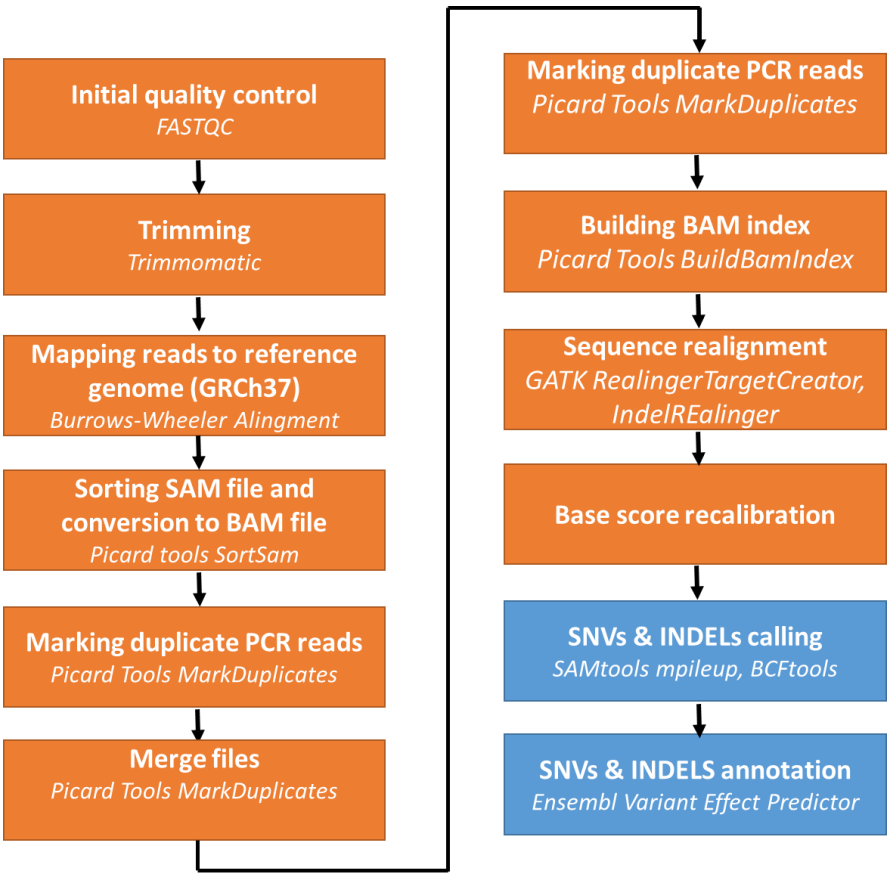


Figure 5.2 Variant calling pipeline

Steps in the in-house built computational pipeline to call variants in the chemo-naive and chemo-resistant TNBC cell lines. Input is the FASTQ file containing reads from exome sequencing analysis, with the output of called variants with variant effect predictor annotation in a VCF file for each of the chemo-naive and chemo-resistant cell lines. Each step is annotated with action performed and the tools used.

The total number of variants for each cell line is shown in Figure 5.3 (numerical data in Appendix A5). A low number of variants are called in HCC38^{r}CDDP^{3000} and HCC38^{r}DOX^{40}, 192393 and 188481 respectively, compared to the other cell lines in the panel, which have variants in the range of 291386 – 398522 called. An exception was also seen in HCC38^{r}GEM^{20}, which had the highest number of variants (630942 variants). The pattern of the number of variants called was observed to be the same as the number of reads available (Figure 5.1). Normalising the variants called against the number of reads available showed that between 3.07% and 4.14% of the number of reads resulted in called variants. This suggests that the calling of variants is dependent on the number of reads available.

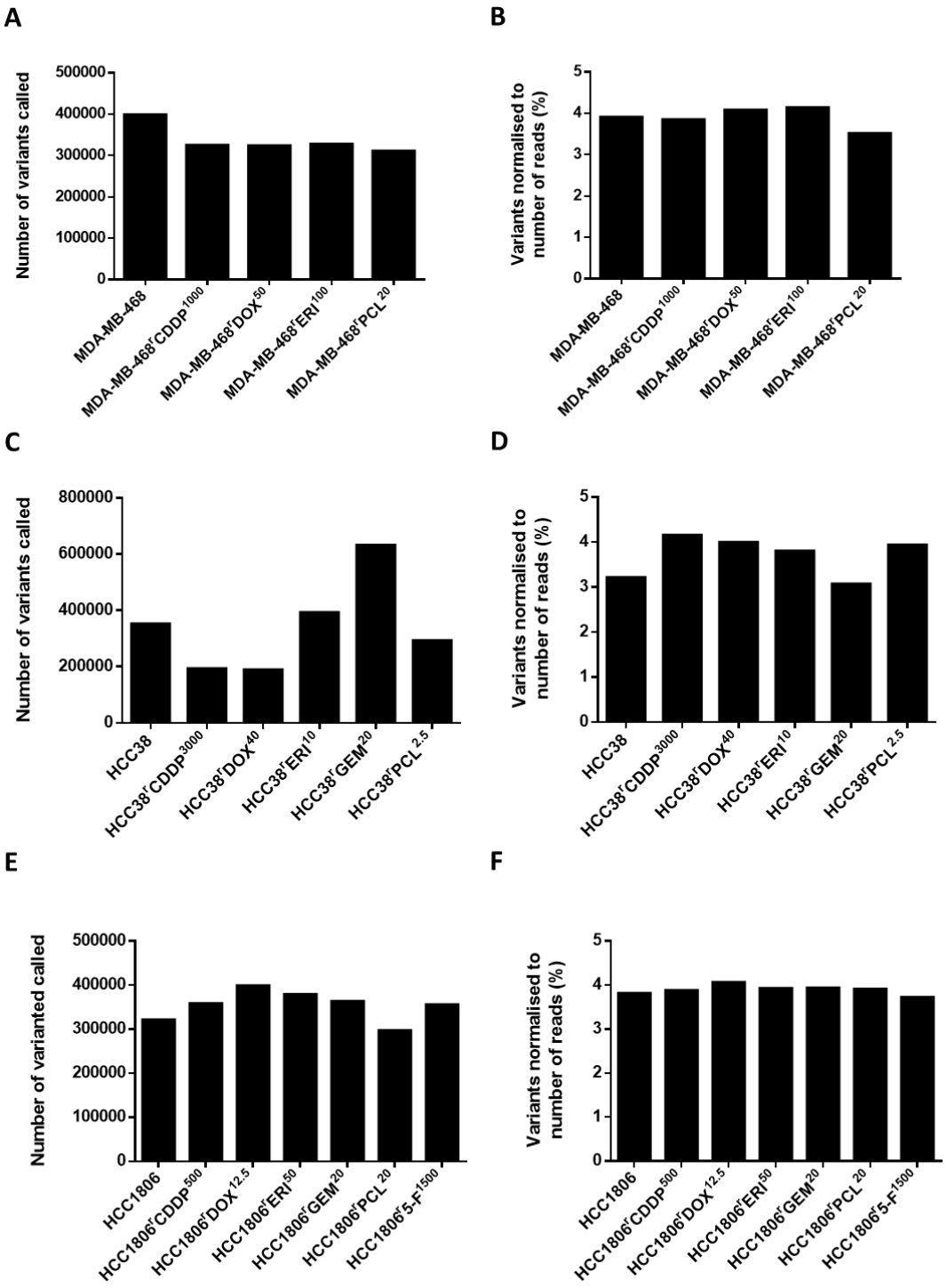


Figure 5.3 Number of variants called in the chemo-naive and chemo-resistant TNBC cell lines

Number of variants called after reads from exome sequencing are aligned against the reference genome (GRCh37), and percentage of called variants normalised against the number of identified reads in in each of the chemo-naive and chemo-resistant cell line A-B) MDA-MB-468, C-D) HCC38, E-F) HCC1806 respectively.

5.2.1.2 Filtering for high confidence variants in the chemo-naive and chemo-resistant TNBC cell lines

Having called the variants in each of the chemo-naive and chemo-resistant TNBC cell lines, the variants next underwent a filtering process in order to identify high confidence somatic variants, as explained below. An in-house variant filtering pipeline was written and partly designed by PhD student Magdalena Antczak, as well as contribution from the work in this chapter. The steps for the filtering of variants are outlined in Figure 5.4.

First, each of the called variants (SNVs and INDELs) are checked for quality or coverage. Step one removed variants with a base call accuracy below 99.9 %, corresponding to a Phred quality score of 30. The Phred quality score is assigned to a given variant during the sequencing process, which is a probability estimate of that base call being a true nucleotide. The likelihood of the accuracy of the base call is then reflected in the Phred Score. Further to this, variants which have less than 10 reads supporting the base call were removed, and of those remaining, variants which have less than three reads supporting the variant call were removed. Step two removed common, germ-like variants in order to obtain somatic variants in the cell lines. This was done by removing variants present at a frequency ≥ 0.001% in the genome aggregation database (gnomAD), a variant data set of 125, 748 exome sequences, and 15,708 whole-genome sequences of unrelated individuals (Karczewski *et al.*, 2019). Often, variants which are considered common in the human population, may have a role in cancer. Following the work of Ghandi et al 2019, if any variant which was identified to a frequency of ≥ 0.001% and were seen in either a) The Cancer Genome Atlas (TCGA) database in at least three samples or b) the Catalogue Of Somatic Mutations In Cancer (COSMIC) database in at least ten samples, these were added back to the list of variants called in the cells lines (Bamford *et al.*, 2004; Weinstein, 2013; Ghandi *et al.*, 2019). Finally, step three removed variants which were not found in the protein sequence, such as upstream variants, as these are outside of the confident sequencing scope. This identified the high confidence somatic variants which were taken forward for analysis.

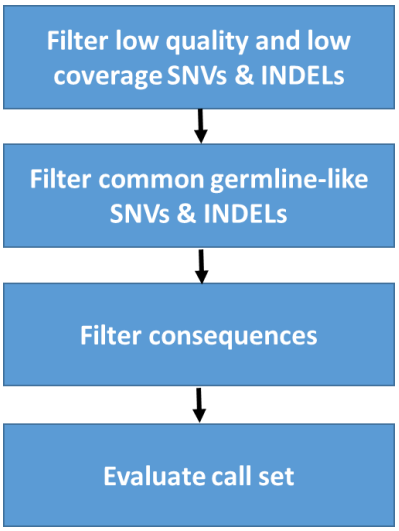


Figure 5.4 Variant filtering pipeline
Steps in the in-house built computational pipeline to filter variants in the chemo-naive and chemo-resistant TNBC cell lines. Input is the called variants with variant effect predictor annotation and the output are high confidence somatic variants for each of the chemo-naive and chemo-resistant cell lines in VCF files.

The removal of the variants as they were processed through the variant filtering is shown in Figure 5.5 (numerical data in Appendix A5). The removal of variants is uniform between MDA-MB-468 and HCC1806 chemo-naive and chemo-resistant cell lines. The HCC38 chemo-naive and chemo-resistant cell lines show a less uniformed filtering, with HCC38^{r}CDDP^{3000} and HCC38^{r}DOX^{40} showing a steeper decrease in variants lost during removal for step one; poor quality and step two; base and variant coverage (Figure 5.5B).

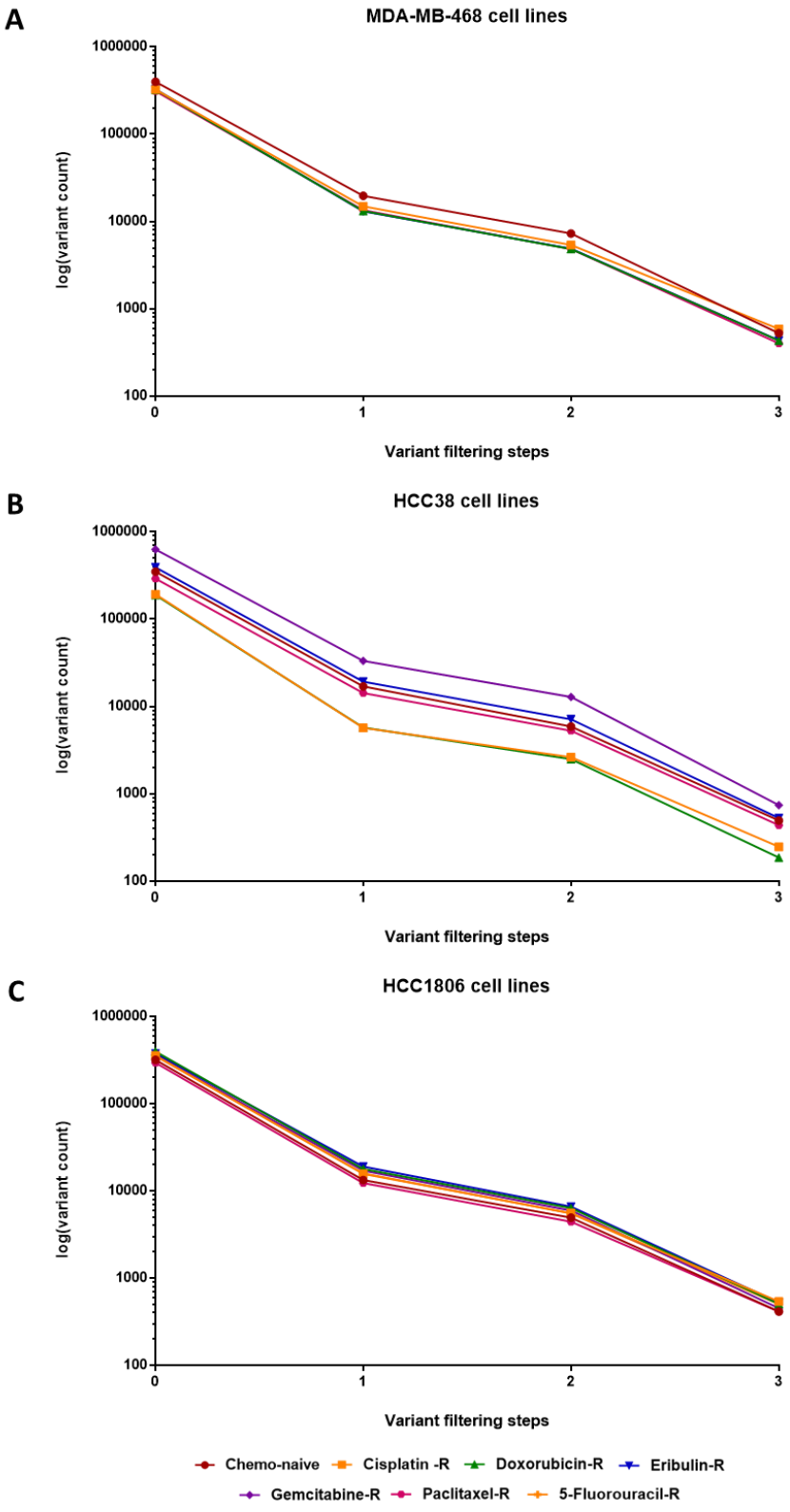


Figure 5.5 Number of variants called at each step of the variant filtering pipeline for chemo-naive and chemo-resistant TNBC cell lines Number of variants called at each step of the variant filtering pipeline. Step 0 indicates called variant input, step 1; after filtering low quality and low coverage variants, step 2; after filtering out germline variants, step 3; after filtering out consequences not within the sequencing remit. Variants called in step 3 are considered to be high confidence somatic variants. A) MDA-MB-468 cell lines B) HCC38 cell lines C) HCC1806 cell lines, whereby colours indicate if cell line is chemo-naive or chemotherapeutic agent – R indicates resistance to stated chemotherapeutic agent. Numerical data in Appendix A5.

Having applied the filtering steps to the called variants, high confidence somatic variants in each of the chemo-naive and chemo-resistant cell lines had been established (Appendix A6 – A23). Variants in which a base change results in an amino acid substitution are termed non-synonymous, whilst variants in which a base change does not result in an amino acid substitution are termed synonymous. It was noted that non-synonymous variants were more prevalent in each of the cell lines compared to synonymous variants.

Considering the MDA-MB-468 derived cell lines, MDA-MB-468 and MDA-MB-468^{r}CDDP^{1000} had the most variants called with 525 and 588 respectively. MDA-MB-468^{r}DOX^{50} and MDA-MB-468^{r}ERI^{50} had a similar number of variants called with 433 and 435 respectively. MDA-MB-468^{r}PCL^{20} only had 402 variants called (Figure 5.6A, Appendix A5). Considering the HCC1806 derived cell lines, HCC1806 and HCC1806^{r}PCL^{20} had a similar number of variants called with 414 and 413 respectively. HCC1806^{r}CDDP^{500}, HCC1806^{r}DOX$^{12.5}$, HCC1806^{r}ERI^{50}, HCC1806^{r}GEM^{20} and HCC1806^{r}5-F^{1500} all had a similar number of variants called with 539, 505, 523, 452 and 503 respectively (Figure 5.6A, Appendix A5). The HCC38 derived cell lines, HCC38, HCC38^{r}ERI^{10} and HCC38^{r}PCL$^{2.5}$ had a similar number of variants called with 495, 528 and 436 respectively. Very few variants were called in HCC38^{r}CDDP^{3000} and HCC38^{r}DOX^{40}, 247 and 186 respectively, whilst 742 variants were called in HCC38^{r}GEM^{20} (Figure 5.6A, Appendix A5). This pattern of low number of variants called in HCC38^{r}CDDP^{3000} and HCC38^{r}DOX^{40}, and high number of variants called in HCC38^{r}GEM^{20} was also seen before the variant filtering step, as well as in the number of reads input before alignment to the reference genome (Figure 5.1B, 5.3B and 5.5B). This suggests that the quality of the sequencing was lower for both HCC38^{r}CDDP^{3000} and HCC38^{r}DOX^{40}, whilst higher depth coverage was achieved when sequencing HCC38^{r}GEM^{20} compared to any of the other cell lines in the panel.

Next, using the established high confidence somatic variants, the cell line mutation burden could be calculated. This is defined by the number of somatic mutations (including both SNVs and INDELs) in the coding region per mega base. As to be expected, the pattern of cell line mutation burden for both synonymous and non-synonymous variants in the cell lines reflected that of the number of variants called (Figure 5.6B, Appendix 5A). Finally, the number of genes which contain the called variants were identified; termed mutated genes (Figure 5.6C, Appendix 5A). It was noted that each of the cell lines showed an unequal number of mutated genes to the number of variants called. This shows that multiple variants were identified within the same genes. Of note, MDA-MB-468^{r}PCL^{20} had 402 variants called but these were found to only be present in 246 different genes, and HCC38^{r}PCL$^{2.5}$ had 436 variants called and these were found to only be in 258 different genes.

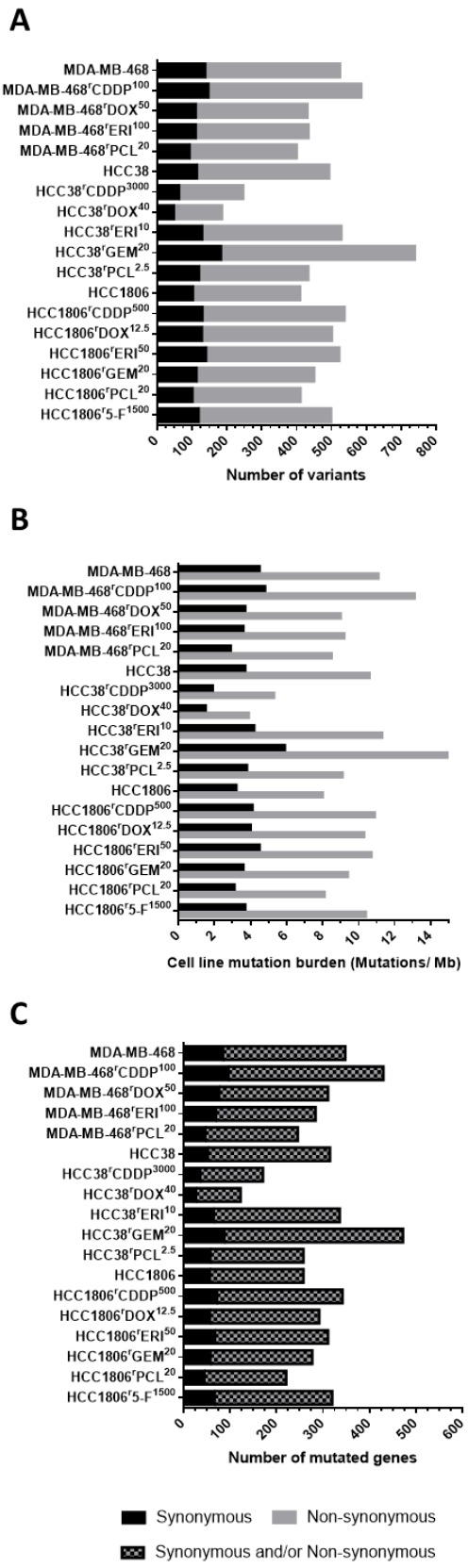


Figure 5.6 Characterisation of high confidence somatic variants in chemo-naive and chemo-resistant cell lines
Basic characterisation of high confidence somatic variants in chemo-naive and chemo-resistant cell lines. A) number of synonymous and non-synonymous variants called in each TNBC cell line. B) Calculated cell line mutation burden (Mutations / Mb) for synonymous and non-synonymous variants in each TNBC cell line. C) Number of genes which contain either synonymous variants or both synonymous and non-synonymous variants (mutated genes) in each TNBC cell line. Numerical data in Appendix A5.

5.2.1.3 Characterisation of variants in the chemo-naive and chemo-resistant TNBC cell lines

Having established the high confidence somatic variants in each of the cell lines, the variants underwent basic characterisation. The number of different variant consequences for each of the cell lines in the panel was calculated (Figure 5.7, Appendix A24). The most common type of consequence in each of the cell lines was missense variants, followed by synonymous variants. Missense variants are considered under the umbrella of non-synonymous variants, whereby a base change results in the single substitution of an amino acid. The next most common consequences seen in each of the cell lines were INDELs and frameshifts. INDELs describe variants which have either the addition or removal of bases in the multiples of three, which results in the insertion or deletion of amino acids within the reading frame. Frameshifts describe variants which have an addition or removal of bases, not as a multiple of three, resulting in a change of the reading frame. Frameshifts are considered more damaging than INDELS, as a shift in the reading frame can result in an altered protein-coding sequence downstream. The consequences; splice acceptor variant or splice donor variants, were found in fewer frequency in the cell lines. These variants are found at the boundary of an exon and intron site, termed the splice site. These can disrupt RNA splicing, resulting in the loss of exons or the inclusion of introns, thereby altering the protein-coding sequence, and consequently, are considered to be very damaging. The two consequences seen the least, are arguably two of the most severe. Very few stop-gain variants were identified in the cell lines, in a range of 3 – 16 stop-gain variants. Stop-gain variants result from a change in a base which introduces a premature stop codon, consequently terminating the protein-coding sequence early, leading to a truncated protein. It must be noted that frameshift variants often result in subsequent stop-gain variants. Finally, stop-lost variants were seen the least in the cell line panel, with none being identified in the MDA-MB-468 and MDA-MB-468 derived chemo-resistant cell lines and HCC38^{r}CDDP^{3000}, and the rest of the cell lines were found to have 1-2 variants of this consequence. Stop-lost variants are a result of a base change within a stop codon, which leads to elongation of the protein-coding sequence.

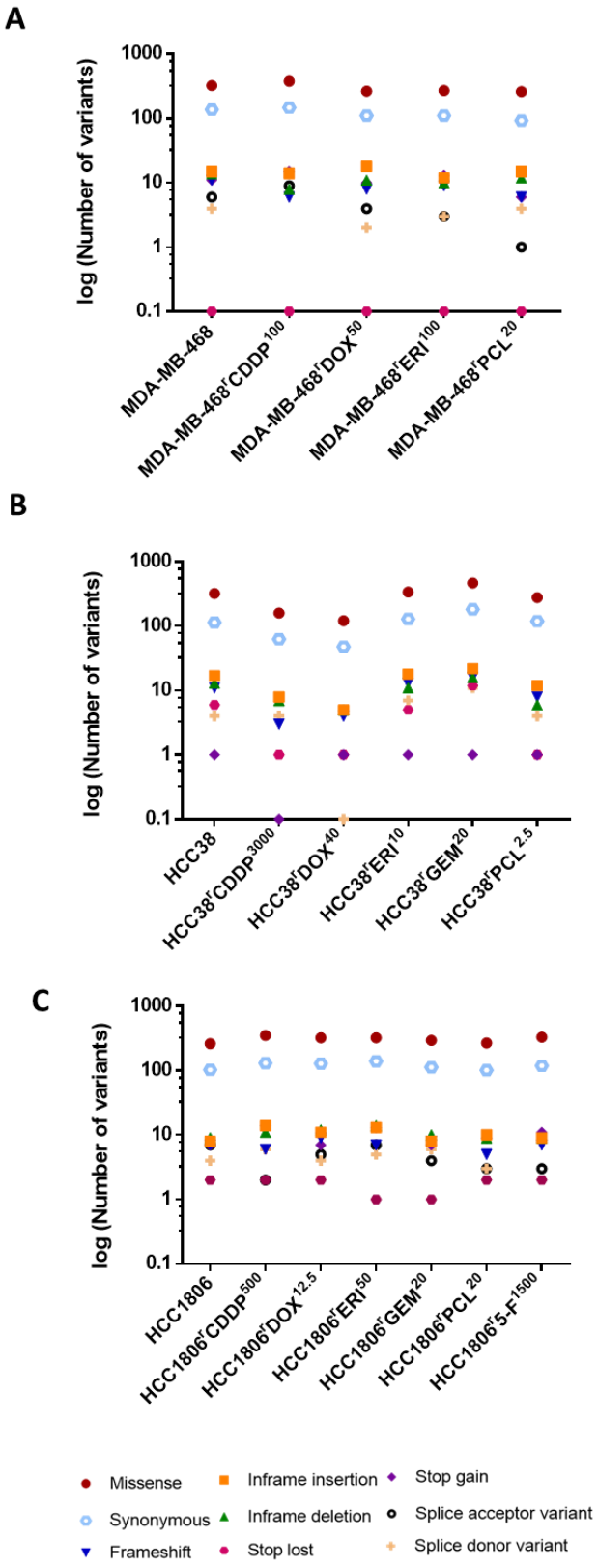


Figure 5.7 Consequence of high confidence somatic variants called in chemo-naive and chemo-resistant TNBC cell lines
Number of variants considered to be the consequences; missense, synonymous, frameshift, inframe insertion, inframe deletion, stop gain, splice acceptor or splice donor in the A) MDA-MB-468, B) HCC38, C) HCC1806 chemo-naive and chemo-resistant TNBC cell lines. Numerical data in Appendix A24.

Next the variants called in the chemo-resistant cell lines were compared to the variants called in the respective chemo-naive cell lines (Figure 5.8, numerical data in Appendix 25). To do this, the first VCF files, which included all called variants prior to filtering, were considered alongside the final VCF files containing high confidence somatic variants.

De novo variants are variants which are called in high confidence in the chemo-resistant cell line, but are never called, even in low confidence, in the respective chemo-naive cell line. MDA-MB-468^{r}CDDP^{1000}, MDA-MB-468^{r}PCL^{20} and HCC38^{r}GEM^{20} all showed to have the highest number of *de novo* variants called with 213, 225 and 203 respectively (Figure 5.A). Even giving the low number of total variants called in HCC38^{r}CDDP^{3000} and HCC38^{r}DOX^{40}, 98 and 31 variants were found to be *de novo* (Figure 5.8A).

Variants which were called in high confidence in the chemo-resistant cell line, and evidence of the variant exists at low confidence in the respective chemo-naive cell line are considered *gained* variants. MDA-MB-468^{r}CDDP^{1000}, MDA-MB-468^{r}PCL^{20} and HCC38^{r}GEM^{20}, have the highest number of *gained* variants called with 286, 276 and 381 (Figure 5.8B). Only 223 *gained* variants were called in HCC38^{r}DOX^{40} (Figure 5.8B).

Next, variants which were called in high confidence in the chemo-naive cell lines, but are never called, even in low confidence, in the respective chemo-resistant cell line were termed as *not called* variants. MDA-MB-468^{r}PCL^{20} showed the highest number, 345, of *not called* variants (Figure 5.8C). The second highest number of *not called* variants were seen in both HCC38^{r}CDDP^{3000} and HCC38^{r}DOX^{40} with 162 *not called* variants each (Figure 5.8C).

Lost variants are called in the chemo-naive cell line, but evidence is found of the variant in low confidence in the respective chemo-resistant cell line. A high number of variants called as *lost* were seen in MDA-MB-468^{r}PCL^{20}, HCC38^{r}CDDP^{3000} and HCC38^{r}DOX^{40} with 398, 362 and 355 variants respectively (Figure 5.8D). It was also noted that out of the 129 *lost* variants called in HCC38^{r}GEM^{20}, only four were synonymous (Figure 5.8D).

Finally, the remaining variants to be considered are the number of variants that are *shared* between the chemo-naive and chemo-resistant cell lines. *Shared* variants are variants that are called in high confidence in both the chemo-naive and chemo-resistant cell lines. HCC338 GEM had the highest number of *shared* variants with HCC38 (368 variants; Figure 5.8E). MDA-MB-468^{r}PCL^{20}, HCC38^{r}CDDP^{3000} and HCC38^{r}DOX^{40} each had the least shared variants 128, 135 and 142 respectively (Figure 5.8E).

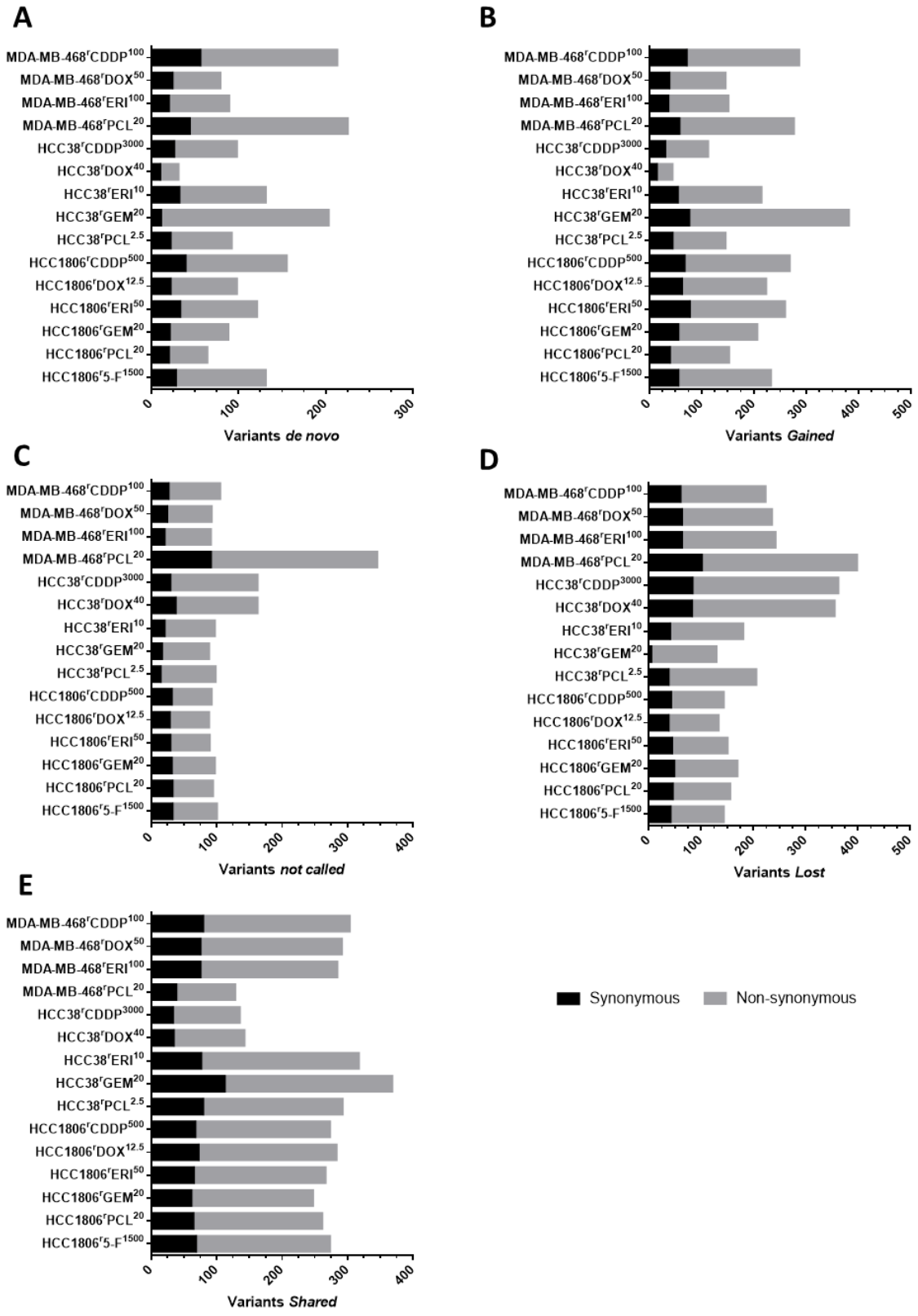


Figure 5.8 Comparison of chemo-resistant cell lines to respective chemo-naive cell lines
Number of synonymous and non-synonymous variants found to be A) *de novo*, B) *gained*, C) *not called* and D) lost in the chemo-resistant cell lines compared to respective chemo-naive cell line. E) Number of synonymous and non-synonymous variants found to be *shared* between chemo-naive and chemo-resistant TNBC cell lines. Numerical data in Appendix A25.

Analysis of the proportion of variant reads, out of the total number of reads at a given locus, can allow for the calculation of the variant allele frequency (VAF). The VAF gives an indication of the prevalence of a called variant within the cell population. For example, a VAF equal to 0 or 1 would be indicative of homozygous to reference or variant respectively, whereas a VAF equal of 0.5 would be indicative of a heterozygous variant in the cell population. Given the heterogeneous population of these cell lines, the latter could indicate either a) each cell in the population are heterozygous for the variant, or b) half the cell population are homozygous for the variant (VAF equal to 1), whilst the other half of the cell population do not have the variant at all (VAF equal to 0).

In order to see a shift in the heterogeneity of the cell population, the VAF of variants shown to be shared between a chemo-resistant cell line with the respective chemo-naive cell line were used to calculate a fold shift. A fold shift is the ratio of chemo-resistant cell line VAF to the respective chemo-naive cell line VAF. The log of the VAF fold shift, and the frequency could be plotted in a density plot to visualise VAF shifts (Figure 5.9-5.12). Plots were made using the online tool DensityPlotter (Spencer, Yakymchuk and Ghaznavi, 2017). A log(VAF fold shift) equal to one is indicative of an two-fold increase of VAF in the chemo-resistant cell line compared to the respective chemo-naive cell line, whilst a log(VAF fold shift) equal to negative one is indicative of a two-fold decrease of VAF in the chemo-resistant cell line compared to the respective chemo-naive cell line.

MDA-MB-468^{r}DOX^{50} and MDA-MB-468^{r}PCL^{20} both had a normal distribution of variants that demonstrated an increase or decrease in VAF (Figure 5.9). Given that not many variants were found to be *de novo* or *gained* (compared to other cell lines in the panel) in MDA-MB-468^{r}DOX^{50}, it's interesting to note that significant shifts in VAF were not seen. MDA-MB-468^{r}ERI^{50} distribution was skewed towards a higher number of variants, which demonstrated an increase in VAF, whilst MDA-MB-468^{r}CDDP^{1000} distribution was skewed towards variants that had a decrease in VAF. This suggests that of the variants shared with MDA-MB-468, the frequency was decreased in the heterogenous population in MDA-MB-468^{r}CDDP^{1000}. This is interesting to note, as earlier in this chapter, MDA-MB-468^{r}CDDP^{1000} was one of the cell lines to demonstrate the highest number of *de novo* and *gained* variants. HCC38^{r}CDDP^{3000} had an almost normal distribution with the number of variants with an increase of decrease in VAF (Figure 5.10). HCC38^{r}DOX^{40}, which had few variants *shared* with HCC38, had variants with an increase in VAF. The distribution of both HCC38^{r}GEM^{20} and HCC38^{r}PCL$^{2.5}$ are both skewed towards a higher number of variants which demonstrated an increase in VAF. The distribution for HCC38^{r}ERI^{10} shows a slight increase in variants which have a decrease in VAF (Figure 5.10). There was no notably significant shift in the VAF of variants between HCC1806 and the respective chemo-resistant cell lines (Figure 5.11).

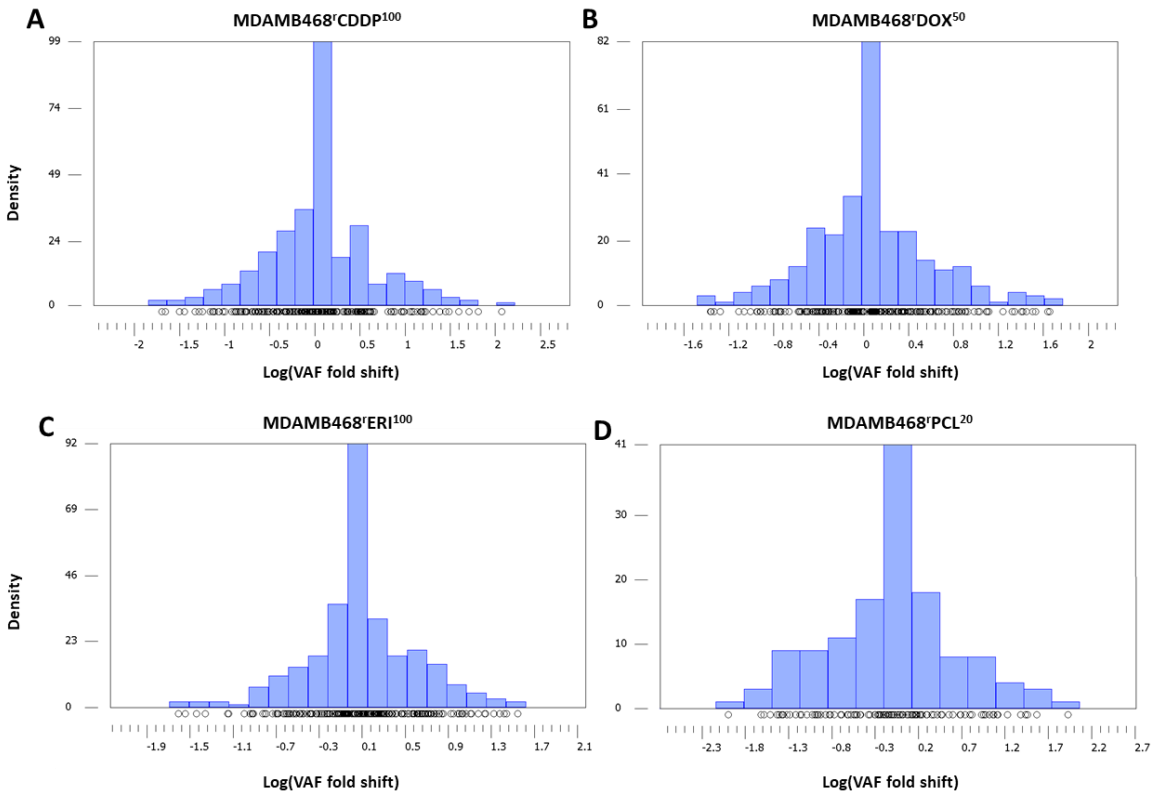


Figure 5.9 Shift in variant allele frequency of variants shared between MDA-MB-468 and MDA-MB-468 derived chemo-resistant cell lines

The variant allele frequency (VAF) was calculated for each variant as the frequency of alternative reads against total reads for each variant. Shift in VAF in variants shared between MDA-MB-468 and MDA-MB-468 derived chemo-resistant cell lines were then calculated by determining fold change in VAF. Density plots created of VAF fold shift for A) MDA-MB-468^{r}CDDP^{1000} B) MDA-MB-468^{r}DOX^{50} C) MDA-MB-468^{r}ERI^{50} D) MDA-MB-468^{r}PCL^{20}

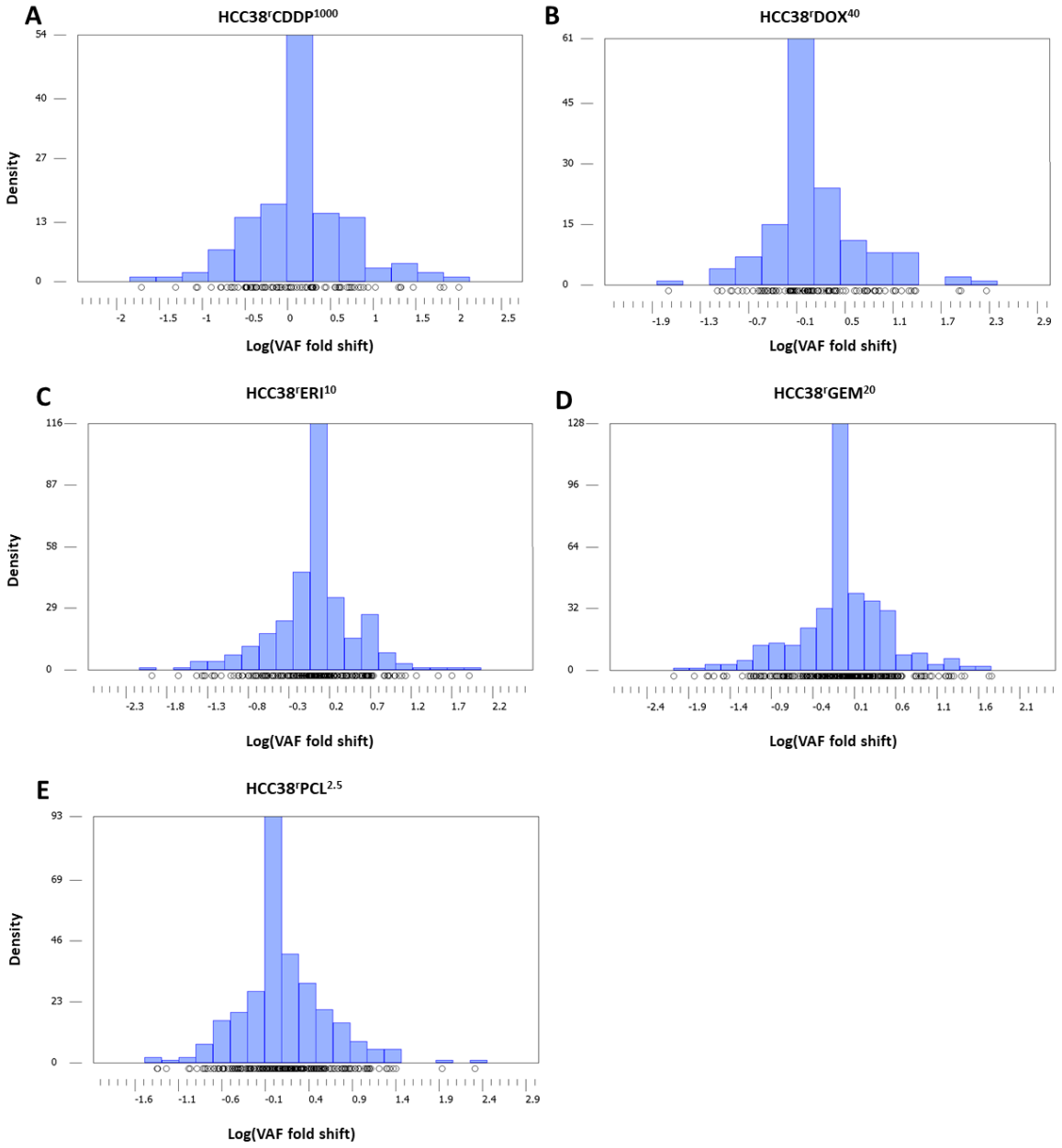


Figure 5.10 Shift in variant allele frequency of variants shared between HCC38 and HCC38 derived chemo-resistant cell lines
The variant allele frequency (VAF) was calculated for each variant as the frequency of alternative reads against total reads for each variant. Shift in VAF in variants shared between HCC38 and HCC38 derived chemo-resistant cell lines were then calculated by determining fold change in VAF. Density plots created of VAF fold shift for A) HCC38^{r}CDDP^{3000} B) HCC38^{r}DOX^{40}C) HCC38^{r}ERI^{10} D) HCC38^{r}GEM^{20} E) HCC38^{r}PCL$^{2.5}$

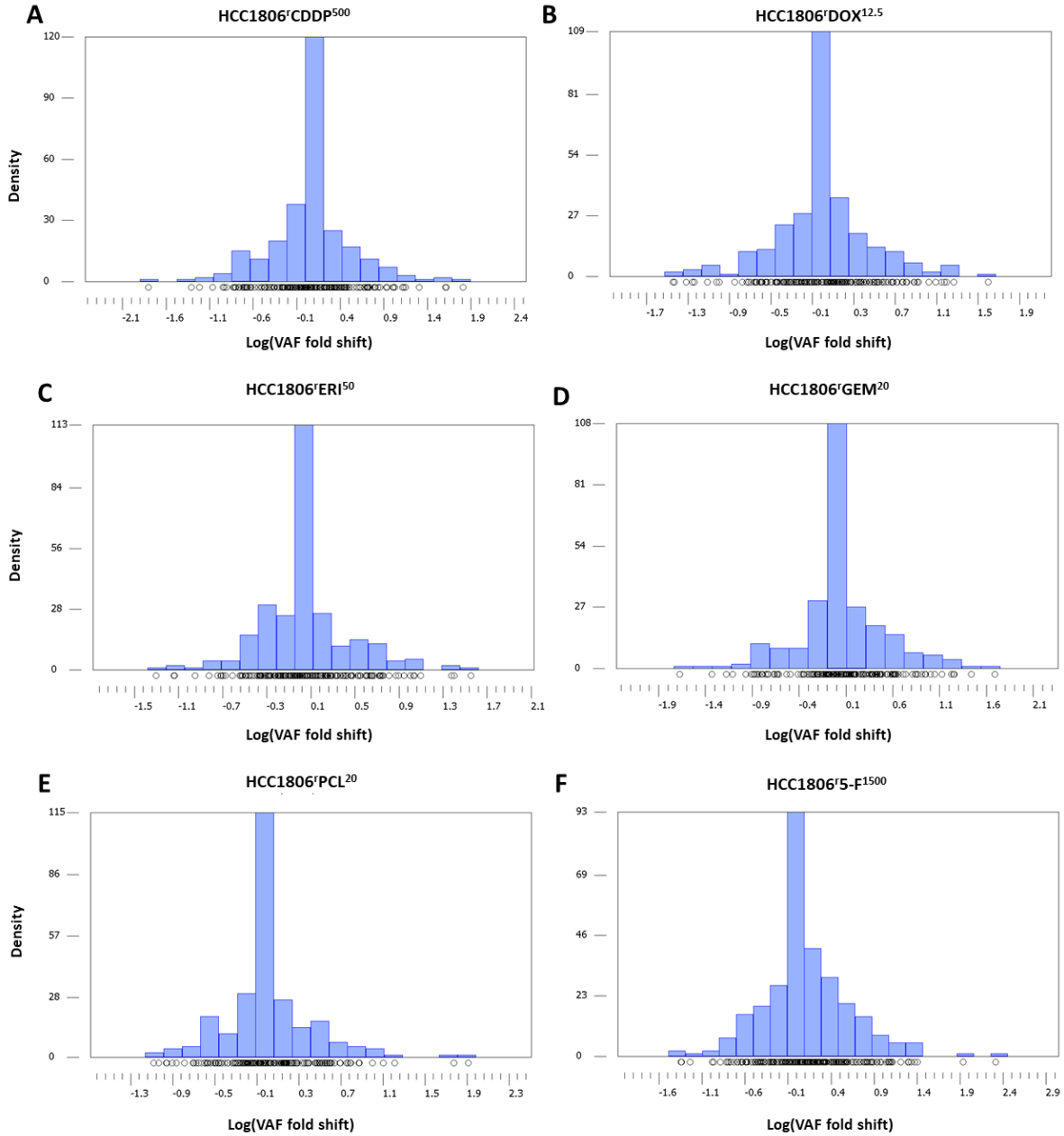


Figure 5.11 Shift in variant allele frequency of variants shared between HCC1806 and HCC1806 derived chemo-resistant cell lines
The variant allele frequency (VAF) was calculated for each variant as the frequency of alternative reads against total reads for each variant. Shift in VAF in variants shared between HCC1806 and HCC1806 derived chemo-resistant cell lines were then calculated by determining fold change in VAF. Density plots created of VAF fold shift for A) HCC1806ʳCDDP^{500} B) HCC1806ʳDOX$^{12.5}$ C) HCC1806ʳERI^{50} D) HCC1806ʳGEM^{20} E) HCC1806ʳPCL^{20} F) HCC1806ʳ5-F^{1500}

5.2.1.4 Mutational patterns and signatures

Somatic variants in cancer genomes can leave mutational profiles. These profiles are mutation patterns which can then be associated to a mutational signature. There are six substitution subtypes; C > A, C > G, C > T, T > A, T > C and T > G. Each of these can also be examined on sequence context by incorporating the bases immediately at 5 prime (5‘) and 3 prime (3’) of each mutated base generating 96 different possible mutation sub-types, whereby; six types of substitution * four types of 5’ base * four types of 3’ base. In order to analyse the base changes, high confidence somatic variants were submitted to the online tool Mutalisk in a VCF (Lee *et al.*, 2018). Mutalisk allowed for the analysis of the identified variants in context of the GRCh37 reference build to identify mutational patterns as a relative proportion of observed trinucleotide frequency.

Across each of the MDA-MB-468 chemo-naive and chemo-resistant cell lines, C > T base changes are the most common, with the least common observed as either T > A or T > G base changes (Figure 5.12). MDA-MB-468^{r}CDDP^{100} has an increase in variants with the base change *5’* C, C > T, C *3’*, *5’* C, C > T, T *3’*, and *5’* G, C > A, C *3’ c*ompared to MDA-MB-468. MDA-MB-468^{r}DOX^{50} had an increase in variants with the base change *5’* G, C>A, A *3’* compared to MDA-MB-468. MDA-MB-468^{r}PCL^{20} had an increase in variants which have *5’* A, C > T, G *3’* and *5’* C, C > T, G *3’* base changes compared to MDA-MB-468. No significant changes were seen in the MDA-MB-468^{r}ERI^{50} cell line compared to MDA-MB-468.

Across each of the HCC38 chemo-naive and chemo-resistant cell lines C > T base changes are the most common, whilst T > A are the least common (Figure 5.13). A small increase in variants with the base change *5’* G, C > T, T *3’* was seen in both HCC38^{r}ERI^{10} and HCC38^{r}PCL$^{2.5}$ cell lines compared to HCC38. MDA-MB-468^{r}ERI^{50} had less variants with the base change *5’* C, T > C, C *3’* compared to HCC38. Given the low number of variants called in HCC38^{r}CDDP^{3000} and HCC38^{r}DOX^{40}, it was not surprising to see that in each case there were often no variants called for an observed context. Even though a disparity in the number of variants called in HCC38^{r}CDDP^{3000}, HCC38^{r}DOX^{40} and HCC38^{r}GEM^{20} the high percentage of C > T and lowest percentage of T > A, T > G base changes remained the same. Both HCC38^{r}CDDP^{3000} and HCC38^{r}DOX^{40} demonstrated a higher percentage of variants called which were related to bases changes C > A, (19.5% and 15.1% respectively) compared to the other HCC38 and HCC38 derived chemo-resistant cell lines.

Across the HCC1806 chemo-naive and chemo-resistant cell lines, C > T and T > C base changes are the most common, with T > A the least common (Figure 5.14). HCC1806^{r}CDDP^{500} had an increase in variants with the base changes; *5’* A, C > A, C *3’*, *5’* G, C > A, C *3’*, and *5’* G, T > C, C *3’* compared to

HCC1806. HCC1806^{r}DOX$^{12.5}$ had an increase in variants with the base change *5'* C, T > C, T *3'* compared to HCC1806. HCC1806^{r}ERI^{50} had an increase in variants with the base changes; *5'* T, C > A, A *3'* and *5'* G, T > C, G *3'* compared to HCC1806. HCC1806^{r}PCL^{20} also demonstrated an increase in variants with a base change *5'* G, T > C, G *3'*, *5'* G, C > T, T *3'* compared to HCC1806. HCC1806^{r}GEM^{20} and HCC1806^{r}5-F^{1500} both had an increase in variants with the base changes; *5'* T, C > G, A *3'*, *5'* A, T > C, G *3'* and *5'* G, T > C, G *3'* compared to HCC1806.

Taken together this data show that C > T base changes are the most common across each of the chemo-naive and chemo-resistant cell lines. When considering mutation patterns between cell lines which were developed to have resistance to the same chemotherapeutic agent, no similarity was seen.

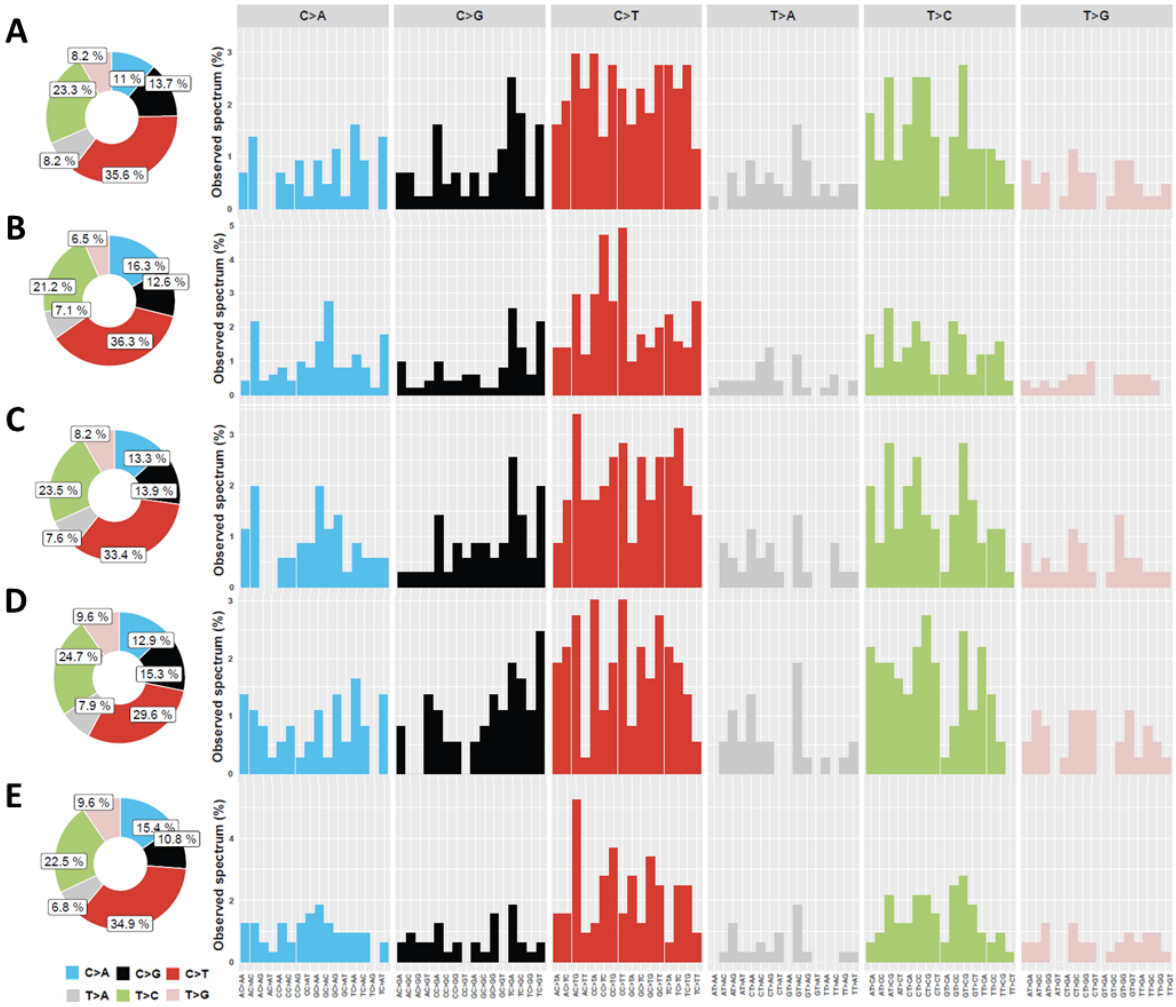


Figure 5.12 Mutational patterns in MDA-MB-468 and MDA-MB-468 derived chemo-resistant cell lines
Frequency of variant base changes and the trinucleotide context based on alignment to human genome reference build GRCh37 for A) MDA-MB-468 B) MDA-MB-468^{r}CDDP^{1000} C) MDA-MB-468^{r}DOX^{50} D) MDA-MB-468^{r}ERI^{50} E) MDA-MB-468^{r}PCL^{20}. Mutational patterns calculated using Mutalisk tool.

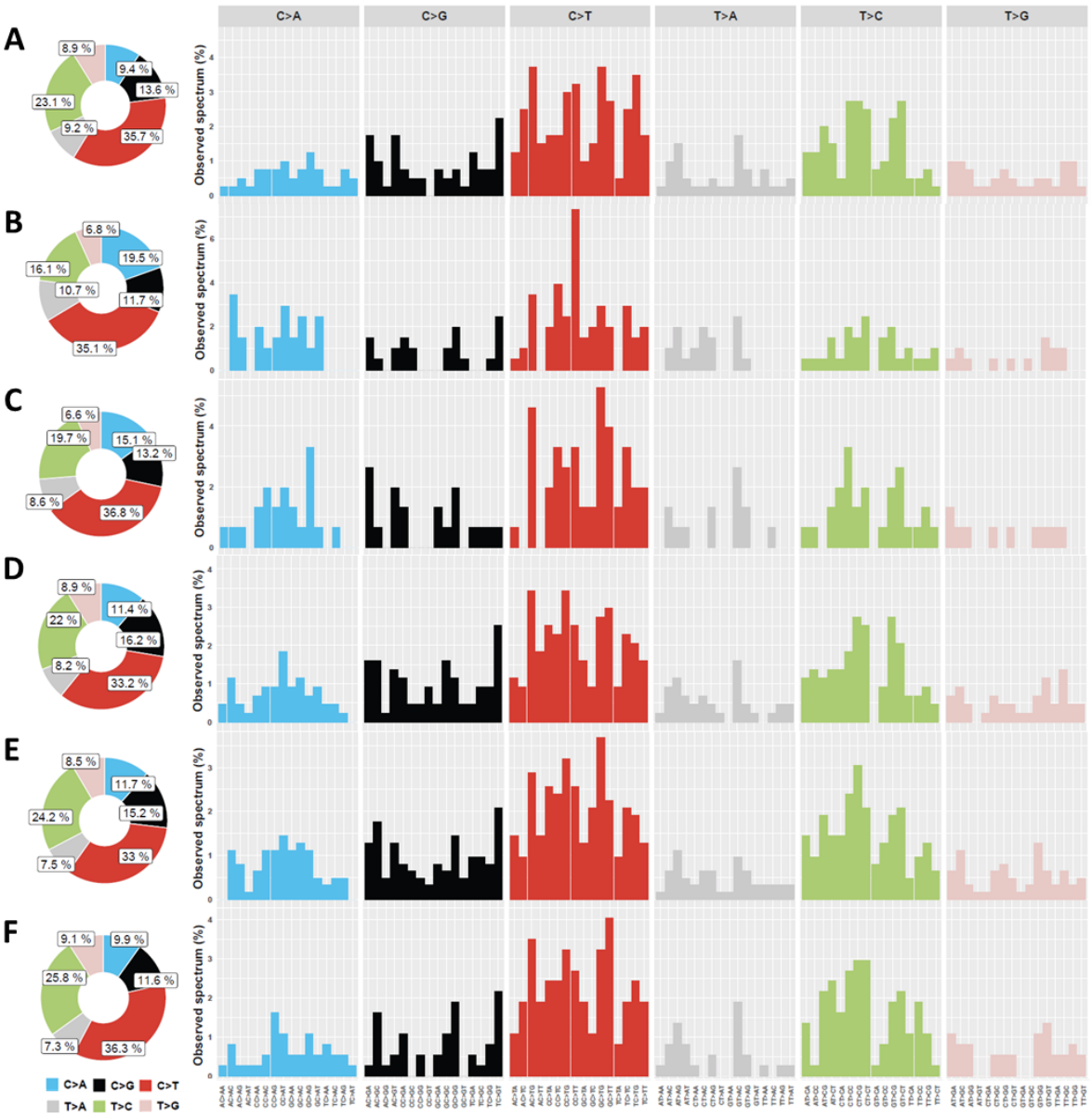


Figure 5.13 Mutational patterns in HCC38 and HCC38 derived chemo-resistant cell lines
Frequency of variant base changes and the trinucleotide context based on alignment to human genome reference build GRCh37 for A) HCC38 B) HCC38^{r}CDDP^{3000} C) HCC38^{r}DOX^{40} D) HCC38^{r}ERI^{10} E) HCC38^{r}GEM^{20} F) HCC38^{r}PCL$^{2.5}$. Mutational patterns calculated using Mutalisk tool.

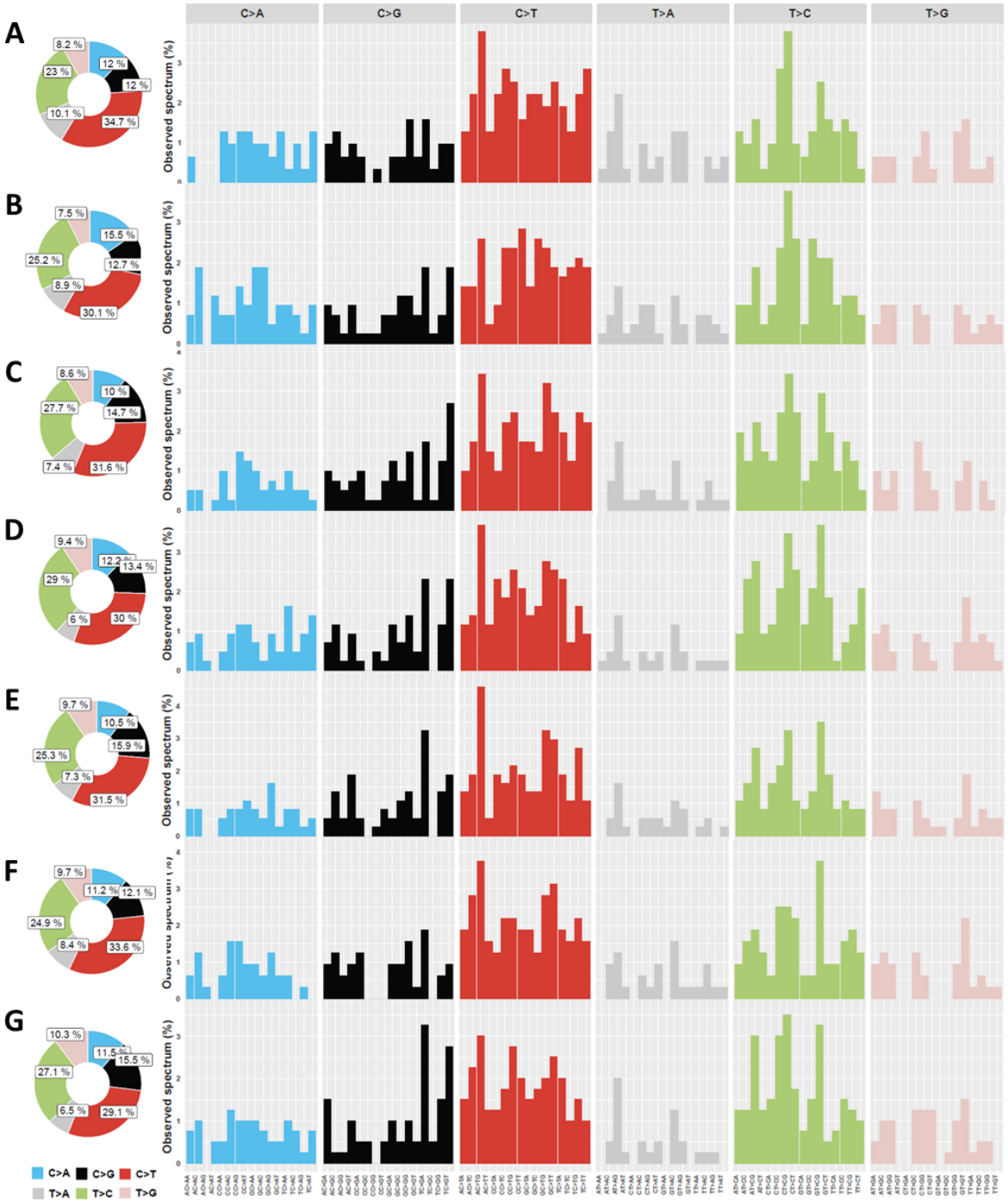


Figure 5.14 Mutational patterns in HCC1806 and HCC1806 derived chemo-resistant cell lines
Frequency of variant base changes and the trinucleotide context based on alignment to human genome reference build GRCh37 for A) HCC1806 B) HCC1806'CDDP^{500} C) HCC1806'DOX$^{12.5}$ D) HCC1806'ERI^{50} E) HCC1806'GEM^{20} F) HCC1806'PCL^{20} G) HCC1806'5-F^{1500}. Mutational patterns calculated using Mutalisk tool.

Having established the mutational patterns in each of the cell lines, a kataegis analysis can be performed. Kataegis describes the phenomenon in which a large number of highly-patterned base pair variants occur in a small region of DNA. This analysis was first used by the Wellcome Trust Sanger institute in order to describe the observations of breast cancer cells by mapping variant clusters across the genome through a rainfall plot to identify clustering patterns of kataegis (Nik-Zainal *et al.*, 2012). It is expected that variants should be evenly dispersed throughout a genome, and so most variants should have approximately one mega base distance between them. In order to perform this analysis, the high confidence somatic variants called in each cell line, were submitted to the online tool Mutalisk in a VCF. This tool produces a rainfall plot in which intermutation distance, i.e., the distance between each somatic variant, and the variant immediately before it, is plotted for each variant. This allows the identification of "mutation showers", which are the base changes that show to have shorter distances (Figure 5.15).

Kataegis is observed in the MDA-MB-468 chemo-naive and chemo-resistant cell lines in chromosomes 3, and 11. MDA-MB-468 also demonstrated kataegis at chromosome 7, 13 and 14. MDA-MB-468^{r}CDDP^{1000} demonstrated kataegis at chromosome 7 and 14, like MDA-MB-468, but also at chromosome 9, 12 in addition. MDA-MB-468^{r}DOX^{50} demonstrated kataegis at chromosome 7, 13 and 14, like MDA-MB-468, but also at chromosome 9. MDA-MB-468^{r}ERI^{50} demonstrated kataegis at chromosome 13, like MDA-MB-468. MDA-MB-468^{r}PCL^{20} did not have any further observed kataegis than already stated. Kataegis is observed in the HCC38 chemo-naive and chemo-resistant cell lines in chromosome 3, 11 and 12. HCC38 and HCC38^{r}ERI^{10}, HCC38^{r}GEM^{20} and HCC38^{r}PCL$^{2.5}$ also all demonstrated kataegis at chromosome 7. In addition, kataegis is observed at in chromosome 8 in HCC38^{r}ERI^{10}, chromosome 17 in HCC38^{r}GEM^{20} and in chromosome 18 in HCC38^{r}PCL$^{2.5}$. In the HCC1806 chemo-naive and chemo-resistant cell lines, kataegis is observed in chromosome 3 and 11. In addition, HCC1806^{r}CDDP^{500} demonstrated kataegis in chromosome 8, and both HCC1806^{r}CDDP^{500} and HCC1806^{r}5-F^{1500} demonstrated kataegis in chromosome 7. Kataegis was demonstrated in chromosome 12 in HCC1806^{r}DOX$^{12.5}$ and HCC1806^{r}ERI^{50}, and HCC1806^{r}ERI^{50} also demonstrated kataegis in chromosome 13. Observed kataegis was also seen in chromosome 13 in HCC1806^{r}GEM^{20} and HCC1806^{r}PCL^{20}.

Taking this data together, it can be seen kataegis was observed at both chromosome 3 and 11 in each of the chemo-naive and chemo-resistant TNBC cell lines, suggesting that this may be associated to TNBC, rather than driving drug-resistance. When considering the cell lines which have been developed to have resistant to the same chemotherapeutic agent, no pattern of kataegis was observed.

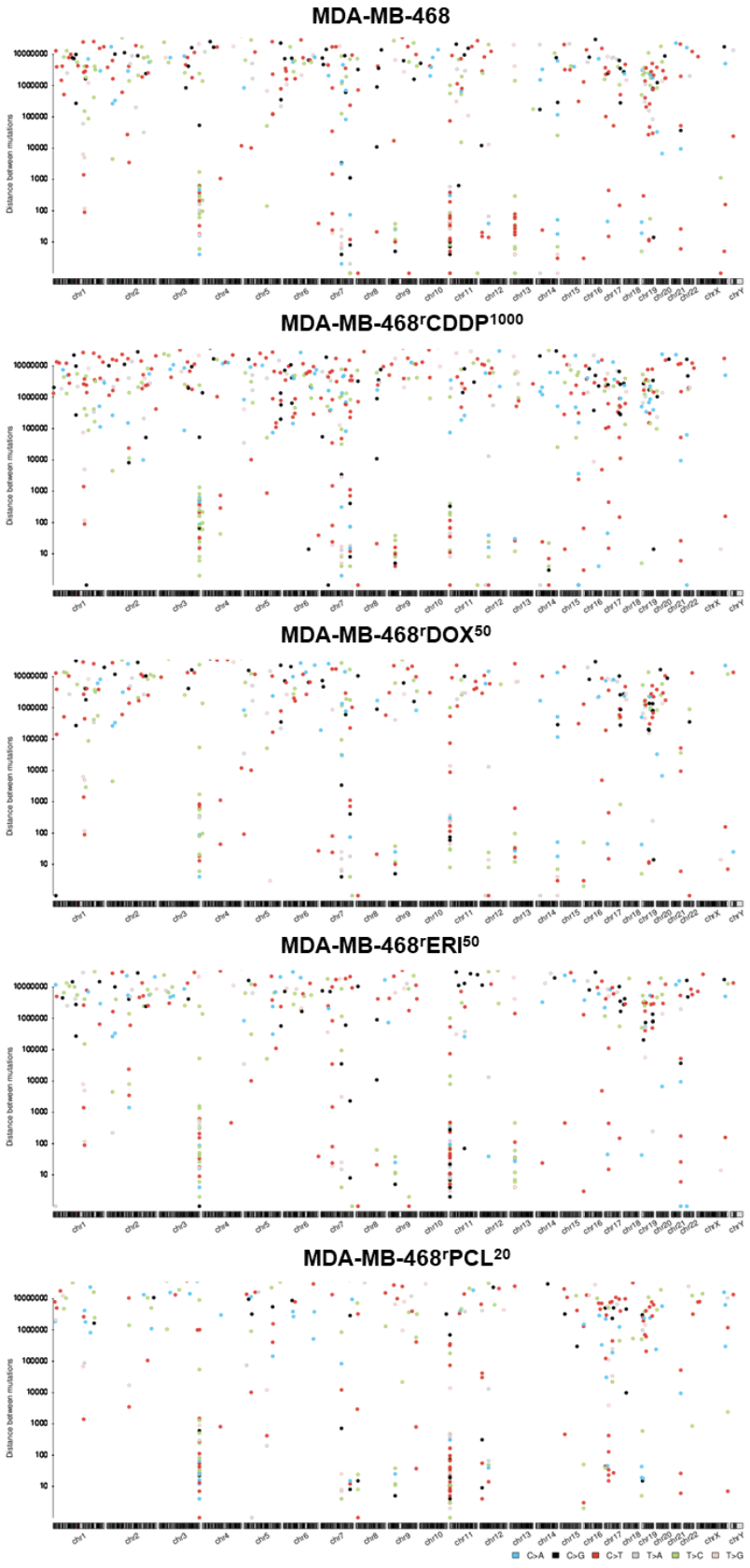


Figure 5.15 Kataegis of each chemo-naive and chemo-resistant TNBC cell line.
Continued on next page

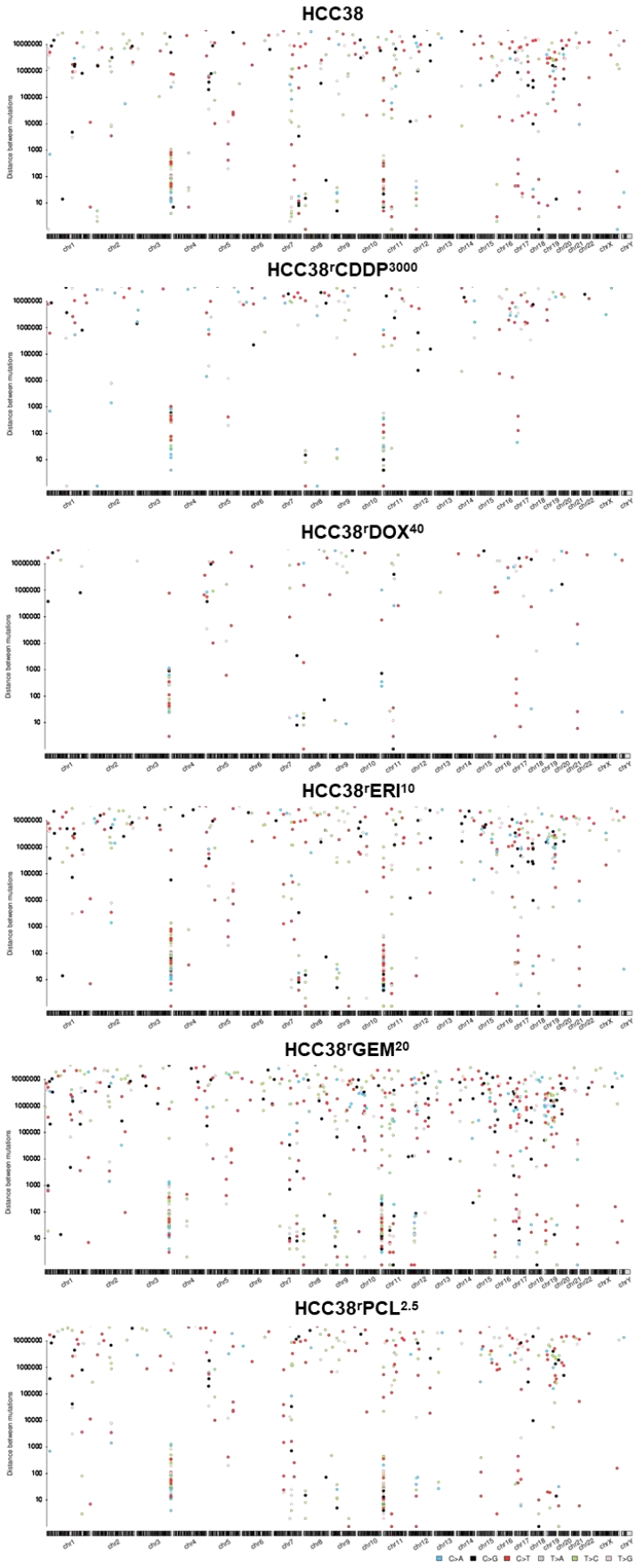


Figure 5.15 Kataegis of each chemo-naive and chemo-resistant TNBC cell line.
Continued on next page

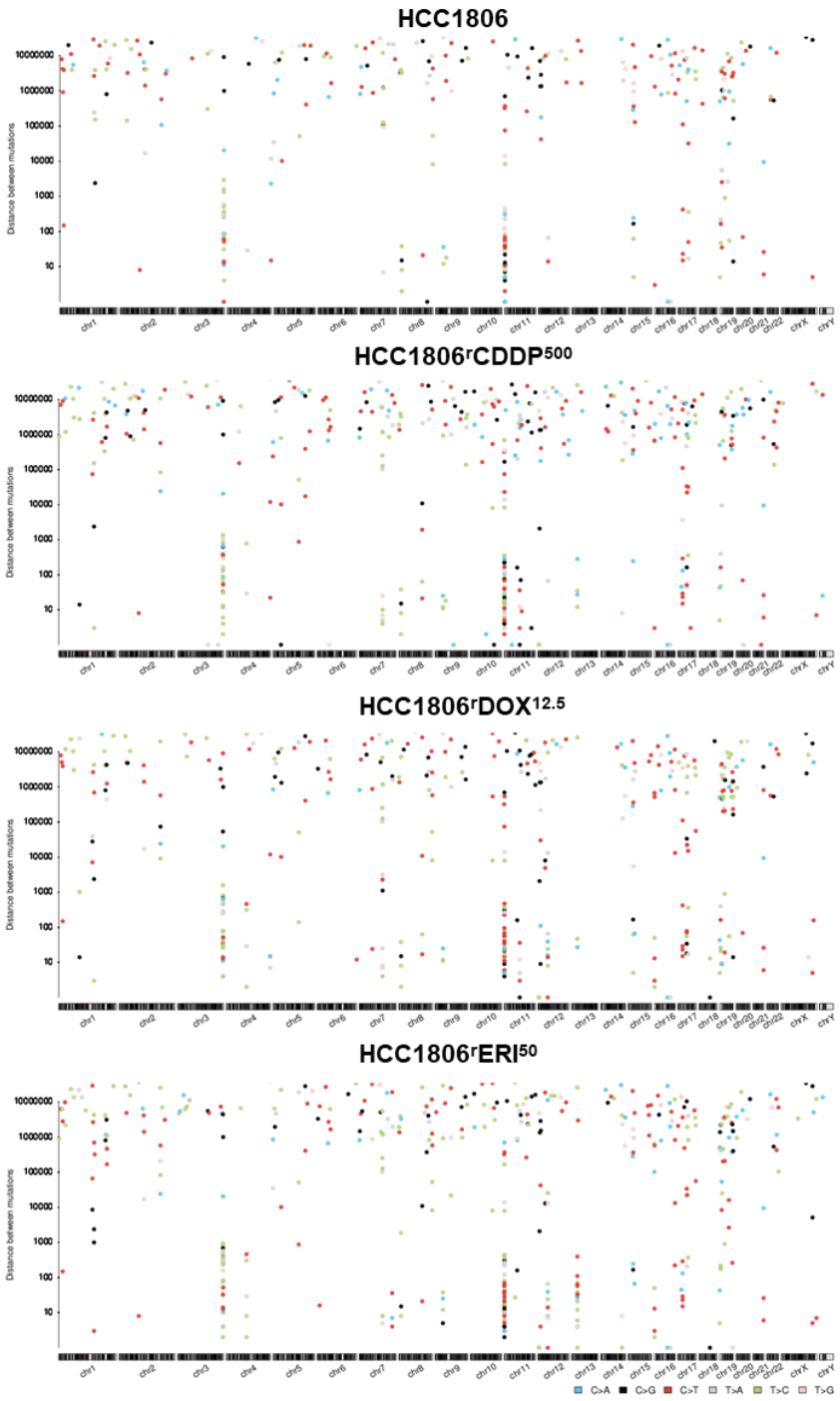


Figure 5.15 Kataegis of each chemo-naive and chemo-resistant TNBC cell line.
Continued on next page

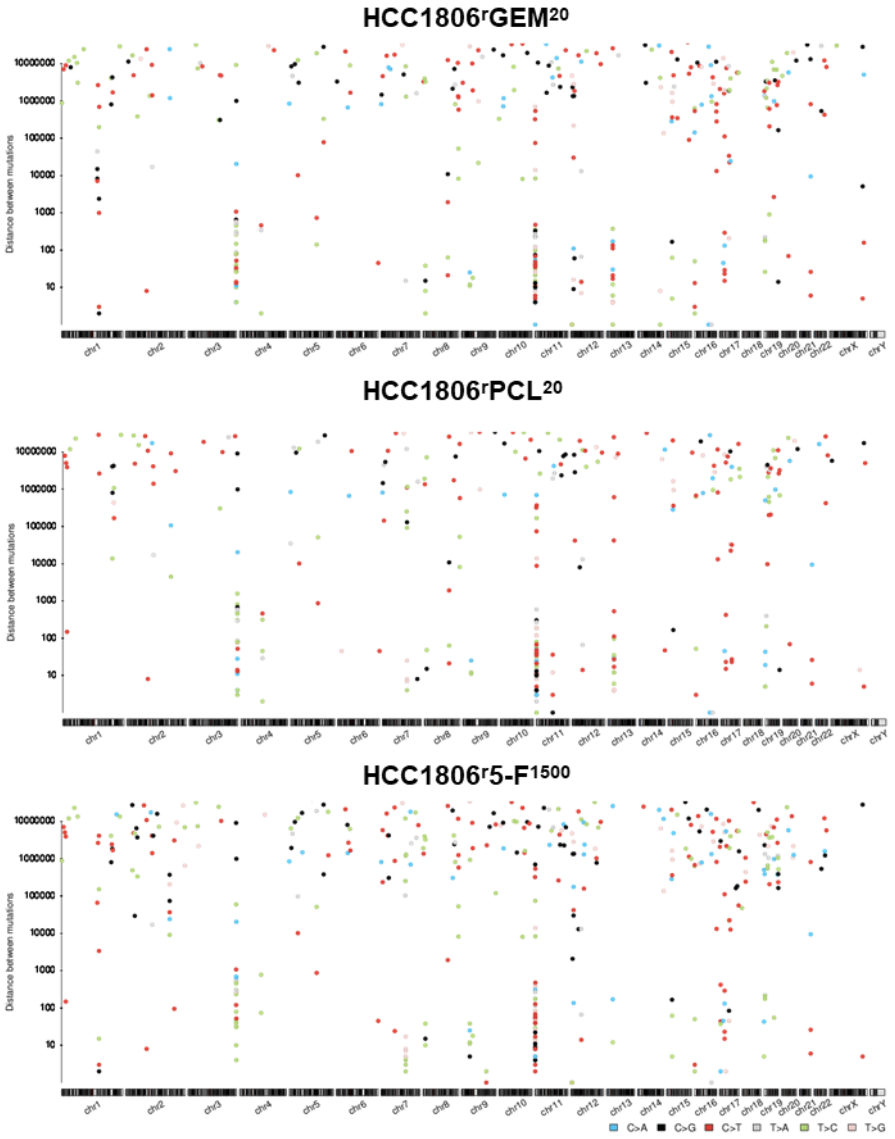


Figure 5.15 Kataegis of each chemo-naive and chemo-resistant TNBC cell line.
Kataegis patterns for each of the chemo-naive and chemo-resistant TNBC cell lines calculated through intermutation distance between each somatic variant and visualised as a rainfall plot. Kataegis calculated using Mutalisk tool with high confidence somatic variant VCF as input.

The Mutalisk tool also allowed for the investigation of known mutational signatures, as described in the work of Alexandrov *et al.*, 2020. Variants may be a consequence of multiple mutational processes including; as a result of infidelity of DNA replication machinery, exogenous or endogenous mutagen exposures (such as chemotherapeutic agents), and the modification of enzymes associated with DNA repair. The mutational signatures were extracted from the 2,780 whole genome variant calls produced by the International Cancer Genome Consortium (ICGC)/TCGA Pan-Cancer Analysis of Whole Genomes (PCAWG) network (International Cancer Genome Consortium, 2010; Weinstein, 2013; Campbell *et al.*, 2020). The signatures are divided into three categories; Single Base Substitution (SBS), Doublet base substitution (DBS) and Small Insertion and Deletion (ID) signatures. Each signature has a proposed aetiology, although some signatures are currently unknown or speculated upon, and these require further evidence, through investigation, to establish the aetiology. Mutalisk identifies seven decomposition models based on SBS, with the best fit based on the ranking by the Bayesian Information Criterion (BIC), which identifies which mutational signatures the cell line being analysed relates to the most. Here the lowest BIC value indicates the best model, and this model was used to identify the mutational signatures each of the TNBC cell lines showed a similarity. The percentage of similarity to the mutational signatures was determined in each TNBC cell line (Figure 5.16).

Many of the identified signatures were associated with defects in the ability of the cells to repair DNA. All of the cell lines, with the exception of HCC38^{r}GEM^{20}, HCC1806^{r}CDDP^{500} and HCC1806^{r}5-F^{1500} demonstrated a similarity to SBS1. The signature SBS1 is thought to be a result of an endogenous mutational process which leads to the deamination of the 5-methylcytosine to generate thymine, consequently generating G:T mismatches in dsDNA. It is thought that the inability to remove the mismatch before DNA replication results in fixation of the T substitution for C. The signature SBS1 is often found correlated with the signature SBS5. Only MDA-MB-468, MDA-MB-468^{r}CDDP^{1000}, MDA-MB-468^{r}DOX^{50}, MDA-MB-468^{r}ERI^{50}, HCC38, HCC38^{r}PCL$^{2.5}$ and HCC1806^{r}DOX$^{12.5}$ demonstrated a similarity to SBS5. The aetiology of SBS5 is unknown, although it has been found to be increased in bladder cancer samples which harbour mutations in the DNA excision repair protein ERCC6, and also in cancers as a result of tobacco smoking. A similarity to the signature SBS3 is found in all of the chemo-naive and chemo-resistant cell lines with the exception of MDA-MB-468^{r}CDDP^{1000}, HCC38, HCC1806^{r}DOX$^{12.5}$ and HCC1806^{r}ERI^{50}. SBS3 is associated with defective HRR. This is identified through small INDELs and genome rearrangements due to abnormal dsDNA repair or base substitution. Similarity to the signature SBS6, which is associated with defective DNA MMR, was seen in HCC38^{r}GEM^{20}, HCC1806^{r}CDDP^{500} and HCC1806^{r}5-F^{1500}. Only

HCC38^{r}DOX^{40} demonstrated similarity to the signature, SBS15, which is also associated with defective DNA MMR.

Two signatures associated to the activity of the AID/APOBEC family were identified in the cell lines. The cell lines; MDA-MB-468, MDA-MB-468^{r}CDDP^{1000}, MDA-MB-468^{r}DOX^{50}, MDA-MB-468^{r}ERI^{50}, HCC1806^{r}CDDP^{500} and HCC1806^{r}5-F^{1500} demonstrated a similarity to the signature SBS13. SBS13 is thought to be as a result of the activity of AID/APOBEC family of cytidine deaminases due to the similar sequence context of cytosine mutations caused by APOBEC found in experimental models. The mutations may be generated directly by DNA replication across uracil or by error prone polymerases, such as REV1 (a deoxycytidyl transferase in DNA repair), in replicating across abasic sites which are generated BER. It was found that the signature SBS13 is closely related to SBS2. The signature SBS2 is attributed to same activity as SBS13, although only HCC1806^{r}GEM^{20} showed a similarity to the SBS2 signature.

Three signatures were identified which have been shown to be associated with treatment of drug agents. The cell lines MDA-MB-468^{r}CDDP^{1000}, and HCC38^{r}CDDP^{3000} had similarity to the signature SBS31 which has been associated with prior chemotherapy treatment with platinum drugs. Importantly both MDA-MB-468^{r}CDDP^{1000} and HCC38^{r}CDDP^{3000} were developed to have resistance to a platinum drug, cisplatin, although it was noted that although HCC1806^{r}CDDP^{500} also had developed resistance to cisplatin, no similarity to SBS31 was seen, even when checking models with a lower BIC value. The cell lines MDA-MB-468^{r}PCL^{20}, HCC38 and HCC38^{r}ERI^{10} demonstrated a similarity to the signature SBS11 which has been related to a mutational pattern associated to alkylating agents. It was shown the SBS11 mutations were found associated to previous treatment with the alkylating agent, temozolomide, when considering patient histories. MDA-MB-468^{r}CDDP^{1000}, MDA-MB-468^{r}PCL^{20} and HCC38^{r}DOX^{40} demonstrated a similarity to the signature SBS24, which has been associated to the exposure of aflatoxin, which are a family of toxins produced by certain fungi found in agricultural crops.

Some signatures found to be associated to the TNBC cell lines are identified only through patterns found in particular cancers. The two cell lines HCC38^{r}GEM^{20} and HCC1806^{r}CDDP^{500} demonstrated a similarity to the signature SBS29 which have been found in cancer samples from individuals with a tobacco chewing habit. Three cell lines were found associated to SBS7. SBS7 is divided into four categories; SBS7a, SBS7b SBS7c and SBS7d. All four have been previously found in skin cancers as a result of sun exposure. It is predicted that they may be the consequence of pyrimidine dimers or 6-4 photoproducts, but there is no evidence of this hypothesis and it is unclear which may be

responsible for the SBS7(a-d) mutational signatures. MDA-MB-468^{r}DOX^{50} demonstrated a similarity to SBS7a, whilst HCC38^{r}DOX^{40} and HCC1806^{r}CDDP^{500} demonstrated a similarity to SBS7b. It was interesting to note that two out of the three doxorubicin resistant cell lines in the panel, showed a similarity to SBS7, and neither of them demonstrate cross-resistance to cisplatin. However, HCC1806^{r}CDDP^{500} has been shown to demonstrate cross-resistance to doxorubicin. This may suggest that the mutational signatures associated to SBS7 may be linked with doxorubicin resistance.

Many of the identified signatures in the TNBC cell lines had no known aetiology. HCC38, HCC38^{r}DOX^{40}, HCC38^{r}ERI^{10}, HCC38^{r}PCL$^{2.5}$, HCC38^{r}GEM^{20}, HCC1806^{r}DOX$^{12.5}$, HCC1806^{r}ERI^{50} and HCC1806^{r}PCL^{20} all showed similarity to the signature SBS39, although aetiology is unknown. HCC38^{r}ERI^{10} and HCC38^{r}PCL$^{2.5}$ show similarity to the signature SBS19, and MDA-MB-468^{r}CDDP^{1000} and HCC1806^{r}ERI^{50} show similarity to SBS40, although the aetiology of these signatures is also unknown. Finally, MDA-MB-468^{r}DOX^{50}, HCC1806^{r}DOX$^{12.5}$, HCC1806^{r}ERI^{50} and HCC1806^{r}GEM^{20} show similarity to the signature SBS23, although the aetiology is unknown. Interestingly each of the latter cell lines all demonstrate resistance or cross-resistance to doxorubicin.

Lastly, three of the signatures identified in the TNBC cell lines are thought to be as a consequence of sequencing artefacts. Similarity to the signature, SBS43, is found in HCC1806, and whilst the aetiology is unknown, it is predicted to be a sequencing artefact. Similarity to the signature SBS46 was found in MDA-MB-468^{r}ERI^{50}, MDA-MB-468^{r}PCL^{20}, HCC38, HCC38^{r}DOX^{40}, HCC38^{r}ERI^{10}, HCC38^{r}PCL$^{2.5}$, HCC38^{r}GEM^{20}, HCC1806^{r}CDDP^{500}, HCC1806^{r}DOX$^{12.5}$, HCC1806^{r}ERI^{50} HCC1806^{r}GEM^{20} and HCC1806^{r}5-F^{1500}. SBS46 is considered to be a sequencing artefact, but has been found commonly in colorectal cancers in early releases of the TCGA database. Finally, similarity to the signature SBS54 was seen in MDA-MB-468, MDA-MB-468^{r}CDDP^{1000}, MDA-MB-468^{r}DOX^{50}, MDA-MB-468^{r}PCL^{20}, HCC38^{r}ERI^{10}, HCC1806 and all of the HCC1806 derived chemo-resistant cell lines. SBS54 is considered as a sequencing artefact which may be as a result of contamination with germline variants.

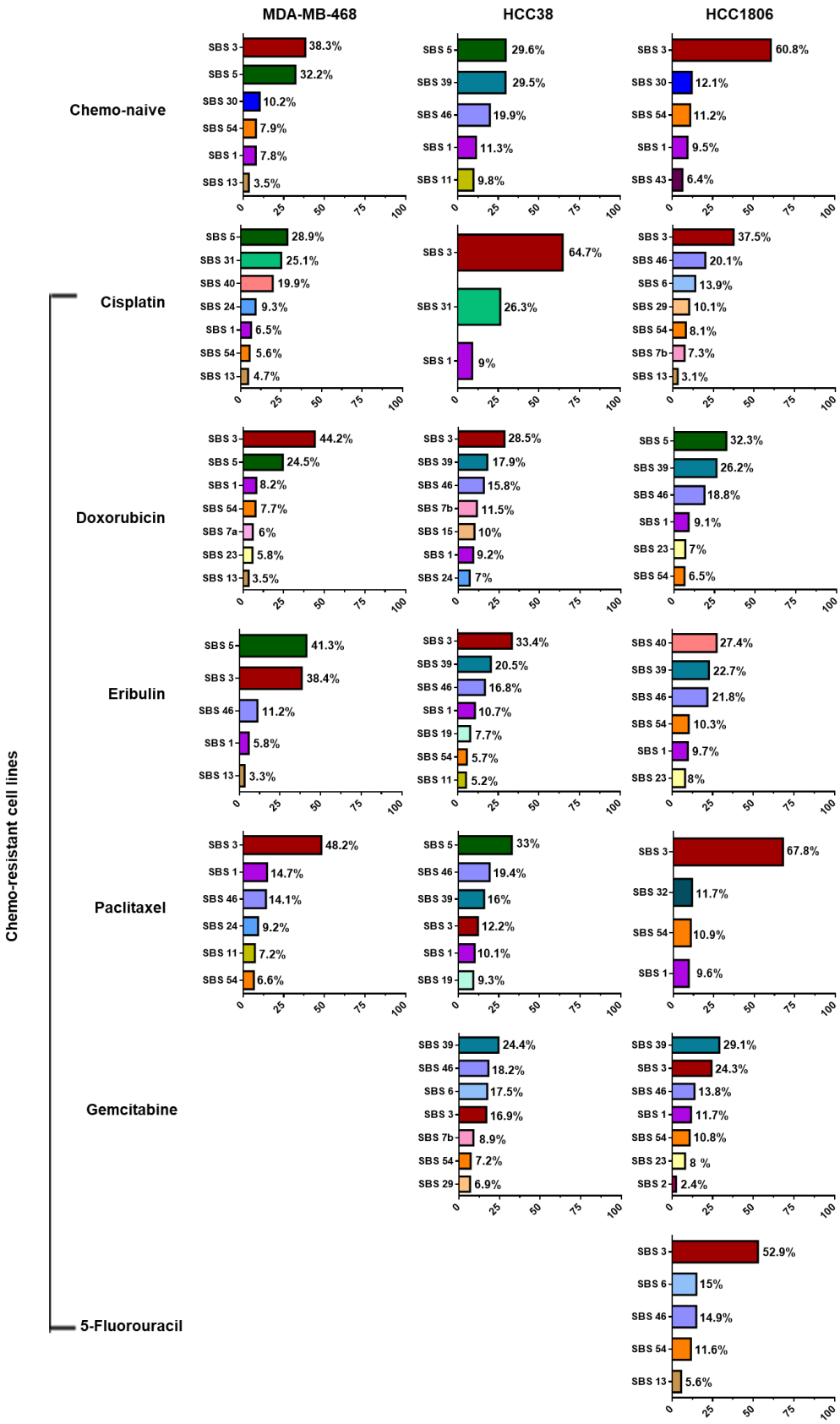
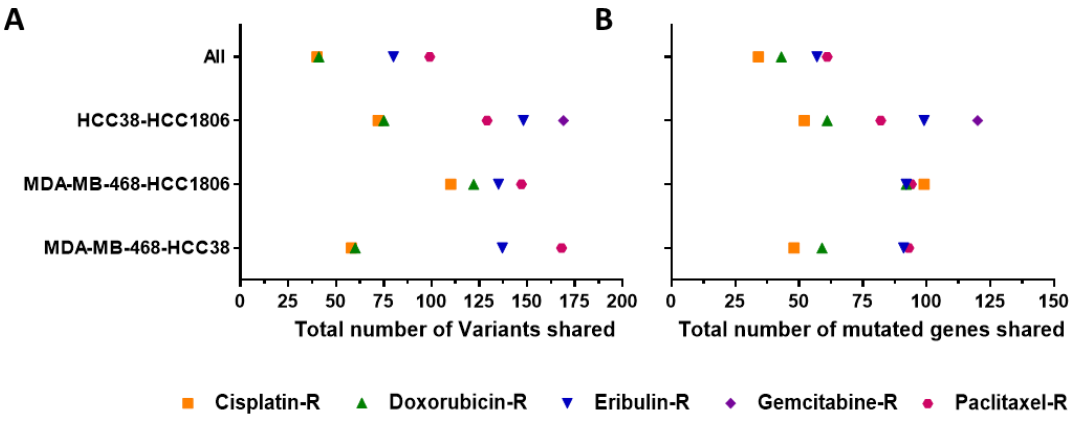


Figure 5.16 Similarity of mutational signatures in the chemo-naive and chemo-resistant TNBC cell lines
Identification of the chemo-naive and chemo-resistant TNBC cell lines similarity to Single Base Substitution (SBS) mutational signatures identified as part of the pan-cancer analysis of whole genomes (PCWAG) network. Similarity of mutational signatures calculated using Mutalisk tool and shown as a percentage similarity. Lowest BIC value identified best fit model selected for each cell line. Colours indicative of an individual signature to allow for cross-comparison between cell lines.

5.2.1.5 Comparison of variants in chemo-resistant cell lines resistant to the same drug

In order to identify variants which may have developed, or been selected for in the chemo-resistant cell lines, cell lines made resistant to the same chemotherapeutic agent were compared. This analysis aimed to identify both exact variants called, which are shared between the chemo-resistant cell lines, and also genes which contain variants (mutated genes). This distinction was made, as although an exact variant may not be shared between the cell lines, a variant in the same gene may have the same or a similar consequence. The number of variants and mutated genes that are shared between chemo-resistant cell lines developed to have resistant to the same chemotherapeutic agent were identified (Figure 5.17, variant and mutated gene data in Appendix A26-A30).



C

		Cisplatin	Doxorubicin	Eribulin	Gemcitabine	Paclitaxel
MDA-MB-468 - HCC38	S Variants	16	17	39	-	46
	NS Variants	42	43	98	-	122
	Total variants	58	60	137	-	168
	Total Genes	48	59	91	-	93
MDA-MB-468 - HCC1806	S Variants	30	35	38	-	41
	NS Variants	80	87	97	-	106
	Total Variants	110	122	135	-	147
	Total genes	99	92	92	-	94
HCC38 - HCC1806	S Variants	21	22	44	56	39
	NS Variants	51	53	104	113	90
	Total Variants	72	75	148	169	129
	Total Genes	52	61	99	120	82
MDA-MB-468 - HCC38 - HCC1806	S Variants	13	14	28	-	30
	NS Variants	27	27	52	-	69
	Total Variants	40	41	80	-	99
	Total Genes	34	43	57	-	61

Figure 5.17 Number of variants and mutated genes shared between chemo-resistant cell lines developed to have resistance to the same chemotherapeutic agent
Identification of all exact variants and mutated genes shared between each of the chemo-resistant cell lines developed to have resistance to the same chemotherapeutic agent. A) total number of variants shared between chemo-resistant cell lines B) total number of mutated genes shared between chemo-resistant cell lines C) table summarises the total number of exact variants where S are synonymous, NS are non-synonymous variants and total variants are the sum of S and NS. Total genes are the total number of mutated genes shared between the chemo-resistant cell lines. Exact variant and mutated gene data in Appendix A26-A30.

The cisplatin and doxorubicin resistant cell lines were seen to share the smallest total number of exact variants (40 and 41 respectively), whilst the gemcitabine resistant cell shared the greatest total number of exact variants (169). This is reflective of the number of variants called in the HCC38^{r}CDDP^{3000}, HCC38^{r}DOX^{40} and HCC38^{r}GEM^{20} cell lines. The paclitaxel resistant cell lines showed the highest number of variants shared between MDA-MB-468^{r}PCL^{20} and HCC38^{r}PCL$^{2.5}$, as well as MDA-MB-468^{r}PCL^{20} and HCC1806^{r}PCL^{20}, with 168 and 147 variants respectively. MDA-MB-468^{r}PCL^{20} had previously been shown to have very few variants shared with the chemo-naive cell line it was derived from.

The number of exact variants which were seen to be shared across each of the resistant sub-types, with the exception to doxorubicin resistant cell lines, were higher than the total number of mutated genes shared. This shows that in many cases multiple variants are present in the mutated genes. It must be noted that this does not necessarily mean every cell in the resistant population has multiple variants in this gene. A sub-population of the resistant cell line may harbour one variant in the gene, whilst another sub-population may harbour a different variant in the same gene.

Table 5.1 shows the genes in which exact variants were found to be shared between all of the cell lines which are resistant to the same chemotherapeutic agent, as well as the mutated genes. Taking the shared mutated genes identified, gene ontology (GO) functional enrichment analysis was conducted using G:profiler (Raudvere *et al.*, 2019). The tool maps genes to known functional information sources, such as Ensembl, and identifies statistically significant enriched terms. The GO hierarchy is divided into three sub-ontologies; Biological Process, Cellular Component and Molecular function, and each of the identified enriched GO terms are associated to one of these sub-ontologies.

All of the mutated genes sets demonstrated an enrichment in the following GO terms; O-glycan processing (GO:0016266), NADH dehydrogenase activity (GO:0003954), mitochondrial ATP synthesis coupled electron transport (GO:0042775), oxidative phosphorylation (GO:0006119), ATP synthesis coupled electron transport (GO:0042773), regulation of production of miRNAs involved in gene silencing by RNA (GO:0070920), and regulation of production of small RNA involved in gene silencing by RNA (GO:0070920). The mutated genes identified to be shared in the doxorubicin resistant cell lines and in the gemcitabine resistant cell lines were found to be enriched in the GO terms; supramolecular polymer (GO:0099081), supramolecular fibre (GO:0099512), supramolecular complex (GO:0099080), keratin filament (GO:0045095) and polymeric cytoskeletal fibre (GO:0099513). The mutated genes found in the gemcitabine resistant cells were also enriched

in the GO terms; DNA binding (GO:0003677) and negative regulation of production of miRNAs involved in gene silencing by miRNA (GO:1903799). The mutated genes shared between the eribulin resistant cell lines were also found to be enriched in the GO term; enzyme binding (GO:0019899). No other statistically significant GO terms were found to be associated with the mutated genes shared between the cisplatin resistant cell lines, or paclitaxel resistant cell lines.

Table 5.1. Summary of shared genes with exact variants and shared mutated genes between chemo-resistant cell lines

		Shared genes with exact variants	Shared mutated genes
Resistant cell lines	Cisplatin	APC, BEST1, CUX1, DSPP, EGFR, GDF15, KCNJ12, KMT2C, MT-ND4, MT-ND6, MUC2, MUC4, MUC6, NBPF1, NBP14, NCOR1, NDUFA6, PRIM2, PRSS3, RP1L1, ZDHHC11	APC, ATXN1, BEST1, C2orf16, CUX1, DSPP, EGFR, FANCA, GDF15, IGFN1, KCNJ12, KMT2C, MT-C03, MT-CYB, MT-ND4, MT-ND5, MT-ND6, MUC2, MUC4, MUC6, NBPF1, NBPF10, NBPF14, NCOR1, NCOR2, NDUFA6, NPIPB15, PRIM2, PRSS3, RGPD4, RP1L1, ZDHHC11
	Doxorubicin	CUX1, CYP2A6, DSPP, E2F4, EGFR, FRG1, FRG1B, IGFN1, IL7R, KCNJ12, KRTAP10-6, LIPT2, MAGEC1, MUC2, MUC3A, MUC4, MUC6, NBPF1, NBPF14, NDUFA6, PRDM9, PRIM2, PRSS1, PRSS3, RGPD4, RP11-492D6.3, RP1L1, SLC25A5, TPSD1, ZDHHC11	APC, ATN1, CDC27, CUX1, CYP2A6, DSPP, E2F4, EGFR, FRG1, FRG1B, IGFN1, KCNJ12, KRTAP10-6, KRTAP10-7, LIPT2, MAGEC1, MT-CO1, MT-CO3, MT-CYB, MT-ND5, MUC2, MUC3A, MUC4, MUC6, NBPF1, NBPF14, NCOR2, NDUFA6, NPIPB15, PRDM9, PRIM2, PRSS1, PRSS3, RGPD4, RP11-492D6.3, RP1L1, RPL19, SLC25A5, TP53, TPSD1, ZDHHC11
	Eribulin	APC, BMP5, BRD9, CUX1, CYP2A6, EGFR, FANCA, FMLN2, FRG1B, GDF15, GXYLT1, IGFN1, IL7R, KCNJ12, KRTAP10-6, KRTAP10-7, KRTAP5-5, LIPT2, MT-ND4, MT-ND5, MT-ND6, MUC2, MUC4, MUC6, NBPF1, NBPF14, NCRO1, NDUFA6, NOTCH3, PRNP, PRSS3, RP11-404K5.2, RP1L1, TBP, TPSD1, ZDHHC11	APC, ATXN1, ATXN2, ATXN3, BMP5, BRD9, CDC27, CUX1, CYP2A6, DSPP, EGFR, EP400, FANCA, FMNL2, FRG1B, GDF15, GOLGA6L2, GXYLT1, IGFN1, IL7R, KCNJ12, KRTAP10-6, KRTAP10-7, KRTAP5-5, LIPT2, MAGEC1, MT-CYB, MT-ND4, MT-ND5, MT-ND6, MUC2, MUC3A, MUC4, MUC6 NBPF1, NBPF10, NBPF14, NCOA3, NCOR1, NCOR2, NDUFA6, NPIPB15, PRNP, PRSS3, RGPD3, RGPD4, RP11-404K5.2, RP11-764K9.4, RP1L1, SLC9B1P1, TBP, TPSD1, TTN, ZDHHC11
	Gemcitabine	AGAP6, AKAP13, ARC, ATXN3, BMP5, CD300LD, CD33, CDK2AP2, CEP57, COCH, CUX1, CYP2A6, DDX1, DNAH, DSPP, ELF3, EPHB1, EPPK1, FAM27E2, FAM86C1, FAM86FP, FANCA, FCGBP, FLNB, FRG1, GBP7, GDF15, GOLGA6L6, GPR20, GXYLT1, IGFN1, IL7R, INADL, KCNJ12, KRTAP10-6, KRTAP10-7, KRTAP5-3, KRTAP5-5, LIPT2, MAGEB4, MAGEC1, MAP2K3, MT-CYB, MT-ND4, MT-ND6, MUC2, MUC3A, MUC4, MUC6, NBPF1, NBPF10, NBPF14, NCLN, NCOR1, NDUFA6, NOTCH3, NPIPB15, PCDHB11, PHF2, POU5F1B, PRDM9, PRIM2, PRNP, PRSS1, PRSS3, PTPLA, RAI1, RGPD4, RHPN2, RP11-492D6.3, RP11-764K9.4, RP1L1, RPL19, SAMD11, SFRP4, SMARCA2, SPRR3, SULT1A2, TBC1D29, TBP, TDG, TPSD1, UTS2, VEZF1, ZDHHC11, ZNF208, ZNF469, ZNF615, ZNF676, ZNF738, ZNF814, ZNF860	ADAMTS18, AGAP6, AKAP13, ANKRD36, ARC, ATN1, ATXN1, ATXN2, ATXN3, BMP5, CD300LD, CD33, CDC27, CDK2AP2, CEP57, CNTNAP3B, COCH, CUX1, CYP2A6, DDX11, DNAH5, DSPP, EGFR, ELF3, EP400, EPHB1, EPPK1, FAM27E2, FAM86C1, FAM86FP, FANCA, FCGBP, FLNB, FRG1, GBP7, GD15, GOLGA6L2, GOLGA6L6, GPR20, GXYLT1, IGFN1, IL7R, INADL, KCNJ12, , KRTAP10-6, KRTAP10-7, KRTAP5-3, KRTAP5-5, LIPT2, LRRIQ1, MAGEB4, MAGEC1, MAP1A, MAP2K3, MSH2, MT-ATP6, MT-CO1, MT-CYB, MT-ND3, MT-ND4, MT-ND5, MT-ND6, MUC16, MUC2, MUC3A, MUC4, MUC6, MYO5B, NBPF1, NBPF10, NBPF14, NCLN, NCOR1, NCOR2, NDUFA6, NEFH, NOTCH3, NPIPB15, OR13C5, PCDHB11, PHF2, POU5F1B, PRDM9, PRIM2, PRNP, PRSS1, PRSS3, PTPLA, RAI1, RGPD3, RGPD4, RHPN2, RNF213, RP11-492D6.3, RP11-764K9.4, RP1L1, RPL19, SAMD11, SFRP4, SMARCA2, SPRR3, SULT1A2, TAS2R46, TBC1D29, TBP, TDG, TP53, TPSD1, TTN, UTS2, VEZF1, ZDHHC11, ZNF208, ZNF469, ZNF615, ZNF676, ZNF738, ZNF814, ZNF860
	Paclitaxel	APC, CD300LD, CEP57, CYP2A6, DNAH5, EGFR, EPHB1, GXYLT1, IGFN1, IL7R, KCNJ12, KRTAP10-6, KRTAP5-3, LIPT2, MT-CO3, MT-CYB, MT-ND4, MT-ND5, MT-ND6, MUC2, MUC3A, MUC4, MUC6, NBPF1, NBPF14, NCOR1, NPIPB15, PHF2, POU5F1B, PRIM2, PRSS1, PRSS3, RGPD4, RHPN2, RP11-492D6.3, RP11-764K9.4, RP1L1, SULT1A2, TBC1D29, TBP, TPSD1, ZNF208, ZNF676, ZNF814	ADAMTS18, APC, ATXN1, CD300LD, CEP57, CYP2A6, DNAH5, DSPP, EGFR, EP400, EPHB1, FRG1B, GOLGA6L2, GXYLT1, IGFN1, IL7R, KCNJ12, KRTAP10-6, KRTAP5-3, LIPT2, MAGEC1, MT-ATP6, MT-CO3, MT-CYB, MT-ND3, MT-ND4, MT-ND5, MT-ND6, MUC16, MUC2, MUC3A, MUC4, MUC6, MYO5B, NBPF1, NBPF10, NBPF14, NCOR1, NCOR2, NPIPB15, PRIM2, PRSS1, PRSS3, RGPD3, RGPD4, RHPN2, RNF213, RP11-492D6.3, RP11-764K9.4, RP1L1, SULT1A2, TBC1D29, TBP, TPSD1, ZNF208, ZNF676, ZNF814

Having established the total number of variants and mutated genes shared between the resistant cell lines made resistant to the same chemotherapeutic agents, the analysis next focused on the *de novo* variants that are shared (Table 5.2).

No *de novo* variants were found to be shared between resistant cell lines derived from each of the three chemo-naive cell lines. The two gemcitabine resistant cell lines; HCC38^{r}GEM^{20} and HCC1806^{r}GEM^{20} shared a *de novo* variant in *DDX11*, an ATP-dependent DNA helicase, although the VAF was low (0.086 and 0.176 respectively). Several *de novo* variants in *CDK2AP2* were identified to be shared between HCC38^{r}DOX^{40} and HCC1806^{r}DOX$^{12.5}$, and several *de novo* variants were identified in *PABPC3* shared between MDA-MB-468^{r}ERI^{50} and HCC1806^{r}ERI^{50}. MDA-MB-468^{r}ERI^{50} and HCC38^{r}ERI^{10} were shown to share non-synonymous *de novo* variants in *RGPD3* and *RGPD4* which encode RanBP2-like and GRIP domain-containing protein 3 and RanBP2-like and GRIP domain-containing protein 4. Both *RGPD3* and *RGPD4* are one of the eight copies of RANBP2 clustered close to the chromosome 2 centromere. RANBP2 is a E3 SUMO-protein ligase which has a role in facilitating SUMO1 and SUMO2 conjugation by UBE2I (Pichler *et al.*, 2002; Reverter and Lima, 2005; Gareau, Reverter and Lima, 2012). Many cell lines were seen to share *de novo* variants in MUC related genes.

Next, genes which were only found to be mutated in the chemo-resistant cell lines, not in the chemo-naive (*de novo* mutated genes), were compared to see if any of these are shared between chemo-resistant cell lines developed to have resistance to the same drug. No *de novo* mutated genes were found to be shared.

The analysis next focused on variants which were *not called* in the chemo-resistant cell lines, and compared these to chemo-resistant cell lines made resistant to the same chemotherapeutic agents (Table 5.3). Many variants identified are associated with MUC genes.

Next, taking the genes which were only found to be mutated in the chemo-naive cell lines, not in the chemo-resistant cell lines (*not called* mutated genes), these were compared to see if any of these are shared between chemo-resistant cell lines developed to have resistance to the same drug. No *not called* mutated genes were found to be shared.

Taken together, these analyses have shown that the variants which may be driving resistance are different in each of the chemo-resistant cell lines, even when made resistant to the same chemotherapeutic agent.

Table 5.2. Summary of shared *de novo* variants between chemo-resistant cell lines

		MDA-MB-468-HCC38-HCC1806	HCC38-HCC1806	MDA-MB-468-HCC38	MDA-MB-468-HCC1806
Chemo-resistant cell lines	**Cisplatin**	0	0	0	Chr7, 100551263T>C, MUC3A, NS, 0.130[a], 0.158
	Doxorubicin	0	Chr11, 67275116G>A, CDK2AP2, NS, 0.214[b], 0.143[c]; Chr11, 67275117A>G, CDK2AP2, S, 0.25[b], 0.188[c]; Chr11, 67275120G>A, CDK2AP2, S, 0.25[b], 0.1875[c]; Chr11, 67275121C>G, CDK2AP2, NS, 0.25[b], 0.188[c]	0	0
	Gemcitabine	-	Chr12, 31237978C>T, DDX11, NS, 0.086[b], 0.176[c]	-	-
	Eribulin	0	Chr1, 145293269G>A,RP11-458D21.5, S, 0.458[b], 0.357[c]; Chr11, 47647265A>G, MTCH2, NS, 0.286[b], 0.313[c]; Chr15, 28518112C>A, HERC2, NS, 0.333[b], 0.188[c]	Chr2, 107049714C>G, RGPD3, NS, 0.357[a], 0.32[b]; Chr2, 108488742G>T, RGPD4, NS, 0.159[a], 0.276[b]; Chr11, 1017068C>T, MUC6, S, 0.118[a], 0.170[b]; Chr11, 1017074A>G, MUC6, S, 0.143[a], 0.164[b]; Chr11, 1017088A>C, MUC6, NS, 0.231[a], 0.173[b]; Chr11, 1017092G>C, MUC6, NS, 0.261[a], 0.177[b]	Chr8, 124382158DelATC, ATAD2, NS, 0.167[a], 0.164[c]; Chr11, 134855496C>G, AP003062.1, NS, 0.172[a], 0.225[c]; Chr13, 25670907C>A, PABPC3, NS, 0.353[a], 0.667[c]; Chr13, 25670919A>G, PABPC3, NS, 0.412[a], 0.684[c]; Chr13, 25670955C>T, PABPC3, NS, 0.357[a], 0.733[c]
	Paclitaxel	0	0	Chr3, 195507193G>A, MUC4, NS, 0.25[a], 0.241[b]; Chr17, 5036823T>C, USP6, NS, 0.158[a], 0.269[b]	Chr11, 1092467C>T, MUC2, NS, 0.214[a], 0.231[c]; Chr11, 1092468A>G, MUC2, S, 0.286[a], 0.308[c]; Chr17, 39274364T>G, KRTAP4-11, NS, 0.346[a], 0.353[c]

Key: Chromosome, variant position, reference base > alternative base, NS/S (non-synonymous/synonymous), VAF MDA-MB-468 = [a], HCC38 = [b], HCC1806 = [c]

Table 5.3. Summary of shared *not called* variants between chemo-resistant cell lines

		MDA-MB-468-HCC38-HCC1806	HCC38-HCC1806	MDA-MB-468-HCC38	MDA-MB-468-HCC1806
Chemo-resistant cell lines	**Cisplatin**	Chr11, 1017592G>C, MUC6, NS, 0.15[a], 0.107[b], 0.196[c]; Chr11, 1017597T>G, MUC6, NS, 0.167[a], 0.107[b], 0.2[c]	Chr2,47641558DELTaaa, MSH2, NS, 0.291[b], 0.24[c]; Chr3, 195505772C>G, MUC4, NS, 0.212[b], 0.122[c]; Chr11, 1017592G>C, MUC6, NS, 0.107[b], 0.196[c]; Chr11, 1017597T>G, MUC6, NS, 0.107[b], 0.2[c]; Chr11, 9015382C>T, MUC16, NS, 0.2[b], 0.131[c]	Chr11, 1016835A>G, MUC6, NS, 0.106[a], 0.085[b]; Chr11, 1017592G>C, MUC6, NS, 0.15[a], 0.107[b]; Chr11, 1017597T>G, MUC6, NS, 0.167[a], 0.107[b]	Chr11,1017573G>C, MUC6, NS, 0.109[a], 0.089[c]; Chr11,1017592G>C, MUC6, NS, 0.177[a], 0.196[c]; Chr11,1017597,T>G, MUC6, NS, 0.167[a], 0.2[c]
	Doxorubicin	0	Chr3, 195505772C>G, MUC4, NS, 0.212[b], 0.122[c]; Chr11, 1017573G>C, MUC6, NS, 0.0869[b], 0.0892[c]; Chr11, 9015382C>T, MUC16, NS, 0.2[b], 0.131[c]	Chr1, 145360608T>A, NBPF10, NS, 0.211[a], 0.235[b]; Chr8, 101719138C>T, PABPC1, NS, 0.25[a], 0.167[b]; Chr11, 1017566G>A, MUC6, S, 0.123[a], 0.188[b]; Chr11, 1017679A>G, MUC6, NS, 0.286[a], 0.333[b]; Chr11, 1092636C>T, MUC2, S, 0.167[a], 0.162[b]; Chr11, 1092647C>T, MUC2, NS, 0.123[a], 0.149[b]	0
	Gemcitabine	-	Chr3, 195505772C>G, MUC4, NS, 0.212[b], 0.122[c]; Chr7, 100677974A>G, MUC17, NS, 0.234[b], 0.125[c]; Chr11, 1017573G>C, MUC6, NS, 0.087[b], 0.089[c]; Chr11, 1017597, T>G, MUC6, NS, 0.107[b], 0.2[c]; Chr17, 45234350C>T, CDC27, S, 0.105[b], 0.148[c]; Chr17, 45234367A>T, CDC27, NS, 0.108[b], 0.2[c]	-	-
	Eribulin	Chr19,58385762,C>G, ZNF814, S, 0.333[a], 0.217[b], 0.294[c]	Chr11,1092599InsAAT, MUC2, S, 0.668[b], 0.529[c]; Chr17,45234350C>T, CDC27, S, 0.105[b], 0.148[c]; Chr17,45234367A>T, CDC27, NS, 0.108[b], 0.2[c]; Chr19,9015382C>T, MUC16, NS, 0.2[b], 0.073[c]; Chr19,58385762C>G, ZNF814, S, 0.217[b], 0.294[c]	Chr7,100677974A>G, MUC17, NS, 0.156[a], 0.155[b]; Chr19,58385762C>G, ZNF814, S, 0.333[a], 0.217[b]	Chr11,1092598InsCAA, MUC2, NS, 0.389[a], 0.567[c]; Chr17,45266534G>A, CDC27, NS, 0.364[a], 0.238[c]; Chr19,58385762C>G, ZNF814, S, 0.333[a], 0.294[c]
	Paclitaxel	0	Chr2,47641558DELTaaa, MSH2, NS, 0.291[b], 0.24[c]; Chr3, 195505772C>G, MUC4, NS, 0.212[b], 0.122[c]	Chr1,145360608,T>A, NBPF10, NS, 0.211[a], 0.235[b]; Chr7,100677974A>G, MUC17, NS, 0.156[a], 0.234[b]; Chr8,101719138C>T, PABPC1, NS, 0.25[a], 0.167[b]; Chr11,1016835A>G, MUC6, NS, 0.106[a], 0.085[b]; Chr17,45214528A>T, CDC27, NS, 0.267[a], 0.455[b]	Chr1,152129115T>C, RPTN, NS, 0.231[a], 0.238[c]; Chr6,55739553A>G, BMP5, S, 0.517[a], 0.231[c], Chr7,56087399G>A, PSPH, NS, 0.286[a], 0.3[c]; Chr7,151970877G>A, KMT2C, NS, 0.418[a], 0.121[c]

Key: Chromosome, variant position, reference base > alternative base, NS/S (non-synonymous/synonymous), VAF MDA-MB-468 = [a], HCC38 = [b], HCC1806 = [c]

5.2.2 Candidate approach

As established in chapter 4, dysregulation of the DDR pathways is a common resistance mechanism for chemotherapeutic agents. Further to this, dysregulation of the way in which drugs are transported or metabolised can also result in chemo-resistance (section 1.3, chapter 4). Here, a candidate approach was taken to analyse variants in genes associated with DDR and drug transport and metabolism in each of the chemo-naive and chemo-resistant cell lines. Five lists of genes with GO terms associated to; Cellular response to DNA damage stimulus (747 genes), DNA repair (469 genes), Cell cycle (1658 genes), Drug transport (208 genes) and Drug metabolic Process (532 genes), were created. These lists were compared to the mutated genes in each of the chemo-naive and chemo-resistant cell lines to identify variants in these pathways. As the severity outcome of synonymous variants is difficult to predict, only non-synonymous variants were used in this analysis. The number of genes with only *de novo* variants were also calculated in the chemo-resistance TNBC cell lines for each category.

The total number of mutated genes for each cell line associated with the five described categories was identified (Figure 5.18A-E). The genes in the DNA damage stimulus and DNA repair categories were considerably overlapping, as many of the proteins derived from the genes have promiscuous roles across these cross-talking pathways. Here, the same number of mutated genes are the same to both of these categories in all of the chemo-naive and chemo-resistant cell lines, with an exception to HCC38^{r}ERI^{10} which identified an extra mutated gene associated to DNA repair, and MDA-MB-468^{r}PCL^{20}, HCC38 and HCC38^{r}CDDP^{3000} which identified an extra mutated gene associated to DNA damage stimulus. All of the HCC1806 derived chemo-resistant cell lines showed an increase in mutated genes associated to the cell cycle compared to HCC1806. HCC38^{r}GEM^{20} had the most mutated genes associated to the cell cycle across all of the cell lines in the panel. MDA-MB-468^{r}CDDP^{1000} had the highest increase in mutated genes associated to the cell cycle compared to other MDA-MB-468 derived cell lines. Very few mutated genes were found in the chemo-resistant cell lines associated to the GO terms; drug transport or drug metabolic process. Both MDA-MB-468^{r}CDDP^{1000} and HCC1806^{r}CDDP^{500} had an increase in mutated genes associated to drug metabolic process, compared to the respective chemo-naive cell line.

The genes which contain *de novo* variants for each of the categories in the chemo-resistant TNBC cell lines were identified (Figure 5.18F). Both the gemcitabine resistant cell lines (HCC38^{r}GEM^{20} and HCC1806^{r}GEM^{20}) had many *de novo* variants in genes associated with the cell cycle. All three cisplatin resistant cell lines (MDA-MB-468^{r}CDDP^{1000}, HCC38^{r}CDDP^{3000} and HCC1806^{r}CDDP^{500}) had the most *de novo* variants in genes associated to drug metabolic processes. Each of the *de novo* variants for each category are summarised with respect to the overlapping GO terms, including the VAF, in Appendix A32.

In order to determine how damaging the *de novo* variants were, several approaches were taken. Those variants which are considered to be protein truncating variants (PTV), such as; frameshifts, stop gains, splice accepter and splice donor variants, were considered to have sudden protein termination, or incorrect protein folding, or chain composition due to changes in the reading frame. This would result in a protein that is predicted to be either; completely no-functional in the cell, ultimately undergoing rapid degradation, or have heavily dysregulated activity. Taking these assumptions, these variants were considered to have loss of function (LOF). Those variants which are considered to be missense variants underwent further analysis using three variant predictors; SIFT, Polyphen and mutational assessor (section 2.4.3.3). The scores obtained from the missense variant predictors do not give any indication as to whether the missense variant could be a gain of function (GOF), or a LOF. For the following analysis, only the variants which are considered to be damaging by at least two out of three predictors are addressed in full.

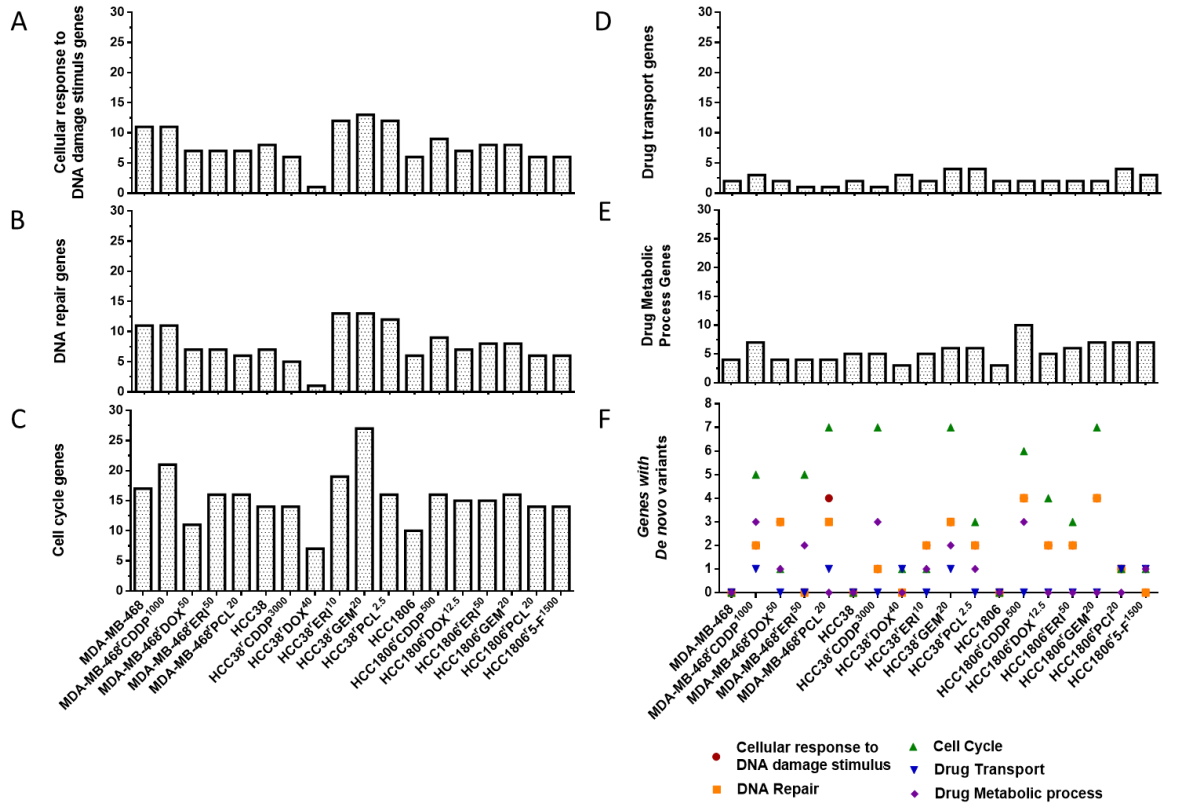


Figure 5.18 Summary of mutated genes identified in the candidate approach
Number of mutated genes in the chemo-naive and chemo-resistant cell lines which are described by the gene ontology terms A) Cellular response to DNA damage stimulus, B) DNA repair, C) Cell cycle, D) Drug transport, E) Drug metabolic processes. F) Identification of genes under the aforementioned gene ontology terms which contain only *de novo* variants in the chemo-resistant cell lines.

5.2.2.1 De novo variants in cisplatin resistant cell lines

Across the three cisplatin resistant cell lines, 33 *de novo* variants in this candidate approach were identified which may be driving cisplatin resistance in the genes; *EEPD1, SMARCAL1, BRD7, EPB41, FER, HNRNPU, KLF11, TERT, SLC25A19, ADH7, DNM2, EPAS1, IDO1, FANCM, IGF1R, MICAL3, SLC16A1, TDRD9, ZBTB17, CYP27B1, GCH1, GCSH, CLOCK, DDX11, NCOA6, RAD23B, PTGS2, RAE1, SDCCAG8, ATP5O* and *SDHA* (Appendix-A31) Table 5.4 shows the results of the missense variant predictors analysis.

Table 5.4. Summary of cisplatin resistant missense variants effect predictor results

Cell line	Gene	SIFT	Polyphen	Mutational assessor			
				Functional impact	Protein bind site	DNA/RNA bind site	Small molecule bind site
MDAMB468^{r}CDDP^{100}	EEPD1	tolerated(0.36)	benign(0.354)	low	-	-	-
MDAMB468^{r}CDDP^{100}	EPB41	tolerated(1)	benign(0)	low	-	-	-
MDAMB468^{r}CDDP^{100}	HNRNPU	deleterious(0)	probably_damaging(0.914)	medium	-	-	-
MDAMB468^{r}CDDP^{100}	TERT	tolerated(0.38)	benign(0.18)	medium	-	-	-
MDAMB468^{r}CDDP^{100}	SLC25A19	tolerated(0.84)	benign(0)	neutral	-	-	CM5
MDAMB468^{r}CDDP^{100}	ADH7	deleterious(0.04)	probably_damaging(0.934)	low	-	-	-
MDAMB468^{r}CDDP^{100}	DNM2	tolerated(0.08)	possibly_damaging(0.699)	-	-	-	1PE
MDAMB468^{r}CDDP^{100}	EPAS1	tolerated(0.07)	benign(0.061)	medium	-	-	-
HCC38^{r}CDDP^{1000}	FANCM	tolerated(0.11)	benign(0.055)	low	-	-	-
HCC38^{r}CDDP^{1000}	IGF1R	tolerated(0.59)	benign(0.164)	low	-	-	NAG
HCC38^{r}CDDP^{1000}	LIN54	deleterious(0.01)	probably_damaging(0.996)	low	-	-	-
HCC38^{r}CDDP^{1000}	MICAL3	deleterious_low_confidence(0)	benign(0.347)	low	-	-	-
HCC38^{r}CDDP^{1000}	SLC16A1	tolerated(0.52)	benign(0.079)	neutral	-	-	-
HCC38^{r}CDDP^{1000}	SLC16A1	deleterious(0.04)	possibly_damaging(0.708)	low	-	-	-
HCC38^{r}CDDP^{1000}	TDRD9	deleterious(0)	probably_damaging(0.999)	medium	-	-	-
HCC38^{r}CDDP^{1000}	ZBTB17	deleterious(0.01)	probably_damaging(0.924)	medium	1	1	-
HCC38^{r}CDDP^{1000}	CYP27B1	tolerated(1)	benign(0.006)	neutral	1	-	-
HCC38^{r}CDDP^{1000}	GCH1	deleterious(0)	probably_damaging(0.988)	high	1	-	-
HCC38^{r}CDDP^{1000}	GCSH	deleterious(0.05)	benign(0.171)	medium	1	-	-
HCC1806^{r}CDDP^{500}	CLOCK	deleterious(0.02)	benign(0.02)	medium	-	-	-
HCC1806^{r}CDDP^{500}	DDX11	deleterious(0.01)	benign(0.424)	low	-	-	-
HCC1806^{r}CDDP^{500}	NCOA6	deleterious(0)	probably_damaging(0.996)	medium	-	-	-
HCC1806^{r}CDDP^{500}	NCOA6	deleterious(0)	probably_damaging(0.997)	medium	-	-	-
HCC1806^{r}CDDP^{500}	RAD23B	tolerated(0.26)	possibly_damaging(0.798)	low	-	-	-
HCC1806^{r}CDDP^{500}	CDK2AP2	deleterious(0)	probably_damaging(0.973)	medium	-	-	-
HCC1806^{r}CDDP^{500}	CDK2AP2	tolerated(0.67)	probably_damaging(0.96)	neutral	-	-	-
HCC1806^{r}CDDP^{500}	CDK2AP2	tolerated(0.27)	benign(0)	neutral	-	-	-
HCC1806^{r}CDDP^{500}	CDK2AP2	tolerated(0.09)	possibly_damaging(0.908)	neutral	-	-	-
HCC1806^{r}CDDP^{500}	PTGS2	deleterious(0)	probably_damaging(0.999)	medium	-	-	PGX, DIF, ACD, EPA
HCC1806^{r}CDDP^{500}	RAE1	deleterious(0)	possibly_damaging(0.693)	medium	-	-	-
HCC1806^{r}CDDP^{500}	ATP5O	deleterious(0)	possibly_damaging(0.539)	low	-	-	-
HCC1806^{r}CDDP^{500}	SDHA	deleterious_low_confidence(0.01)	possibly_damaging(0.52)	high	1	-	-

A stop gain variant in *SMARCAL1* was identified in MDA-MB-468^{r}CDDP^{1000} cell line at W277* with a VAF of 0.124. SMARCAL1 (SWI/SNF-related matrix-associated-actin-dependent regulator of chromatin subfamily A-like protein 1), is an ATP-dependent annealing helicase that binds selectively to DNA, and is involved in fork reversal activity. Depletion of SMARCAL1 restores fork protection in BRCA1 or BRCA2 depleted cells, and increased drug resistance to both cisplatin and PARP inhibitors

(Cantor and Calvo, 2017; Kolinjivadi *et al.*, 2017; Taglialatela *et al.*, 2017). It could be predicted the stop gain variant leads to LOF of SMARCAL1, prevention of fork reversal, restoring fork protection and appropriate DNA repair to overcome cisplatin lesions.

MDA-MB-468^{r}CDDP^{100} harbours a splice acceptor variant in *BRD7* with a VAF of 0.23. BRD7 interacts and negatively regulates YB-1 phosphorylation level. YB-1 has been found to be implicated in response to multiple drugs such as cisplatin (Niu *et al.*, 2020). It has been shown that YB-1 is overexpressed in human cancer cell lines which are resistant to cisplatin, and YB-1 can function as a recognition protein for cisplatin intercalated-DNA (Ohga *et al.*, 1996; Tomoko *et al.*, 1999). It could be predicted that a splice acceptor variant leads to LOF of BRD7, preventing negative regulation of YB-1, and allowing YB-1 to respond and promote DNA repair to cisplatin lesions.

A missense variant in *HNRNPU* was identified in MDA-MB-468^{r}CDDP^{100} at F299Y. HNRNPU (heterogenous nuclear ribonucleoprotein U) has many roles associated with nuclear chromatin organisation and mitotic cell progression to name but a few. HNRNPU has been shown to enhance cisplatin-induced apoptosis through up-regulation of GADD45A, which functions to stimulate DNA excision repair and inhibits entry of cells into S phase (Li *et al.*, 2017). SIFT predicts this variant to be deleterious, Polyphen; probably damaging and mutational assessor; medium (Table 5.4). The variant was identified to be in the B30.2/SPRY domain, which is predicted to function through protein-protein interaction, and are normally highly conserved domains (Woo *et al.*, 2006). It could be predicted that a LOF in HNRNPU, preventing cisplatin induced apoptosis, which could result in cisplatin resistance.

A splice donor variant in *IDO1* was identified in MDA-MB-468^{r}CDDP^{100}. IDO1 (Indoleamine 2,3-dioxygenase 1) is found in the kynurenine pathway, and catalyses the first and rate limiting step of the catabolism of the amino acid tryptophan (Metz *et al.*, 2007). Cisplatin resistant lung cancer cells have been shown to activate the kynurenine pathway to cope with excessive reactive oxygen species, and increased IDO1 activities are found in these cisplatin resistant cells. The group also established that shRNA knockout of IDO1 in cisplatin resistant cells resulted in a significant growth inhibitory effect (Nguyen *et al.*, 2020). Here predictions of LOF variant in *IDO1* as a mechanism of cisplatin resistance contradict what has previously been found. This mechanism may be context specific.

A *de novo* missense variant was found in *GCH1* (G108S) in HCC38^{r}CDDP^{3000} with a VAF of 0.2. GCH1 (GTP cyclohydrolase 1), has been found to positively regulated nitric oxide (NO) synthesis. Increased NO production has previously been implicated in cisplatin drug resistance through regulation of Bcl-

2, and cisplatin treatment of rats reported a beneficial effect of blockade of NO pathways during cisplatin chemotherapy (Srivastava *et al.*, 1996; Chanvorachote *et al.*, 2006). All three variant predictors identify the variant in *GCH1* to be damaging, and mutation assessor identifies it to be in a protein binding site, and next to a known conserved site (Table 5.4). The variant could be a GOF variant which leads to increase in NO resulting in cisplatin resistance.

A *de novo* missense variant was found in *PTGS2* (G512E) in HCC1806^{r}CDDP^{500} with a VAF of 0.3. PTGS2 (prostaglandin G/H synthase 2), is involved in prostanoid synthesis. Upregulation of PTGS2 has been associated to cell adhesion and resistance to apoptosis through the production of prostaglandin E2. Drugbank, a bioinformatics tool for drug target discovery and drug docking, has shown that cisplatin binds to PTGS2 (Wishart, 2006). It has also been shown that transactivation of PTGS, by PAX5 signalling, potentiates cisplatin resistance (Dong *et al.*, 2018). All three variant predictors identify the variant to be damaging in PTGS2, with mutation assessor identifying the variant to be in a small molecule binding site (Table 5.4).

A *de novo* missense variant was found in *SDHA* (A449V) in HCC1806^{r}CDDP^{500} with a VAF of 0.3. SDHA (succinate dehydrogenase [ubiquinone] flavoprotein subunit, mitochondrial) is a protein involved in complex II of the mitochondrial electron transport chain (Burnichon *et al.*, 2010). Cisplatin is known to induce a mitochondrial ROS response, of which contributes to the cytotoxicity of the drug (Marullo *et al.*, 2013b). The three variant predictors identified this to be damaging, with mutational assessor identifying the variant to be in a protein binding site (Table 5.4). This variant may reduce the damage of the ROS in the mitochondria, resulting in reduced DNA damage and reduced cell death as a possible cisplatin resistance mechanism.

Several *de novo* missense variants, which are predicted to be damaging, were also identified in the cisplatin resistant cells lines, of which the resistance mechanism could not be hypothesised. MDA-MB-468^{r}CDDP^{100} was identified to have a stop gain variant in FER (W20*) and KLF11 (S442*). FER (tyrosine-protein kinase FER), is a cytoplasmic tyrosine protein kinase that acts downstream of cell surface receptors. KLF11 (krueppel-like factor 11) is a transcription factor which can repress transcription of SMAD7, antagonist of TGF-beta signalling.

A *de novo* missense variant was identified in *LIN54* (P448T) in HCC38^{r}CDDP^{3000} with a VAF of 0.5. LIN54 (protein lin-54 homolog), is a key component of the DREAM complex, a complex which acts as a coordinator of cell-cycle-dependent gene expression, either as an activator or a repressor depending on the context (Litovchick *et al.*, 2007; Schmit *et al.*, 2007; Sadasivam and DeCaprio, 2013). Two of the variant predictors consider it to be a damaging variant, although the variant is

not present in any known domains (Table 5.4). The DREAM complex is part of the p53-p21-DREAM-E2F/CHR pathway. It is known that $p21^{CIP1/WAF1}$ is required for the activation of the DREAM complex (Engeland, 2018).The $HCC38^{r}CDDP^{3000}$ cell line was previously investigated in chapter 3 and it was found to have elevated $p21^{CIP1/WAF1}$ expression at both gene and protein level. However, the link between the elevated $p21^{CIP1/WAF1}$ levels in $HCC38^{r}CDDP^{3000}$ and the *de novo* missense variant in *LIN54* and the implication in cisplatin resistance remains unknown.

Interestingly, another *de novo* variant identified in $HCC38^{r}CDDP^{3000}$, has been found to have a role implicated with $p21^{CIP1/WAF1}$. A *de novo* missense variant was found in *ZBTB17* (T439P) in $HCC38^{r}CDDP^{3000}$ with a VAF of 0.2. ZBTB17 (zinc finger and BTB domain -containing protein 17), is a transcription factor which can act as both an activator or repressor and targets negative regulators of cell cycle progression, including inducing transcription of $p21^{CIP1/WAF1}$ /*CDKN1A* and the retinoblastoma gene *RB1* upon exposure to genotoxic stress (Jeon *et al.*, 2014). All three variant predictors consider the variant to be damaging, and mutation assessor predicted the variant to be located in a protein and a DNA/RNA binding site, of which is identified to be in a zinc finger (C2H2-type 5) (Table 5.4). It could be predicted that this variant could result in higher transcription of $p21^{CIP1/WAF1}$ /*CDKN1A*, resulting in the higher expression levels of $p21^{CIP1/WAF1}$ seen experimentally in $HCC38^{r}CDDP^{3000}$, although the implication of the *de novo* missense variant in *ZBTB17* role in cisplatin resistance is unknown.

A *de novo* missense variant was found in *TDRD9* (L273H) in $HCC38^{r}CDDP^{3000}$ with a VAF of 0.4. TDRD9 (ATP-dependent RNA helicase TDRD), is predominantly associated with meiotic nuclear division. The variant is considered to be damaging by all three variant predictors, and can be found to be in the helicase ATP-binding domain (Table 5.4). Its implication in cisplatin resistance is unknown.

A *de novo* variant was found in *RAE1* (P228T) in $HCC1806^{r}CDDP^{500}$ with a VAF equal of 0.3. RAE1 (mRNA export factor) is known to have a role in the mitotic bipolar spindle formation (Wong, Blobel and Coutavas, 2006). All three variant predictors identify this variant as damaging, and it is identified to be directly next to a threonine residue at position 229, which is known to be phosphorylated (Table 5.4). Its implication in cisplatin resistance is unknown.

A missense variant in *ADH7* was identified in MDA-MB-$468^{r}CDDP^{100}$ at G129V. ADH7 (all-trans retinol dehydrogenase) catalyses the NAD-dependent oxidation of all transoretinol, alcohol, and omega-hydroxy fatty acids and their derivatives. SIFT predicted the variant to be deleterious, Polyphen; probably damaging and mutational assessor; low (Table 5.4). Its implication in response to cisplatin is unknown.

5.2.2.2 De novo variants in doxorubicin resistant cell lines

Across the three doxorubicin resistant cell lines, 10 *de novo* variants in this candidate approach were identified which may be driving doxorubicin resistance in the genes; *CHEK2, MTA1, XRN2, MTRR, CSNK1D, TOR1A, CLOCK, TOP2A, CDK2AP2* and *SOX2* (Appendix-A31). Table 5.5 shows the results of the missense variant predictors analysis.

Table 5.5. Summary of doxorubicin resistant missense variants effect predictor results

Cell line	Gene	SIFT	Polyphen	Mutational assessor			
				Functional impact	Protein bind site	DNA/RNA bind site	Small molecule bind site
MDAMB468^{r}DOX^{50}	CHEK2	deleterious(0)	probably_damaging(0.949)	neutral	1	-	TPO
MDAMB468^{r}DOX^{50}	MTA1	tolerated(0.51)	benign(0.003)	neutral	-	-	-
MDAMB468^{r}DOX^{50}	XRN2	deleterious(0)	probably_damaging(0.995)	medium	-	-	-
HCC38^{r}DOX^{40}	CSNK1D	deleterious_low_confidence(0.01)	benign(0.149)	medium	1	-	-
HCC38^{r}DOX^{40}	TOR1A	tolerated(1)	benign(0.009)	neutral	-	-	-
HCC1806^{r}DOX$^{12.5}$	CLOCK	deleterious(0.02)	benign(0.02)	medium	-	-	-
HCC1806^{r}DOX$^{12.5}$	CDK2AP2	deleterious(0.02)	benign(0.258)	low	-	-	-
HCC1806^{r}DOX$^{12.5}$	CDK2AP2	deleterious(0)	probably_damaging(0.973)	medium	-	-	-
HCC1806^{r}DOX$^{12.5}$	CDK2AP2	tolerated(0.67)	probably damaging(0.96)	neutral	-	-	-
HCC1806^{r}DOX$^{12.5}$	CDK2AP2	tolerated(0.27)	benign(0)	neutral	-	-	-
HCC1806^{r}DOX$^{12.5}$	CDK2AP2	tolerated(0.09)	possibly damaging(0.908)	neutral	-	-	-

HCC1806^{r}DOX$^{12.5}$ carried a frameshift mutation in *TOP2A* with a VAF of 0.8. TOP2A (topoisomerase 2 alpha) is a known target of doxorubicin and its dysregulation is known to play a role in doxorubicin resistance (Burgess *et al.*, 2008). The frameshift occurs at PI843-844PIYX, introducing an early nonsense mutation, resulting in the loss of sites associated with the interaction of TOP2A with DNA, and many protein translation modification sites (Wendorff, Timothy J.; Schmidt, Bryan H.; Heslop, Pauline; Austin, Caroline A.; Berger, 2012). Loss of the drug target reduces the way in which doxorubicin cytotoxicity can occur. It can be predicted that the *de novo* frameshift variant in *TOP2A* directly results in doxorubicin resistance.

A frameshift variant in *SOX2* (LL131-132X) was also identified in HCC1806^{r}DOX$^{12.5}$. SOX2, is a transcription factor which, through a functional study, was shown to be a target for *TOP2A* (Lachmann *et al.*, 2010). It has been further implicated to be a target for *TOP2A* through a multidimensional protein identification technology (MudPIT) analysis (Gao *et al.*, 2012). Here it could be predicted that the identified frameshift in *SOX2* could result in a LOF phenotype, resulting in reduced transcription of *TOP2A*. This could result in a reduction of available target for doxorubicin binding, resulting in doxorubicin resistance.

HCC38^{r}DOX^{40} has a missense variant S101N in *CSNK1D* with a VAF of 0.6. CSNK1D (casein kinase 1, delta), an essential serine/threonine protein kinase, and a central component of the circadian clock, has been shown to phosphorylates TOP2A at S1106 which enhances TOP2A activity and results in sensitivity to Topo II – targeted drugs in vivo (Chikamori *et al.*, 2003; Grozav *et al.*, 2009). The missense variant changes a serine, to an asparagine, and although not previously identified as a site of phosphorylation, the variant was identified to be in a protein binding site by mutation assessor (Table 5.5). It could be hypothesised that the missense variant in *CSNK1D*, could lead to a downregulation in CSNK1D activity, or reduction of protein-protein-interaction, resulting in the hypophosphorylation of TOP2A, and ultimately resistance to doxorubicin.

Given the role of CSNK1D in the circadian clock, it was interesting to find a *de novo* missense variant, E165A, in HCC1806^{r}DOX$^{12.5}$ in the gene; *CLOCK*. CLOCK (circadian locomotor output cycles kaput), is a transcriptional activator which forms a core component of the circadian clock, to the best of knowledge has not be implicated in doxorubicin resistance. Two of the variant predictors identify this variant to be damaging (Table 5.5). Further to this, a *de novo* missense variant in *MTA1*, H614N, in MDA-MB-468^{r}DOX^{50} was identified. MTA1 (metastasis associated 1), is a transcriptional coregulator, and has been found to recruit CLOCK-BMAL1 heterodimer to its promotors to promote transcription (Li *et al.*, 2013). However, each of the variant predictors identified the variant to be tolerated/benign/neutral.

A *de novo* missense variant was also identified in *XRN2* (R362G) in MDA-MB-468^{r}DOX^{50}. XRN2 (5'-3' exoribonuclease 2) has been implicated in promotion of the termination of transcription by RNA polymerase II (West, Gromak and Proudfoot, 2004). The missense variant has been predicted to be damaging by all variant predictors (Table 5.5), but its role in doxorubicin resistance is yet to be determined.

As dysregulation of TOP2A has been established in literature as a mechanism of resistance of TOP2A targeted inhibitors, western blotting analysis of TOP2A was conducted in HCC1806 and HCC1806^{r}DOX$^{12.5}$ (Burgess *et al.*, 2008). TOP2A protein expression was reduced in HCC1806^{r}DOX$^{12.5}$ cell compared to HCC1806 at basal levels (Appendix A32, Figure A32.1). In order to investigate if reduced protein expression of TOP2A results in doxorubicin resistance, TOP2A targeted siRNA was reverse transfected into HCC1806. The conditions for transfection were optimised for HCC1086, as before in section 4.2.6.2 (Appendix A32, Figure A32.2). Here it was determined that a concentration of 0.05% lipofectamine 2000 is required for transfection, and cells are to be plated at 12800 cells/well (Appendix A32, Figure A32.2). Under these conditions, four TOP2A targeted siRNA's were

reverse transfected into HCC1806. Oligonucleotides 3 and 4 showed knockdown of TOP2A, and these were taken forward for further investigation (Appendix A32, Figure A32.3).

To determine if loss of TOP2A expression can induce resistance to doxorubicin in these cell lines, a siRNA knockdown and dose response was performed. TOP2A knockdown was performed in 96 well plates with HCC1806, as per the optimal conditions previously established, and treated with doxorubicin, in a dose response manner, and incubated for an additional 72 hours before cell viability was assessed with an SRB assay (Figure 5.19A-B). Simultaneously, HCC1806 cells were plated into 6-well plates, without drug treatment, in order to observe knockdown at several time point throughout the assay (Figure 5.19C).

NT transfection was not observed to reduce TOP2A levels across all time points, and both the mock and NT demonstrated consistent TOP2A levels compared to HC1806 cells used as a positive control (Figure 5.19C). Successful knockdown of TOP2A was observed with transfection of 3 and 4 oligonucleotides across the full 96-hour period, demonstrating that knockdown occurred during the 72-hour drug incubation period (Figure 5.19C). Knockdown of TOP2A, with siRNA 3 (TOP2A_3), showed an almost two-fold RF to the treatment of doxorubicin (Figure 5.19A). An unpaired t test was performed which showed a statistically significant difference between the GI_{50} values obtained across the biological repeat for both Mock and TOP2A_3 (Figure 5.19D). These data show that siRNA-mediated knockdown of TOP2A can induce a resistance phenotype in HCC1806 when treated with doxorubicin.

A

		M	NT	TOP2A_3	TOP2A_4
Doxorubicin	GI_{50} (µM)	0.039	0.053	0.07	0.064
	RF	-	1.358	1.799	1.64

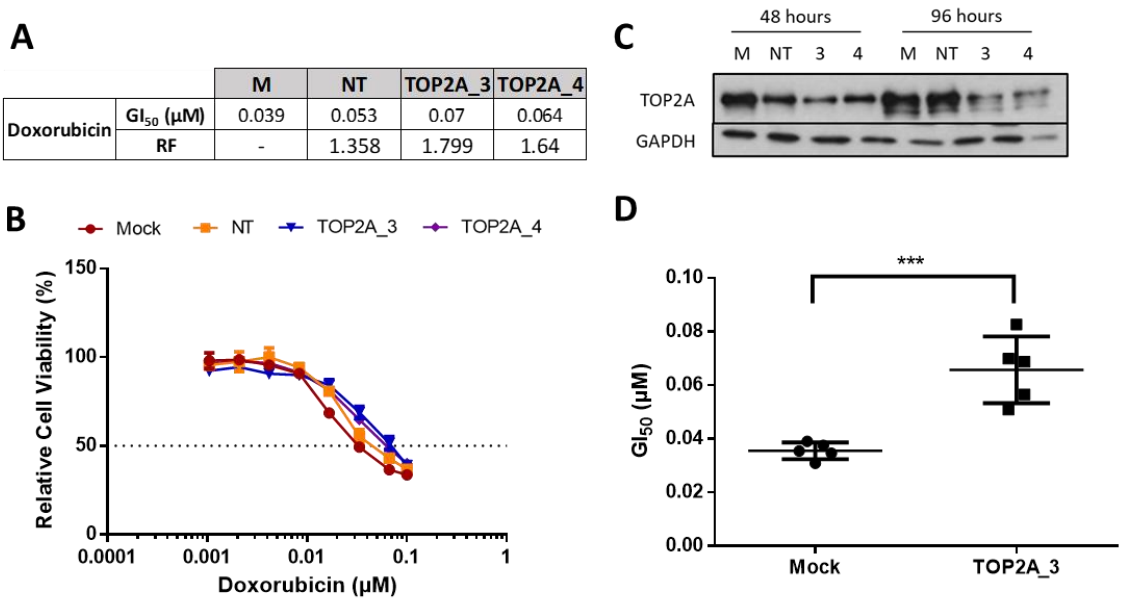


Figure 5.19 TOP2A siRNA knockdown in HCC1806 and response to doxorubicin
HCC1806 cells were reverse-transfected in 96 well plates with RNAse free water (mock; M), 25 nM non-targeting Allstar negative control oligonucleotide (NT), or 10 nM of TOP2A oligonucleotides (3; TOP2A_3, 4; TOP2A_4) using 0.05% Lipofectamine 2000. 24 hours after transfection, cells were treated with either a serial dilution of doxorubicin. A-C) HCC1806 cells were reverse-transfected as above in 6-well dishes simultaneously with 96 well plate transfection as knockdown effciency controls. Plates were incubated for 24, 48, 96 hours as indicated. Cells were used and analysed by western blotting as per Figure 4.18. GAPDH was used as a loading control. Dose response curves generated and half-maximal growth inhibition concentrations (GI_{50}) were calculated using GraphPad Prism 6 for treatment with doxorubicin. Growth curves were normalised to untreated control, for each transfection condition. Dotted line indicates the GI_{50}. A) Summary of GI_{50} values for response to doxorubicin in transfected HCC1806 cells. Data representative of ≥ 3 independent experiments. RF = resistance factor. D) unpaired t test of Mock an TOP2A_3 GI_{50} values across independent experiments performed in GraphPad Prism 6. *** indicates P ≤ 0.0001.

5.2.2.3 De novo variants in gemcitabine resistant cell lines

Across the two gemcitabine resistant cell lines, 20 *de novo* variants in this candidate approach were identified which may be driving gemcitabine resistance in the genes; *BCLAF1, ERCC6, TDG, APPL2, CSNK1D, MYOCD, PARD3, SMAD3, SMC1B, TUBB3, ABCC1, DSE1, CLOCK, DDX11, EGLN3, WDR76, CDK2AP2, CEP152, MAPK7* and *RB1* (Appendix A31). Table 5.6 shows the results of the missense variant predictors analysis.

Table 5.6. Summary of gemcitabine resistant missense variants effect predictor results

Cell line	Gene	SIFT	Polyphen	Mutational assessor			
				Functional impact	Protein bind site	DNA/RNA bind site	Small molecule bind site
HCC38^{r}GEM^{20}	BCLAF1	deleterious(0.02)	probably_damaging(0.996)	medium	-	-	-
HCC38^{r}GEM^{20}	ERCC6	deleterious(0)	possibly_damaging(0.903)	low	-	-	-
HCC38^{r}GEM^{20}	ERCC6	deleterious(0)	probably_damaging(1)	low	-	-	-
HCC38^{r}GEM^{20}	TDG	tolerated(0.12)	possibly_damaging(0.686)	low	1	1	-
HCC38^{r}GEM^{20}	APPL2	tolerated(0.23)	benign(0.022)	medium	-	-	-
HCC38^{r}GEM^{20}	CSNK1D	deleterious_low_confidence(0.01)	possibly_damaging(0.534)	medium	1	-	-
HCC38^{r}GEM^{20}	MYOCD	deleterious(0)	probably_damaging(0.912)	medium	-	-	-
HCC38^{r}GEM^{20}	PARD3	tolerated(0.11)	benign(0.084)	medium	-	-	-
HCC38^{r}GEM^{20}	SMAD3	tolerated(0.28)	benign(0.007)	low	1	-	SEP
HCC38^{r}GEM^{20}	TUBB3	tolerated_low_confidence(0.31)	benign(0.117)	neutral	-	-	G2N, POD, CN2, LOC
HCC38^{r}GEM^{20}	DSEL	tolerated(0.59)	possibly_damaging(0.672)	medium	-	-	-
HCC1806^{r}GEM^{20}	CLOCK	deleterious(0.02)	benign(0.02)	medium	-	-	-
HCC1806^{r}GEM^{20}	DDX11	deleterious(0.01)	benign(0.424)	low	-	-	-
HCC1806^{r}GEM^{20}	EGLN3	tolerated(0.08)	benign(0.109)	medium	-	-	-
HCC1806^{r}GEM^{20}	WDR76	deleterious(0)	possibly_damaging(0.76)	medium	-	-	-
HCC1806^{r}GEM^{20}	CDK2AP2	deleterious(0)	probably_damaging(0.973)	medium	-	-	-
HCC1806^{r}GEM^{20}	MAPK7	tolerated(0.74)	benign(0.01)	neutral	-	-	-
HCC1806^{r}GEM^{20}	RB1	tolerated(0.08)	probably_damaging(0.961)	medium	-	-	-

A *de novo* frameshift variant has been identified in *CEP152* (EQV372EX) in HCC1806^{r}GEM^{20} with a VAF of 0.129. CEP152 (centrosomal protein 152 kDa), is an essential protein required for centrosome duplication. In complex with CEP63, CDK5RAP2 and WDR62 at the centrosome, it recruits CDK2 for centriole duplication (Kodani *et al.*, 2015). The mechanism in which it may be linked to gemcitabine resistance is not understood.

A *de novo* missense variant has been identified in *WDR76* in (H578Y) in HCC1806^{r}GEM^{20} with a VAF of 5.6. WDR76 (WD repeat-containing protein 76) specifically binds 5-Hydroxymethylcytosine, a DNA pyrimidine nitrogen base derived from cytosine. It has also been shown to be a RAS binding protein that functions as a tumour suppressor through RAS degradation (Jeong *et al.*, 2019). All

three variant predictors identify this variant as damaging, and the variant is located between repeated parts of the protein (WD 6 and WD 7) (Table 5.6). The role of this variant in gemcitabine resistance is unclear.

HCC38^{r}GEM[20] harbours a *de novo* missense variant in *BCLAF1* (E603G) with a VAF of 0.25. BCLAF1 (Bcl-2-associated transcription factor 1), is a death-promoting transcriptional repressor, with a role in cyclin-D1/*CCND1* mRNA stability (Bracken *et al.*, 2009). The variant in BCLAF1 is predicted to be damaging by all three variant impact predictors, although is not located in any known binding sites (Table 5.6). The role of this variant in *BCLAF1* is unclear in gemcitabine resistance.

Taking this data together, the gemcitabine resistant cell lines did not have high VAF *de novo* variants that are considered to be damaging associated to the five categories investigated. This may suggest that the resistance mechanism to gemcitabine does not lie in DDR pathways or drug transport and metabolism.

5.2.2.4 De novo variants in eribulin resistant cell lines

Across the three eribulin resistant cell lines, 15 *de novo* variants in this candidate approach were identified which may be driving eribulin resistance in the genes; ***C2CD3***, ***KIF2C***, *MIS18BP1,* ***MYH9***, ***RDX***, *CX3CR1, NT5E, ERCC6, HERC2,* ***NIN***, *DSEL, WDR35, CLOCK, FANCD2* and ***CDK2AP2*** (Appendix A31). It is first important to note that the six genes highlighted in bold have been identified from the GO Cell cycle category and have a direct role in either; centriole elongation/cohesion, attachment of mitotic spindle microtubules to the kinetochore, regulation of the mitotic cell cycle or cytoskeleton organisation. Table 5.7 shows the results of the missense variant predictors analysis.

Table 5.7. Summary of eribulin resistant missense variants effect predictor results

Cell line	Gene	SIFT	Polyphen	Mutational assessor			
				Functional impact	Protein bind site	DNA/RNA bind site	Small molecule bind site
MDA-MB-468^{r}ERI^{50}	KIF2C	deleterious(0.01)	benign(0.23)	low	-	-	-
MDA-MB-468^{r}ERI^{50}	MIS18BP1	deleterious(0)	benign(0.432)	medium	-	-	-
MDA-MB-468^{r}ERI^{50}	CX3CR1	deleterious(0.03)	probably_damaging(0.956)	low	1	-	OLC,MHA,BOG
MDA-MB-468^{r}ERI^{50}	NT5E	deleterious(0.03)	benign(0.076)	low	-	-	-
HCC38^{r}ERI^{10}	ERCC6	deleterious(0)	possibly_damaging(0.903)	low	-	-	-
HCC38^{r}ERI^{10}	ERCC6	deleterious(0)	probably_damaging(1)	low	-	-	-
HCC38^{r}ERI^{10}	HERC2	tolerated(0.17)	benign(0.055)	low	-	-	-
HCC38^{r}ERI^{10}	NIN	deleterious(0)	possibly damaging(0.849)	medium	-	-	-
HCC38^{r}ERI^{10}	DSEL	tolerated(0.59)	possibly damaging(0.672)	medium	-	-	-
HCC38^{r}ERI^{10}	WDR35	tolerated(0.15)	benign(0.009)	low	-	-	-
HCC1806^{r}ERI^{50}	CLOCK	deleterious(0.02)	benign(0.02)	medium	-	-	-
HCC1806^{r}ERI^{50}	CDK2AP2	deleterious(0)	probably_damaging(0.973)	-	-	-	-

MDA-MB-468^{r}ERI^{100} carried a stop gain mutation at Q556* in *C2CD3*, with a VAF of 0.2. C2CD3, C2 calcium-dependent domain containing 3, promotes centriole elongation (Thauvin-Robinet *et al.*, 2014). It has been found that overly long centrioles form over-active centrosomes that nucleate more microtubules and perturb chromosome segregation (Marteil *et al.*, 2018). Therefore, the LOF variant in *C2CD3* could result in a reduction of centriole elongation, and reduced nucleation of microtubules. This seems counterintuitive given the depolymerising activity of eribulin and the role of *C2CD3* in eribulin resistance is unknown.

Also identified in MDA-MB-468^{r}ERI^{100} is a *de novo* missense variant in *KIF2C* with a VAF of 0.5. KIF2C, kinesin family member 2C, when in complex with KIF18B (kinesin-like protein KIF18B), constitutes the major microtubule plus end depolymerising activity in mitotic cells (Tanenbaum *et al.*, 2011). The variant, E238Q, is found in the 207-238 region which has been identified for negative regulator of microtubule binding (by similarity in Uniprot). The variant has only been considered to be deleterious with the variant predictor SIFT, but not damaging by Polyphen or mutation assessor (Table 5.7). It could be predicted that the missense variant *KIF2C* prevent depolymerising activity, reducing the net depolymerising activity induced by eribulin.

Also found in MDA-MB-468^{r}ERI^{100} is a missense variant in *NT5E* at N333D, a site which normally undergoes N-linked (GlcNAc) asparagine glycosylation with a VAF of 0.14. NT5E (5-nucleotidase, alternative name CD73) has been shown to hydrolyse extracellular nucleotides into membrane permeable nucleosides, and have a role in adenosine biosynthetic process (Tanenbaum *et al.*, 2011). Through its role associated with adenosine, it has been shown that a knockout of NT5E/CD73

correlates with higher levels of detyrosinated microtubules in mice, which is consistent with a role for endogenous adenosine in reducing microtubule stability (Fassett *et al.*, 2009). Only SIFT predicted this variant to be damaging (Table 5.7). It could be predicted that the variant prevents glycosylation at the site N333, which may be regulating the protein. It could potentially lead to higher levels of detyrosinated microtubules, resulting in stabilised microtubules and thereby overcoming the de-stabilising effect of eribulin.

HCC38^{r}ERI^{10} harboured a missense variant in *NIN* E616Q with a VAF of 0.363. NIN (Ninein) is a centrosomal protein required for microtubule organisation, anchoring at the centrioles and microtubule nucleation. It has been shown that overexpression of NIN leads to a mislocalisation of γ-tubulin, recruiting it to ectopic sites which are not active in nucleating microtubules (Stillwell, Zhou and Joshi, 2004). All three variant predictors have identified it as damaging, however its implication in eribulin resistance cannot be predicted (Table 5.7)

Two *de novo* variants were identified in *ERCC6* (L610R, I611F) in HCC38^{r}ERI^{10}, both with a VAF of 0.2. ERCC6 (DNA excision repair protein ERCC-6) is an essential factor involved in transcription-coupled nucleotide excision repair. Both variants are predicted to be deleterious and damaging by SIFT and Polyphen respectively, although it has a predicted low functional impact when considered with mutational assessor (Table 5.7). The variants are both present in the helicase ATP-binding domain of the protein, however the implications of these variants in *ERCC6* for eribulin resistance cannot be predicted.

5.2.2.5 De novo variants in paclitaxel resistant cell lines

Across the three paclitaxel resistant cell lines, 18 *de novo* variants in this candidate approach were identified which may be driving paclitaxel resistance in the genes; ***FZR1, MLK1, SMARCAL1, THOC, INHBA, PTPN3, PTPRC, SPECC1L, TSC1****, FXR1, HERC2, HTT, IGF1R, PPP1R12,* *SIRPA, TDG,* ***CDK2AP2*** and *LRP2* (Appendix A31). The 15 genes highlighted in bold have been identified from the GO Cell cycle category and have a direct role in either; regulation or arrest of the cell cycle and regulation of DNA damage checkpoints. Table 5.8 shows the results of the missense variant predictors analysis.

Table 5.8. Summary of paclitaxel resistant missense variants effect predictor results

Cell line	Gene	SIFT	Polyphen	Mutational assessor			
				Functional impact	Protein bind site	DNA/RNA bind site	Small molecule bind site
MDA-MB-468^{r}PCL^{20}	FZR1	deleterious_low_confidence(0)	benign(0.131)	medium	-	-	-
MDA-MB-468^{r}PCL^{20}	MLH1	tolerated(0.33)	benign(0.06)	low	-	-	-
MDA-MB-468^{r}PCL^{20}	SMARCAL1	deleterious(0)	benign(0.162)	neutral	-	1	-
MDA-MB-468^{r}PCL^{20}	PTPN3	deleterious(0.02)	possibly_damaging(0.823)	medium	1	-	-
MDA-MB-468^{r}PCL^{20}	PTPRC	tolerated(0.21)	benign(0.025)	neutral	1	-	-
MDA-MB-468^{r}PCL^{20}	SPECC1L	deleterious(0)	probably_damaging(0.987)	medium	-	-	-
MDA-MB-468^{r}PCL^{20}	TSC1	tolerated(0.48)	benign(0)	neutral	-	-	-
HCC38^{r}PCL$^{2.5}$	FXR1	tolerated(0.24)	benign(0.036)	neutral	-	-	-
HCC38^{r}PCL$^{2.5}$	HERC2	tolerated(0.17)	benign(0.055)	low	-	-	-
HCC38^{r}PCL$^{2.5}$	IGF1R	deleterious(0.01)	possibly_damaging(0.629)	medium	-	-	-
HCC38^{r}PCL$^{2.5}$	PPP1R12	deleterious(0.02)	probably_damaging(0.979)	medium	-	-	-
HCC38^{r}PCL$^{2.5}$	SIRPA	deleterious(0.01)	benign(0.437)	medium	1	-	1PE,NAG
HCC38^{r}PCL$^{2.5}$	SIRPA	tolerated(0.39)	benign(0.013)	neutral	1	-	NAG
HCC38^{r}PCL$^{2.5}$	SIRPA	tolerated(0.24)	benign(0.007)	neutral	1	-	NAG
HCC38^{r}PCL$^{2.5}$	SIRPA	tolerated(0.09)	possibly_damaging(0.474)	low	1	-	NAG
HCC38^{r}PCL$^{2.5}$	SIRPA	tolerated(1)	benign(0.001)	neutral	1	-	NAG
HCC1806^{r}PCL^{20}	TDG	tolerated_low_confidence(0.31)	benign(0)	neutral	-	-	-
HCC1806^{r}PCL^{20}	CDK2AP2	deleterious(0)	probably_damaging(0.973)	medium	-	-	-
HCC1806^{r}PCL^{20}	CDK2AP2	tolerated(0.67)	probably_damaging(0.96)	neutral	-	-	-
HCC1806^{r}PCL^{20}	CDK2AP2	tolerated(0.27)	benign(0)	neutral	-	-	-
HCC1806^{r}PCL^{20}	CDK2AP2	tolerated(0.09)	possibly_damaging(0.908)	neutral	-	-	-
HCC1806^{r}PCL^{20}	LRP2	tolerated(0.35)	benign(0.025)	neutral	-	-	-

MDA-MB-468^{r}PCL^{20} harbours a missense variant in *FZR1* at R7G with a VAF of 0.3. FZR1 (fizzy and cell division cycle 20 related 1) is a substrate-specific adapter for the anaphase promoting complex/cyclosome (APC/C). FZR1 associates with APC/C in late mitosis, and activates it during anaphase and telophase (Fang, Hongtao and Kirschner, 1998). FZR1 has also been implicated in promoting NHEJ repair over RBBP8 mediated HRR through its regulation of RBBP8/CtIP turnover (Lafranchi *et al.*, 2014). The variant has been predicted to be deleterious and to have a medium functional impact by SIFT and mutational assessor respectively (Table 5.8). It could be predicted that misregulation of the of the cell cycle checkpoints through FZR1 may be implicated in paclitaxel resistance.

MDA-MB-468^{r}PCL^{20} also harbours a missense variant in *SPECC1L*. The variant, K195T, is within a coiled coil region of the protein, and has a VAF of 0.5. SPECC1L (cytospin-A), is involved in microtubule stabilisation and has shown to be a cross-linking protein which interacts with both microtubules and the actin cytoskeleton (Saadi *et al.*, 2011). The variant has been predicted to be damaging by all three of the variant predictors (Table 5.8). It could be predicted that a damaging variant in SPECC1L may lead to destabilising of the microtubules, overcoming the stabilising effect of paclitaxel.

A *de novo* stop gain variant was identified in *INHBA* (Q325*) in MDA-MB-468^{r}PCL^{20} with a VAF of 0.5. INHBA (inhibin beta A chain) is a TGF-beta superfamily member that have roles in reproduction and development (Brown *et al.*, 2000). It has been shown to be translationally regulated by TGFβ (Brown *et al.*, 2000). The TGFβ signalling pathway is a key inducer of the epithelial-to-mesenhymal transition process (EMT). Having not been implicated with paclitaxel resistance, this could be a novel biomarker for paclitaxel resistance.

A *de novo* missense variant was also identified in *PTPN3* (Q574H) in MDA-MB-468^{r}PCL^{20} with a VAF equal to 0.5. PTPN3 (Tyrosine-protein phosphatase non-receptor type 3) has been shown to have tyrosine phosphatase activity at junctions between the membrane and the cytoskeleton. All three variant predictors identified the variant as damaging, with mutation assessor also predicting the variant to be present in a protein binding site (Table 5.8). The variant was found to be in the PDZ domain, a domain that plays a key role in anchoring receptor proteins in the membrane to cytoskeletal components. Interestingly, PTPN3 has been found to be overexpressed in cell lines which demonstrate resistant to cisplatin and doxorubicin, of which MDA-MB-468^{r}PCL^{20} has shown (Chapter 3), although the variant implications in paclitaxel resistance cannot be predicted (S. Li *et al.*, 2016).

HCC38^{r}PCL$^{2.5}$ carried a missense variant in *HERC2* at S280I with a VAF equal to 0.4. HERC2 (E3 ubiquitin-protein ligase HERC2) is a E3 ubiquitin protein ligase that regulates ubiquitin dependent retention of repair proteins on damaged chromosomes (Bekker-Jensen *et al.*, 2010). HERC2 has been shown to regulate the insulin-like growth factor receptor signalling pathway with its essential role in insulin-like growth factor 1 receptor (IGF1R) ubiquitination for degradation (Osorio *et al.*, 2016). Interestingly, HCC38^{r}PCL$^{2.5}$ also harbours a missense variant in *IGF1R* at N819Y with a VAF equal to 0.3. The variant in *IGF1R* has been considered damaging by all three variant predictors. The variant is located extracellularly, and is part of the ligand binding site, fibronectin type III3 (Table 5.8). Importantly, it has been shown that inhibition of IGF1R sensitises non-small cell lung

cancer cells to Paclitaxel (Spiliotaki *et al.*, 2011). The implications of the variants present in *HERC2* and *IGFR* cannot be predicted.

HCC38^{r}PCL$^{2.5}$ carried a *de novo* inframe deletion (18DelQQQ) in *HTT* with a VAF of 1. HTT (huntingtin) plays a role in microtubule-mediated transport, and has been shown to interact with β-tubulin (Hoffner, Kahlem and Djian, 2002). Mutant HTT is implicated in Huntington's disease, and has been shown to result in microtubule destabilisation (Trushina *et al.*, 2003). HCC1806^{r}PCL^{20} had a *gained* insertion variant (44InsPPP) in HTT with a VAF of 0.4. It could be predicted that these variants in HTT may lead to destabilisation to counteract the stabilising effect of paclitaxel.

5.2.2.6 De novo variants in 5-Fluorouracil resistant cell lines

Across the 5-fluorouracil resistant cell lines, 4 *de novo* variants in this candidate approach were identified which may be driving 5-fluorouracil resistance in the genes; *ANP32B, ARG2* and *TLR5* Appendix A31). Table 5.9 shows the results of the missense variant predictors analysis.

Table 5.9. Summary of 5-Fluorouracil resistant missense variants effect predictor results

Cell line	Gene	SIFT	Polyphen	Mutational assessor			
				Functional impact	Protein bind site	DNA/RNA bind site	Small molecule bind site
HCC1806^{r}5-F^{1500}	ANP32B	tolerated(1)	benign(0)	neutral	-	-	-
HCC1806^{r}5-F^{1500}	ANP32B	deleterious(0.04)	possibly_damaging(0.451)	medium	-	-	-
HCC1806^{r}5-F^{1500}	ARG2	tolerated(0.58)	possibly_damaging(0.742)	neutral	1	-	-
HCC1806^{r}5-F^{1500}	TLR5	tolerated(0.05)	possibly_damaging(0.642)	medium	-	-	-

ANP32B (acidic leucine-rich nuclear phosphoprotein 32 family member B), is a multifunctional protein required for the progression from the G1 to the S phase of the cell cycle (Tochio *et al.*, 2010; Yang *et al.*, 2016). *ANP32B* mRNA has also be shown to be translationally upregulated in HCT-116 colorectal carcinoma cells upon treatment with 5-Fluorouracil (Bash-Imam *et al.*, 2017). HCC1806^{r}5-F^{1500} has two *de novo* missense variants in *ANP32B*, G234S and G235R with VAF of 0.3 each, with the second predicted to be damaging by all three variant predictors (Table 5.9). The variants are not predicted to be in any known functional domains, and the implication in 5-fluorouracil resistance cannot be predicted.

A *de novo* missense variant in *ASRG2* was identified in HCC1806^{r}5-F^{1500} S91L with a VAF of 0.4. ARG2 (arginase type II) catalyses the hydrolysis of l-arginine to ornithine and urea, reducing l-arginine availability. L-arginine is a vital precursor for the synthesis of polyamines, creatine, proline,

ornithine and nitric oxide (NO). NO is a key modulator of many key processes, and high concentrations can result in cell cycle arrest, mitochondrial respiration and apoptosis (Lunt and Vander Heiden, 2011; Napoli *et al.*, 2013; Ma *et al.*, 2015). Combined treatment of L-arginine and 5-Fluorouracil led to increased concentration of NO resulting in apoptosis (Jahani *et al.*, 2017). The variant has been identified to be damaging by Polyphen, and predicted to be in a protein binding site (Table 5.9). It could be predicted that the missense variant in *ARG2* as a GOF, which leads to the reduction in the availability of L-arginine, decreasing the pool of NO and ultimately preventing cell cycle arrest and apoptosis.

A *de novo* missense variant in *TLR5* was identified in HCC1806^{r}5-F^{1500} K844N with a VAF of 0.33. TLR5 (toll-like receptor 5) is a pattern recognition receptor on the cell surface which plays an important role in the activation of innate immunity (Hayashi *et al.*, 2001). It has been shown that an agonist of TLR5, entolimod, broadens the therapeutic window of 5-Fluorouracil by reducing its toxicity to normal tissues in mice (Kojouharov *et al.*, 2014). The variant has been identified to be damaging by Polyphen (Table 5.9).

5.3.2 Beneficially selected variants vs not beneficially selected in heterogenous population

Having established that very few variants, or mutated genes, are seen to be shared between the chemo-resistant cell lines made resistant to the same drug, the next analysis aimed to identify shifts in the heterogeneity of each resistant cell line population. Here variants which were either selected for, or developed, in the chemo-resistant cell line were considered to be beneficially selected for resistance, whilst variants which are not called or are at a lower VAF are considered to be not beneficially selected for resistance. Those variants which are considered to be beneficially selected include; *de novo* and *gained* variants and also variants which are seen to be shared between the chemo-naive and chemo-resistant cell line of which demonstrate a two-fold increase in VAF in the chemo-resistant cell line. Those considered to be not beneficially selected include; *not called* and *lost* variants, as well as variants which are seen to be shared between the chemo-naive and chemo-resistant cell line of which demonstrate a two-fold decrease in VAF in the chemo-resistant cell line. Table 5.10 summarises the variants allocation in the two lists.

Table 5.10. Allocation of variants beneficially selected or not beneficially selected lists

Beneficially selected	Not beneficially selected
De novo variants	*Not called* variants
Gained variants	*Lost* variants
Two-fold increase in VAF of *shared* variants	Two-fold decrease in VAF of *shared* variants

The *de novo, gained, not called* and *lost* variants and the calculated VAF had already been established in this chapter (Section 5.2.1.3). From the VAF fold shift analysis, variants were selected for the beneficially selected list when the VAF fold shift ≥ 2 and variants were selected for the not beneficially selected list when the VAF fold shift ≤ 0.5. Identification of these variants can be seen when the VAF of the chemo-naive cell line is plotted against the VAF of the chemo-resistant cell line.

When considering the MDA-MB-468 derived chemo-resistant cell lines, it can be seen that a linear relationship was found between the VAF of the variants in the MDA-MB-468 to the chemo-resistant (Figure 5.2.20). MDA-MB-468^{r}CDDP^{1000} and MDA-MB-468^{r}ERI^{50} had many variants which demonstrated both a two-fold increase and decrease. MDA-MB-468^{r}PCL^{20}, had very few variants shared with MDA-MB468, and of those that were shared a significant drop in VAF was observed. Several variants which are seen to have a VAF equal to 1 in the MDA-MB-468 cell line are seen to be < 0.5 VAF in MDA-MB-468^{r}PCL^{20}.

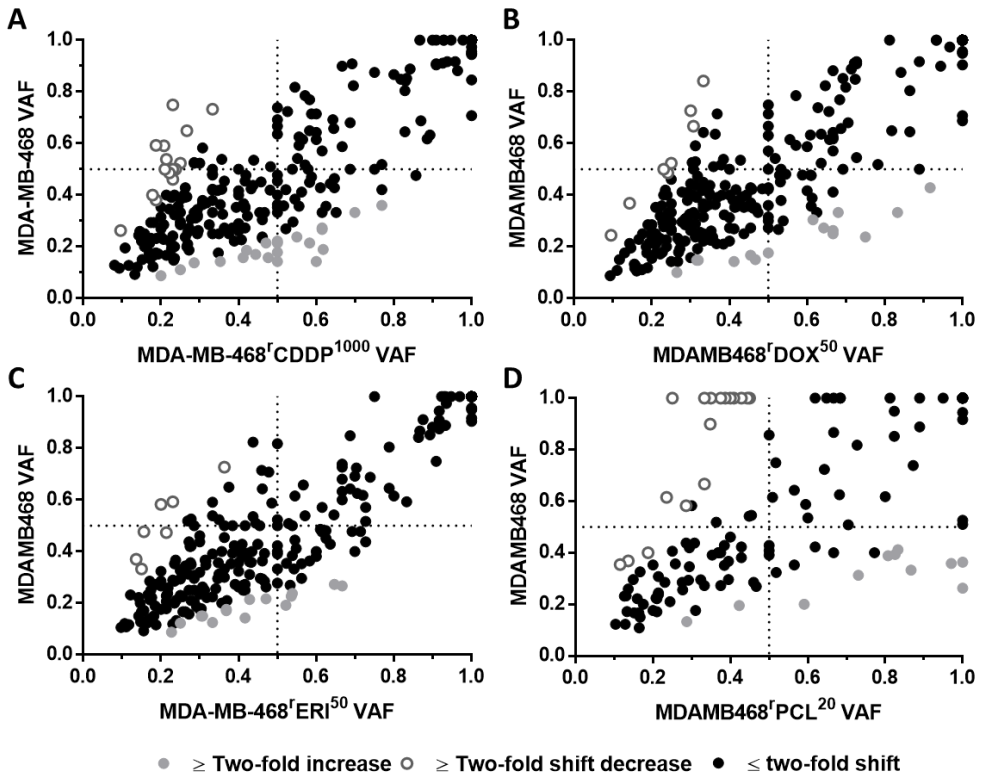


Figure 5.20 Scatter plots of variant allele frequency of shared variants between MDA-MB-468 and MDA-MB-468 derived chemo-resistant cell lines
Variant allele frequency (VAF) was calculated for each shared variant, as the frequency of alternative reads against total reads for each variant. Plots show the VAF for all variants shared between MDA-MB-468 and A) MDA-MB-468^{r}CDDP^{1000}, B) MDA-MB-468^{r}DOX^{50}, C) MDA-MB-468^{r}ERI^{50} and D) MDA-MB-468^{r}PCL^{20}. Black dots indicate ≤ two-fold shift in VAF, grey dots indicate ≥ two-fold increase in VAF in chemo-resistant cell line, open grey dots indicate ≥ two-fold decrease in VAF in chemo-resistant cell line.

A linear relationship was found between the VAF of the variants in HCC38 to the HCC38 derived chemo-resistant cell lines (Figure 5.21). Both HCC38^{r}CDDP^{3000} and HCC38^{r}DOX^{40} had very few variants shared with HCC38, but it was observed that many had a two-fold increase in VAF. HCC38^{r}GEM^{20} had the most variants shared with HCC38, and although a linear relationship was found, the majority of the shared variants had a very low VAF in both cell lines. Very few variants were observed to show a decrease in VAF in HCC38^{r}PCL$^{2.5}$.

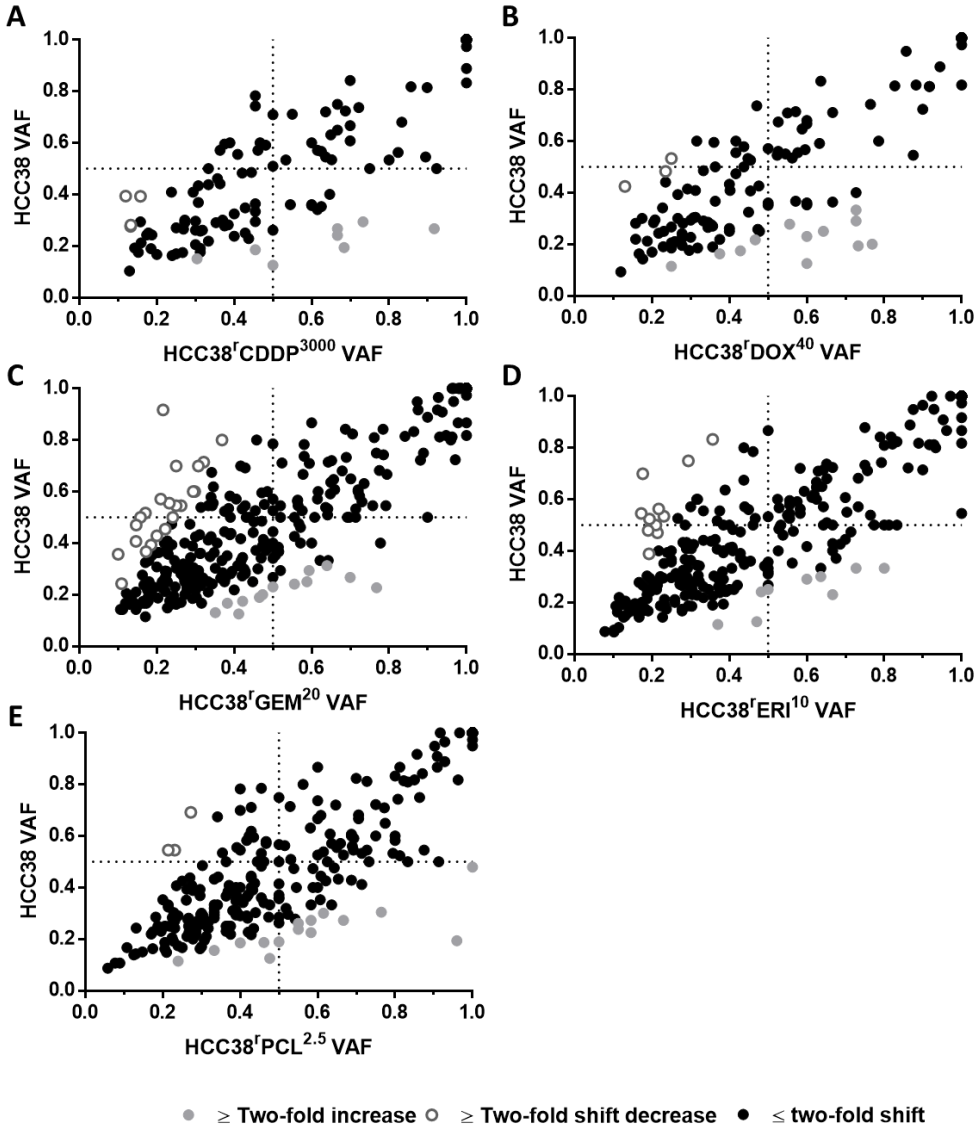


Figure 5.21 Scatter plots of variant allele frequency of shared variants between HCC38 and HCC38 derived chemo-resistant cell lines
Variant allele frequency (VAF) was calculated for each shared variant, as the frequency of alternative reads against total reads for each variant. Plots show the VAF for all variants shared between HCC38 and A) HCC38^{r}CDDP^{3000}, B) HCC38^{r}DOX^{40}, C) HCC38^{r}ERI^{10}, D) HCC38^{r}GEM^{20} and E) HCC38^{r}PCL$^{2.5}$. Black dots indicate ≤ two-fold shift in VAF, grey dots indicate ≥ two-fold increase in VAF in chemo-resistant cell line, open grey dots indicate ≥ two-fold decrease in VAF in chemo-resistant cell line.

A linear relationship was found between the VAF of the variants in HCC1806 to the HCC1806 derived chemo-resistant cell lines, although here the relationship is much tighter compared to the others in the panel (Figure 5.22). Very few variants demonstrated either a two-fold increase or decrease in VAF in HCC1806^{r}PCL^{20}, showing that very few changes are seen in the frequency of variants which are shared with HCC1806.

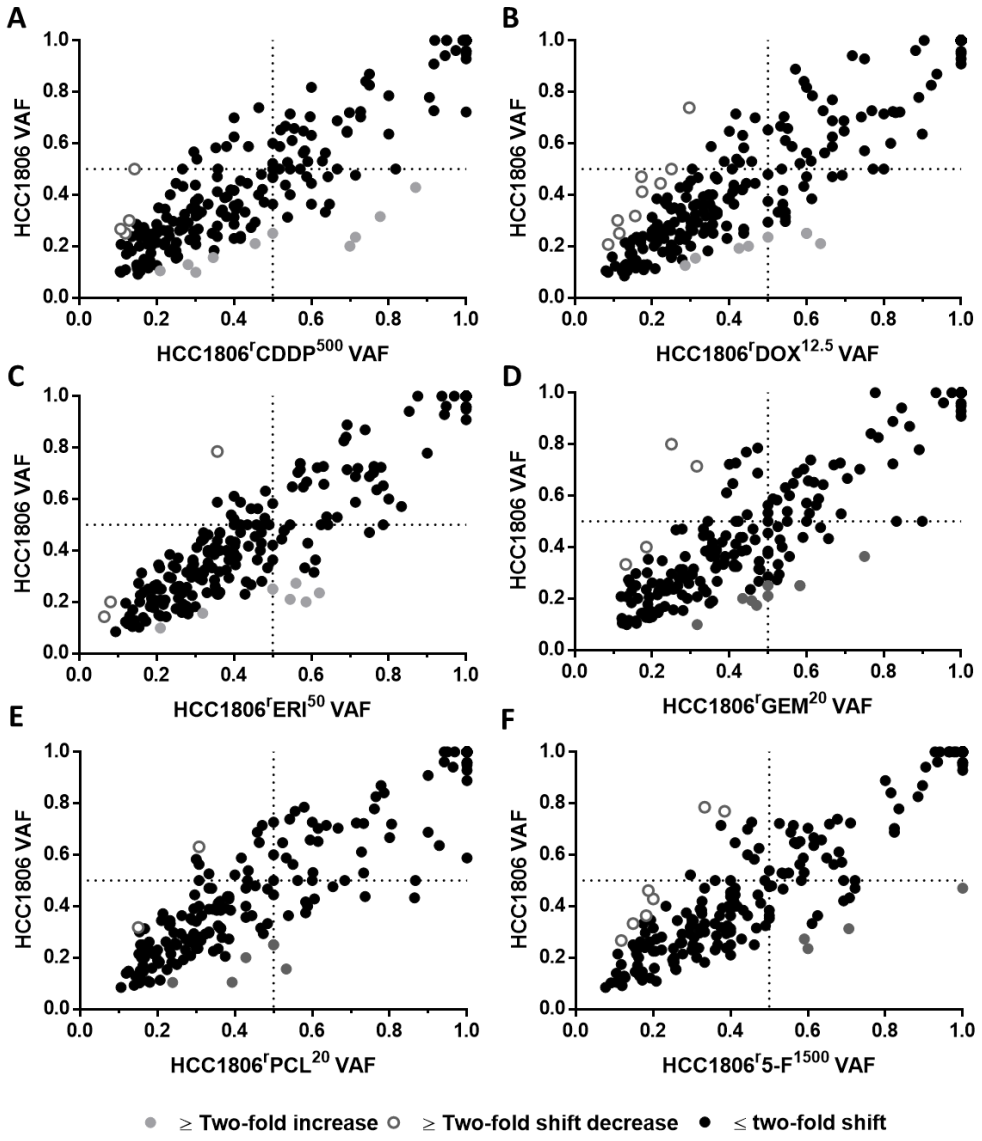


Figure 5.22 Scatter plots of variant allele frequency of shared variants between HCC1806 and HCC1806 derived chemo-resistant cell lines

Variant allele frequency (VAF) was calculated for each shared variant, as the frequency of alternative reads against total reads for each variant. Plots show the VAF for all variants shared between HCC1806 and A) HCC1806^{r}CDDP^{500}, B) HCC1806^{r}DOX$^{12.5}$, C) HCC1806^{r}ERI^{50}, D) HCC1806^{r}GEM^{20}, E) HCC1806^{r}PCL^{20} and F) HCC1806^{r}5-F^{1500}. Black dots indicate ≤ two-fold shift in VAF, grey dots indicate ≥ two-fold increase in VAF in chemo-resistant cell line, open grey dots indicate ≥ two-fold decrease in VAF in chemo-resistant cell line.

The total number of variants, and mutated genes which were determined to be beneficially selected, and not beneficially selected for each of the chemo-resistant cell lines is summarised in Figure 5.23. Mutated genes are summarised in Appendix A33, A34. Many variants were observed within the same gene. This could indicate that multiple variants are present in the gene in the cell population, or that several sub-populations of the cell population have a different variant in the same gene.

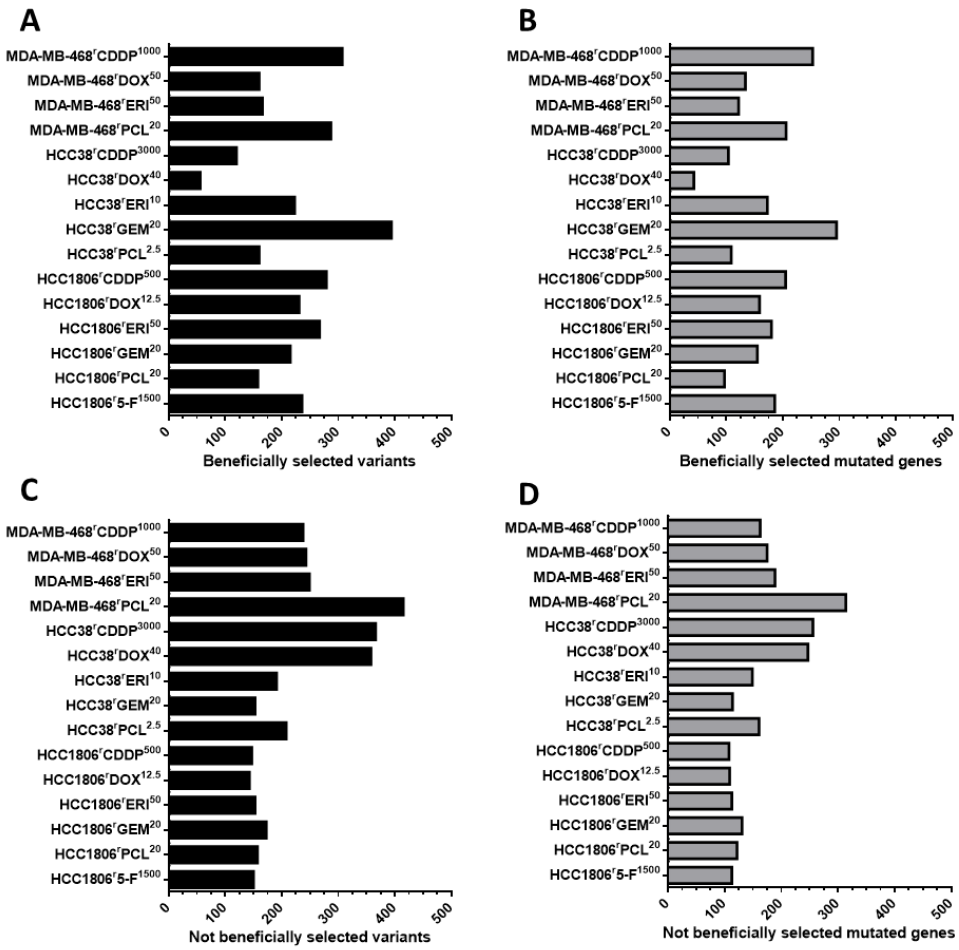


Figure 5.23 Number of variants and mutated genes in beneficially selected or not beneficially selected heterogenous populations of chemo-resistant cell lines
Analysis of changes in the heterogenous populations of the chemo-resistant cell lines in which A) variants and B) mutated genes are beneficially selected for chemo-resistance, whilst C) variants and D) mutated genes are not beneficially selected for chemo-resistance.

Having established both the beneficially selected and not beneficially selected lists, two separate GO analyses were conducted to understand the function of the mutated genes associated in the shifts of heterogeneity in the chemo-resistant cell lines. Firstly, KEGG BRITE pathway was used to label the gene lists with biological and cellular functions. KEGG BRITE is a collection of hierarchical classification systems which incorporates different types of relationships including; genes and proteins, compounds and reactions, drugs, diseases and organisms and cells. KEGG BRITE is the reference database for BRITE mapping in KEGG Mapper (Kanehisa, 2000; Kanehisa *et al.*, 2019). Here each gene list was considered to determine changes in the identified classifications between beneficially selected and not beneficially selected groups. Secondly, analysis using g:profiler was conducted, as per the description earlier in this chapter, to identify enriched GO terms (Raudvere *et al.*, 2019). Beneficially selected and not beneficially selected are termed (+) and (-) respectively.

First, the KEGG BRITE analysis of MDA-MB-468 derived chemo-resistant cell lines were considered (Figure 5.24). Analysis of MDA-MB-468^{r}CDDP^{1000} showed changes in the terms; ubiquitin systems (+12/-5), transcription factors (+11/-5), membrane trafficking (+26/-9), exosomes (+18/-9), spliceosome (+9/-16) and chromosome associated proteins (+18/-12) (Figure 5.24A). Analysis of MDA-MB-468^{r}DOX^{50} showed changes in the terms; transcription factors (+9/-7), chromosomal and associated proteins (+18/-8), messenger RNA biogenesis (+4/-8) and DNA replication (0/-4) (Figure 5.24B). Analysis of MDA-MB-468^{r}ERI^{50} showed changes in the terms; g-protein coupled receptors (+2/-10), ion channels (+5/-3), exosome (+9/-12) and transcription factors (0/-6) (Figure 5.24C). Analysis of MDA-MB-468^{r}PCL^{20} demonstrated a change in the terms; Cytochrome P450 (+4/-1), Ubiquitin system (+5/-11), Cytoskeleton proteins (+19/-13), exosome (+13/-18), transporters (+5/-10) and ion channels (+6/-5) (Figure 5.24D).

Next, using G: profiler, enriched GO terms were identified in both the beneficially selected and not beneficially gene lists for MDA-MB-468 derived chemo-resistant cell lines (Table 5.11). MDA-MB-468^{r}CDDP^{1000} demonstrated to be enriched in GO terms associated to extracellular matrix in the beneficially selected list, whilst MDA-MB-468^{r}DOX^{50} not beneficially selected gene list was enriched in GO terms associated with NADH dehydrogenase activity. It was noted that MDA-MB-468^{r}PCL^{20} had the longest enriched GO term list for both beneficially selected and not beneficially selected terms, and several of the same enriched GO terms that were seen in both gene lists.

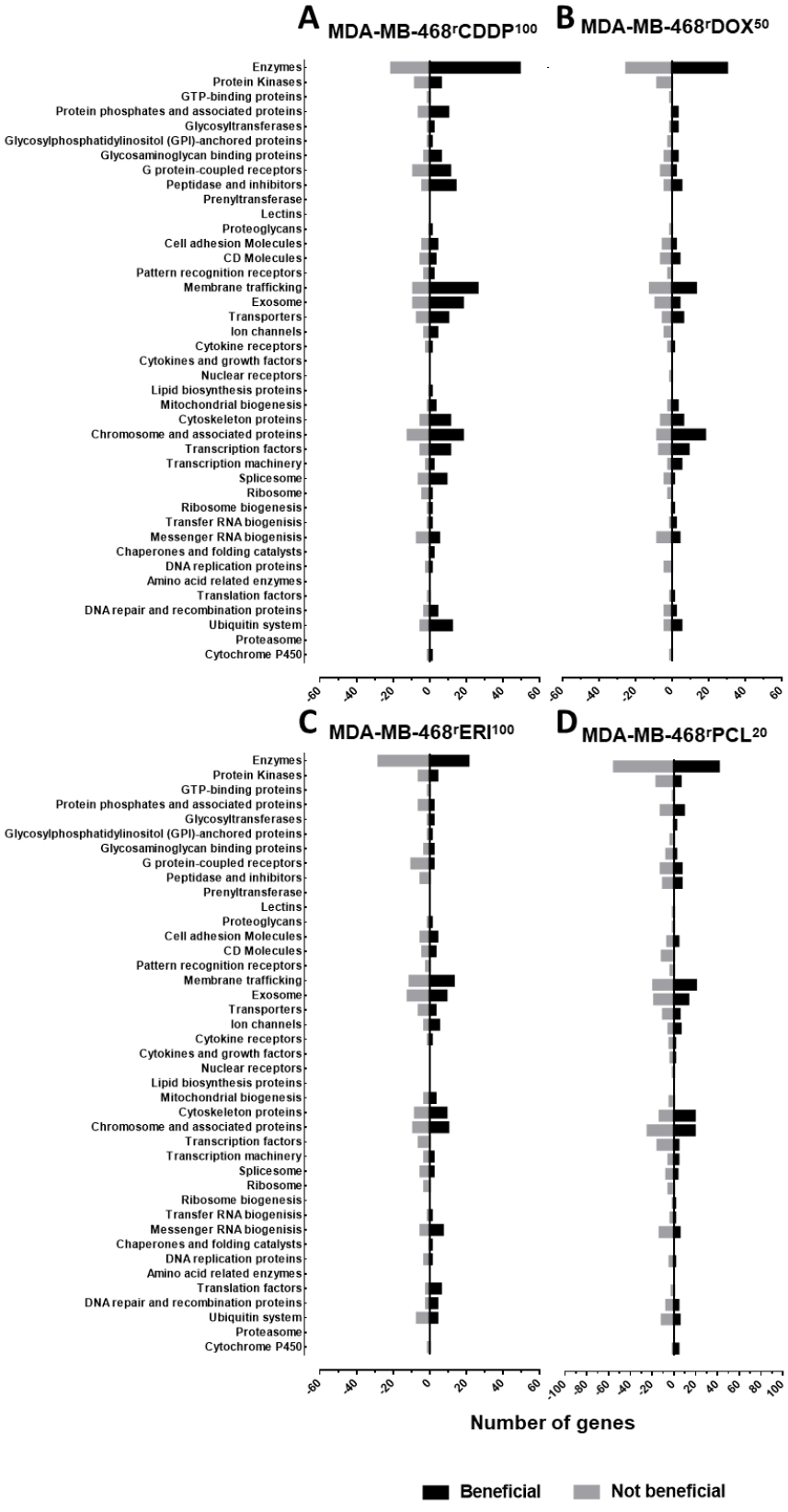


Figure 5.24 KEGG BRITE pathway analysis for MDA-MB-468 derived chemo-resistant TNBC cell lines
KEGG BRITE hierarchical classification of mutated genes considered to beneficially selected or not beneficially selected in the MDA-MB-468 chemo-resistant derived cell lines; A) MDA-MB-468^{r}CDDP^{1000}, B) MDA-MB-468^{r}DOX^{50}, C) MDA-MB-468^{r}ERI^{50} and D) MDA-MB-468^{r}PCL^{20}

Table 5.11. G:profiler analysis of beneficially selected or not beneficially selected variants in chemo-resistant MDA-MB-468 derived cell lines

	Beneficially selected	Not beneficially selected
MDA-MB-468^{r}CDDP^{1000}	extracellular matrix structural constituent (GO:0005201), extracellular matrix structural constituent conferring tensile strength (GO:0030020), extracellular matrix constituent, lubricant activity (GO:0030197), extracellular matrix (GO:0031012), collagen-containing extracellular matrix (GO:006202)3, complex of collagen trimers (GO:0098644), collagen trimer (GO:0005581), structural molecule activity (GO:0005198, O-glycan processing (GO:0016266)	O-glycan processing (GO:0016266)
MDA-MB-468^{r}DOX^{50}	extracellular matrix structural constituent (GO:0005201), extracellular matrix (GO:0031012), structural molecule activity (GO:0005198)	extracellular matrix constituent, lubricant activity (GO:0030197), O-glycan processing (GO:0016266), NADH dehydrogenase activity (GO:0003954), NADH dehydrogenase (ubiquinone) activity (GO:000813)7, NADH dehydrogenase (quinone) activity (GO:0050136), oxidoreductase activity, acting on NAD(P)H, quinone or similar compound as acceptor (GO:0016655), Respirasome (GO:0070469), Golgi lumen (GO:0005796).
MDA-MB-468^{r}ERI^{50}	O-glycan processing (GO:0016266), epithelium development (GO:0060429), supramolecular complex (GO:0099080), supramolecular polymer (GO:0099081), intermediate filament (GO:0005882), supramolecular fibre (GO:0099512), intermediate filament cytoskeleton (GO:0045111)	extracellular matrix constituent, lubricant activity (GO:0030197), O-glycan processing (GO:0016266)
MDA-MB-468^{r}PCL^{20}	O-glycan processing (GO:0016266), supramolecular polymer (GO:0099081), supramolecular fibre (GO:0099512), supramolecular complex (GO:0099080), polymeric cytoskeletal fibre (GO:0099513), cellular component organisation (GO:0016403), muscle myosin complex (GO:0005859), intermediate filament (GO:0005882), myosin II complex (GO:0016460), intermediate filament cytoskeleton (GO:0045111), Golgi lumen (GO:0005796), protein-containing complex binding (GO:004487)7, calmodulin binding (GO:0005516), organelle organization GO:0006996, isotype switching to IgG isotypes (GO:0048291), stimulatory C-type lectin receptor signalling pathway (GO:0002223), innate immune response activating cell surface receptor signalling pathway (GO:0002220).	O-glycan processing (GO:0016266), extracellular matrix constituent, lubricant activity (GO:0030197) structural molecule activity (GO:0005198), supramolecular complex (GO:0099080), respiratory chain complex (GO:0098803), respirasome (GO:0070469), protein-protein complex (GO:0032991), electron transport coupled proton transport (GO:0015990), energy coupled proton transmembrane transport, against electrochemical gradient (GO:0015988), protein tyrosine kinase activity (GO:0004713), transmembrane receptor protein kinase activity (GO:0019199), transmembrane receptor protein tyrosine kinase activity (GO:0004714).

Next, the KEGG BRITE analysis of HCC38 derived chemo-resistant cell lines were considered (Figure 5.25). Analysis of HCC38^{r}CDDP^{3000} demonstrated changes in the terms; exosome (+12/-11), chromosome and associated proteins (+3/-17), messenger RNA biogenesis (0/-6), DNA repair and recombination proteins (+2/-7), ubiquitin system (+4/-9) (Figure 5.25A). Analysis of HCC38^{r}DOX^{40} demonstrated a change in the terms; ubiquitin system (+1/-7), DNA repair and recombination proteins (0/-9), messenger RNA biogenesis (0/-8), membrane trafficking (+4/-26), transcription factors (+2/-15) (Figure 5.25B). Analysis of HCC38^{r}ERI^{10} demonstrated a change in the terms; cytoskeleton proteins (+11/-7), proteasome (+1/-1), cytochrome P450 (+2/0), ion channels (+5/0), mitochondrial biogenesis (+5/0), transporters (+6/-6) (Figure 5.25C). Analysis of HCC38^{r}GEM^{20} demonstrated a change in the terms; chromosomal; and associated proteins (+22/-9), transcription factors (+20/-6), DNA repair and recombination proteins (+8/-6), ubiquitin system (+6/-3), ion channels (+8/-10), glycosaminoglycan binding proteins (+4/0) (Figure 5.25D). Analysis of HCC38^{r}PCL$^{2.5}$ demonstrated a change in the terms; ubiquitin system, (+5/-6), DNA repair and recombination proteins (+3/-7), transcription factors (+6/-8), chromosome and associated proteins (+6/-14), cytoskeleton proteins (+6/-11) (Figure 5.25E).

Next, using G: profiler, enriched GO terms were identified in both the beneficially selected and not beneficially gene lists for HCC38 derived chemo-resistant cell lines (Table 5.12). It was noted that the GO terms found associated with the beneficially selected gene lists for HCC38^{r}CDDP^{3000} were large umbrella terms, and not very specific. Many of the same GO terms were identified to be enriched in both beneficially selected and not beneficially selected gene lists for HCC38^{r}ERI^{10}.

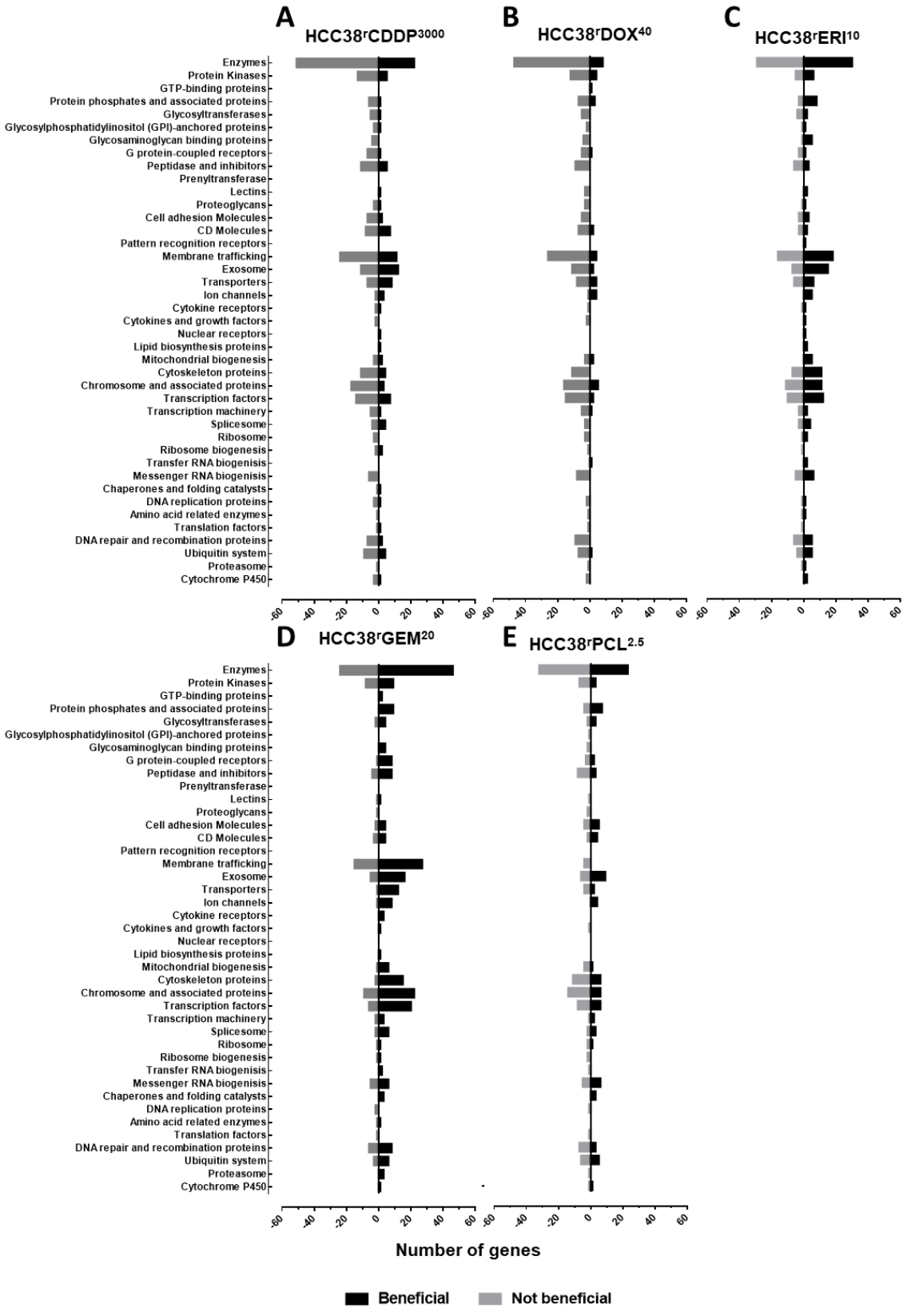


Figure 5.25 KEGG BRITE pathway analysis for HCC38 derived chemo-resistant TNBC cell lines
KEGG BRITE hierarchical classification of mutated genes considered to beneficially selected or not beneficially selected in the HCC38 chemo-resistant derived cell lines; A) HCC38^{r}CDDP^{3000}, B) HCC38^{r}DOX^{40}, C) HCC38^{r}ERI^{10}, D) HCC38^{r}GEM^{20} and E) HCC38^{r}PCL$^{2.5}$

Table 5.12. G:profiler analysis of beneficially selected or not beneficially selected variants in chemo-resistant HCC38 derived cell lines

	Beneficially selected	Not beneficially selected
HCC38^{r}CDDP^{3000}	multicellular organismal process (GO:0032501), response to external stimulus (GO:0009605), extracellular region (GO:0005576).	O-glycan processing (GO:0016266), extracellular matrix structural constituent (GO:0005201), extracellular matrix constituent, lubricant activity (GO:0030197), Golgi lumen (GO:0005796), stimulatory C-type lectin receptor signalling pathway (GO:0002223), innate immune response activating cell surface receptor signalling pathway (GO:00022220)
HCC38^{r}DOX^{40}	tissue homeostasis (GO:0001894), regulation of production of miRNAs involved in gene silencing by miRNA (GO:1903798), regulation of production of small RNA involved in gene silencing by RNA (GO:0070920)	O-glycan processing (GO:0016266), extracellular matrix structural consistent (GO:0005201), extracellular matrix constituent, lubricant activity (GO:0030197), structural consistent of synapse (GO:0098918, structural molecule activity (GO:0005198), stimulatory C-type lectin receptor signalling pathway (GO:0002223) Golgi lumen (GO:0005796), innate immune response activating cell surface receptor signalling pathway (GO:00022220), electron transport coupled proton transport (GO:0015990), energy coupled proton transmembrane transport, against electrochemical gradient (GO:0015988)
HCC38^{r}ERI^{10}	O-glycan processing (GO:0016266), stimulatory C-type lectin receptor signalling pathway (GO:0002223), innate immune response activating cell surface receptor signalling pathway (GO:00022220), Golgi lumen (GO:0005796), structural molecule activity (GO:0005198)	O-glycan processing (GO:0016266), stimulatory C-type lectin receptor signalling pathway (GO:0002223), innate immune response activating cell surface receptor signalling pathway (GO:00022220), Golgi lumen (GO:0005796), extracellular matrix constituent, lubricant activity (GO:0030197), protein O-linked glycosylation (GO:0006493)
HCC38^{r}GEM^{20}	O-glycan processing (GO:0016266), stimulatory C-type lectin receptor signalling pathway (GO:0002223), innate immune response activating cell surface receptor signalling pathway (GO:00022220), extracellular matrix constituent, lubricant activity (GO:0030197), epidermis development (GO:0008544), protein O-linked glycosylation (GO:0006493), innate immune response-activating signal transduction (GO:0002758), Golgi lumen (GO:0005796), supramolecular polymer (GO:0099081), supramolecular fibre (GO:0099512), supramolecular complex (GO:0099080), axon initial segment (GO:0043194)	O-glycan processing (GO:0016266), stimulatory C-type lectin receptor signalling pathway (GO:0002223), innate immune response activating cell surface receptor signalling pathway (GO:00022220), extracellular matrix constituent, lubricant activity (GO:0030197), transmembrane receptor tyrosine kinase activity (GO:0004714), transmembrane receptor protein kinase activity (GO:0019199, extracellular matrix structural constituent (GO:0005201) and multivesicular body, internal vesicle (GO:0097487)
HCC38^{r}PCL$^{2.5}$	O-glycan processing (GO:0016266)	O-glycan processing (GO:0016266), stimulatory C-type lectin receptor signalling pathway (GO:0002223), innate immune response activating cell surface receptor signalling pathway (GO:00022220), extracellular matrix constituent, lubricant activity (GO:0030197), extracellular matrix structural constituent (GO:0005201), Golgi lumen (GO:0005796)

Next, the KEGG BRITE analysis of HCC1806 derived chemo-resistant cell lines were considered (Figure 5.26). Analysis of HCC1806^{r}CDDP^{500} demonstrated a change in the terms; ubiquitin system (+8/-9), ion channels (+6/-2), transporters (+7/ -7), chromosome and associated proteins (+20/-11), ribosome (+3/0), exosome (+9/-5) (Figure 5.26A). Analysis of HCC1806^{r}DOX$^{12.5}$ demonstrated a change in the terms; transcription factors (+15/-5), transporters (+5/-7), messenger RNA biogenesis (+3/-7), DNA repair and recombination proteins (+4/-2), ubiquitin system (+8/-6) (Figure 5.26B). Analysis of HCC1806^{r}ERI^{50} demonstrated a change in the terms; membrane trafficking (+19/-9), cytoskeleton proteins (+5/-3), transporters (+9/-7), ubiquitin system (+6/-7) (Figure 5.26C). Analysis of HCC1806^{r}GEM^{20} demonstrated a change in the terms; transcription factors (+12/-8), ubiquitin system (+3/-8), transporters (+3/-8), chromosome and associated proteins (+14/-15) (Figure 5.26D). Analysis of HCC1806^{r}PCL^{20} demonstrated a change in the terms; membrane trafficking (+18/-12), transcription factors (+4/-9), mitochondrial biogenesis (+6/-4) (Figure 5.26E). Analysis of HCC1806^{r}5-F^{1500} demonstrated a change in the terms; ubiquitin system (+8/-7), messenger RNA biogenesis (+7/-4), exosome (+12/-5), membrane trafficking (+17/-11) (Figure 5.26F).

Finally, using G: profiler, enriched GO terms were identified in both the beneficially selected and not beneficially gene lists for HCC1806 derived chemo-resistant cell lines (Table 5.13). Of note, HCC1806^{r}GEM^{20} had very few GO terms found enriched in the beneficially selected list, but many GO terms were found enriched in the not beneficially selected list. It was also observed that very few GO terms were enriched in the gene lists derived from HCC1806^{r}5-F^{1500}, and of those identified, were often the same GO term in each list.

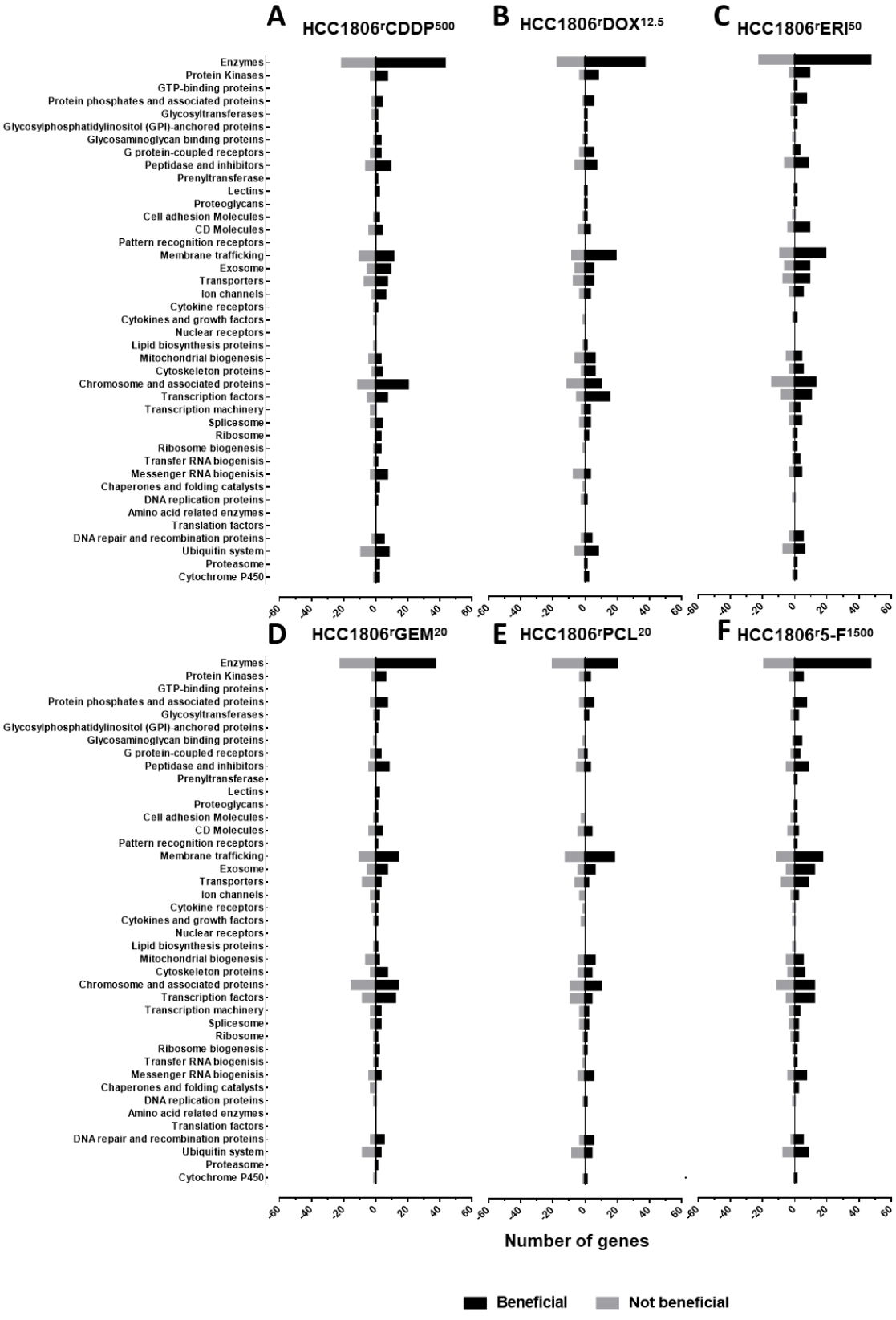


Figure 5.26 KEGG BRITE pathway analysis for HCC1806 derived chemo-resistant TNBC cell lines
KEGG BRITE hierarchical classification of mutated genes considered to beneficially selected or not beneficially selected in the HCC1806 chemo-resistant derived cell lines; A) HCC1806^{r}CDDP^{500}, B) HCC1806^{r}DOX$^{12.5}$, C) HCC1806^{r}ERI^{50}, D) HCC1806^{r}GEM^{20}, E) HCC1806^{r}PCL^{20} and F) HCC1806^{r}5-F^{1500}

Table 5.13. G:profiler analysis of beneficially selected or not beneficially selected variants in chemo-resistant HCC1806 derived cell lines

	Beneficially selected	Not beneficially selected
HCC1806^{r}CDDP^{500}	O-glycan processing (GO:0016266), Golgi lumen (GO:0005796), protein O-linked glycosylation (GO:0006493), NADH dehydrogenase (quinone) activity (GO:0050136), NADH dehydrogenase (ubiquinone) activity (GO:0008137), mitochondrial ATP synthesis coupled electron transport(GO:004277)5, ATP synthesis coupled electron transport (GO:0042773), ATP metabolic process (GO:0046034), oxidative phosphorylation (GO:0006119), electron transport chain (GO:0022900), respiratory electron transport chain (GO:0022904), respirasome (GO:0070469), respiratory chain complex (GO:0098803)	O-glycan processing (GO:0016266), Golgi lumen (GO:0005796), protein O-linked glycosylation (GO:0006493)
HCC1806^{r}DOX$^{12.5}$	enzyme binding (GO:0019899), extracellular matrix constituent, lubricant activity (GO:0030197), intermediate filament (GO:0005882), keratin filament 9GO:0045095), Golgi lumen (GO:0005796), intermediate filament cytoskeleton (GO:0045111), O-glycan processing (GO:0016266), stimulatory C-type lectin receptor signalling pathway (GO:0002223), innate immune response activating cell surface receptor signalling pathway (GO:00022220), protein O-linked glycosylation (GO:0006493), positive regulation of response to external stimulus (GO:0032103).	NADH dehydrogenase activity (GO:0003954), NADH dehydrogenase (ubiquinone) activity (GO:0008137), NADH dehydrogenase (quinone) activity (GO:0050136), oxidoreductase activity, acting on NAD(P)H, quinone or similar compound as acceptor (GO:0016655), protein C-terminus binding (GO:0008022), oxidative phosphorylation (GO:0006119), mitochondrial electron transport, NADH to ubiquinone (GO:0006120), mitochondrial ATP synthesis coupled electron transport (GO:0042775), ATP synthesis coupled electron transport (GO:0042773), respiratory chain complex (GO:0098803), mitochondrial respirasome (GO:0005746), respiratory chain complex I (GO:0045271), mitochondrial respiratory chain complex I (GO:0005747), NADH dehydrogenase complex (GO:0030964), respirasome (GO:0070469), catalytic complex (GO:1902494), oxidoreductase complex (GO:1990204), mitochondrial protein complex (GO:0098798), inner mitochondrial membrane protein complex (GO:0098800)
HCC1806^{r}ERI^{50}	extracellular matrix structural constituent (GO:0005201), O-glycan processing (GO:0016266), negative regulation of intracellular steroid hormone receptor signalling pathway (GO:0033144), protein tyrosine kinase activity (GO:0004713)	O-glycan processing (GO:0016266), electron transport coupled proton transport (GO:00015990), energy coupled proton transmembrane transport, against electrochemical gradient (GO:0015988), oxidative phosphorylation (GO:0006119), mitochondrial ATP synthesis coupled electron transport (GO:0042775), ATP synthesis coupled electron transport (GO:0042773), respiratory chain complex GO:0098803, catalytic complex (GO:1902494), respirasome (GO:0070469), mitochondrial respirasome (GO:0005746)
HCC1806^{r}GEM^{20}	O-glycan processing (GO:0016266) and protein phosphatase binding (GO:0019903)	NADH dehydrogenase activity (GO:0003954), NADH dehydrogenase (ubiquinone) activity (GO:0008137), NADH dehydrogenase (quinone) activity (GO:0050136), oxidoreductase activity, acting on NAD(P)H, quinone or similar compound as acceptor (GO:0016655), oxidative phosphorylation (GO:0006119), mitochondrial ATP synthesis coupled electron transport (GO:0042775), ATP synthesis coupled electron transport (GO:0042773), respiratory chain complex (GO:0098803), extracellular matrix constituent, lubricant activity (GO:0030197), O-glycan processing (GO:0016266), ATP metabolic process (GO:0046034), electron transport coupled proton transport (GO:0015990), energy coupled proton transmembrane transport, against electrochemical gradient (GO:0015988), electron transport chain (GO:0022900), cellular respiration (GO:0045333), protein O-linked glycosylation (GO:0006493), energy derivation by oxidation of organic compounds (GO:0015980), mitochondrial electron transport, NADH to ubiquinone (GO:0006120), respiratory chain complex (GO:0098803), respirasome (GO:0070469), mitochondrial respirasome (GO:0005746), inner mitochondrial membrane protein complex (GO:0098800), catalytic complex (GO:1902494), mitochondrial respiratory chain complex I (GO:0005747), NADH dehydrogenase complex (GO:0030964), respiratory chain complex I (GO:0045271), Golgi lumen (GO:0005796), cytochrome complex (GO:0070069), and mitochondrial membrane (GO:0031966).
HCC1806^{r}PCL^{20}	O-glycan processing (GO:0016266), protein O-linked glycosylation (GO:0006493), stimulatory C-type lectin receptor signalling pathway (GO:0002223), innate immune response activating cell surface receptor signalling pathway (GO:00022220),Golgi lumen (GO:0005796), keratin filament (GO:0045095), respirasome (GO:0070469)	O-glycan processing (GO:0016266)
HCC1806^{r}5-F^{1500}	O-glycan processing (GO:0016266)	O-glycan processing (GO:0016266), protein O-linked glycosylation (GO:0006493), Golgi lumen (GO:0005796)

5.3.3 Identification of candidate biomarkers of resistance using TCGA database

The TCGA is a cancer genomics program which has molecularly characterised over 20,000 primary cancers across thirty-three cancer types (Weinstein, 2013). The majority of the data is publicly available for analysis. Briefly, variant data was extracted via the GDC Data portal and the Bioconductor R package *TCGAbiolinks* was used to obtain clinical data (Colaprico *et al.*, 2016; Grossman *et al.*, 2016). Chromosomal locations of patient variants were remapped from GRCh38 to GRCh37 using the NCBI Genome Remapping service to allow for direct comparison between TCGA variants and variants called in the chemo-resistant cell lines.

To further understand how a variant may be related to patient response to the chemotherapeutic agents, analysis of gene expression when patients were treated with the chemotherapeutic agent was conducted. Gene expression and survival data was available for 772 patients treated with cisplatin, 614 patients treated with doxorubicin, 458 patients treated with gemcitabine, 903 patients treated with paclitaxel and 467 patients treated with 5-Fluorouracil. No gene expression data was available for eribulin treated patients. This analysis was conducted using pan-cancer types as there was not enough data available to consider only breast invasive carcinoma alone.

For each gene containing a PTV in a chemo-resistant cell line; survival analyses were conducted to determine the response of the patient treated with the chemotherapeutic agent for when the gene expression was high or low. Cox proportional hazards regression was used to calculate the hazard ratio for cohorts expressing high vs low expression levels of the given gene. The '*surv*_cutpoint' function of the package *survminer* in R allowed for the identification of the optimal expression cut-off point to give the lowest p-value for high vs low expression. The cut-off selected was between the 20^{th} and 80^{th} percentiles of gene expression values as previously described by Uhlen *et al.*, 2017. The calculations used overall survival as the measure of clinical outcome. Overall survival is defined as days to last medical follow up or death as was previously described by Ng *et al.*, 2016. The calculations were performed using the R *survminer* and *survival* packages. From this, Kaplan-Meier survival curves were generated using the R package *ggsurvplot*. Statistical analysis using the Wald test (or log rank (Mantel-Cox) test) was performed to obtain p-value of significance for each Kaplan-Meier graph. Hazard ratios were also calculated which refer to values for "low" (below median) expression for each given gene, with values >1 indicative of increased hazard (a reduced overall survival) and values < 1 are indicative of decreasing hazard (an increased overall-survival).

5.3.3.1 Comparison of beneficially selected variants to variants in the TCGA

The first analysis aimed to identify if selected variants are present in the TCGA database. The variants considered were those which have been determined to be beneficially selected in the chemo-resistant cell lines earlier in this chapter (section 5.3.2). Two types of variants were considered; exact variants and same consequence variants. Exact variants are variants which demonstrate the same chromosomal position and base change to those in the TCGA database. Same consequence variants are variants which demonstrated the same chromosomal position, with a different base change, but which resulted in the same variant consequence as those seen in the TCGA database, i.e., both are indicative of a missense variant. For the identified variants, expression data was obtained for the genes in relation to the chemotherapeutic agent the cell line is resistant to. These were then filtered based on statistical significance whereby $p \leq 0.05$ is considered significant, as well as those which demonstrated ≤ 0.2 false discovery rate (FDR). For the resulting variants, Kaplan-Meir graphs were produced and considered significant if $p \leq 0.05$. Exact or same consequence data along with filtering by statistical significance is summarised in Appendix A35.

It could be reasoned that low gene expression demonstrates a similarity to that of a LOF variant, whereby loss of available protein or loss of functional protein can result in a similar phenotype. It could also be reasoned that high gene expression demonstrates a similarity to that of a GOF variant, whereby gain of available protein, or a hyper-functional protein can result in a similar phenotype. Both scenarios can be a mechanism of drug resistance, and this would be indicative of a poor patient outcome upon treatment with the drug. Figure 5.27 summaries this concept.

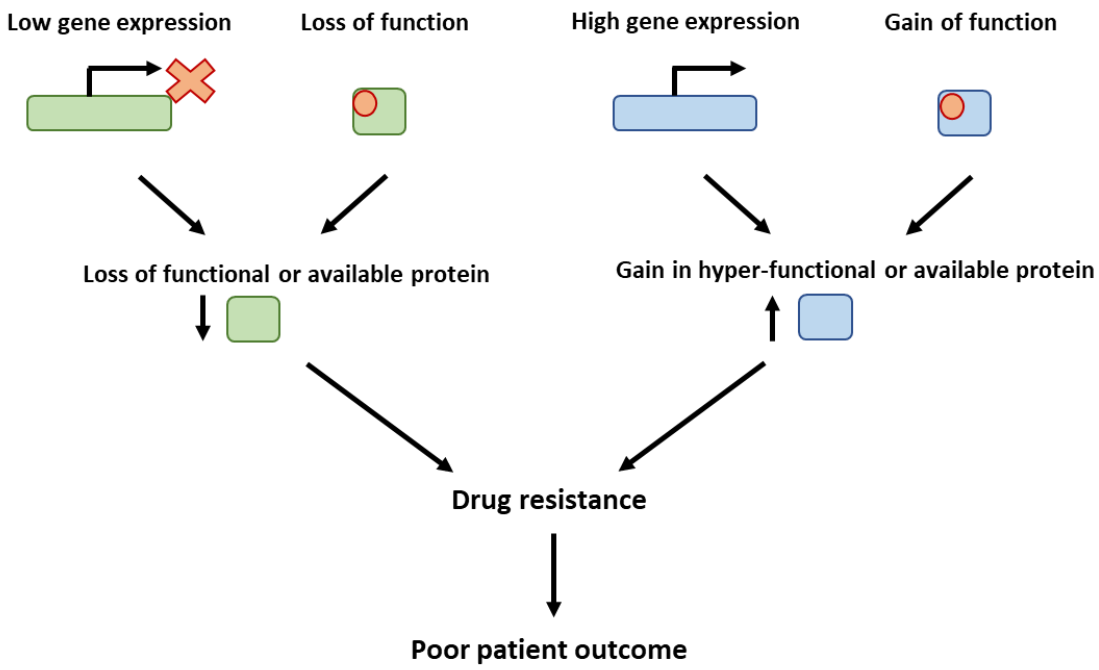


Figure 5.27 Loss of function variant or gain of function variant can lead to drug resistance in the same way as low and high gene expression
Low gene expression results in reduced availability of functional protein to carry out mechanisms of action, just as a loss of function variant in the protein is unable to carry out mechanism of action. This can result in drug resistance and poor patient outcome when treated with drug. High gene expression results in increased availability of protein to carry out mechanism of action, just as a gain of function variant in the protein could become hyper-functional. This can result in drug resistance and poor patient outcome when treated with drug. Red circle indicates variant on protein.

Analysis of the cisplatin resistant cell lines identified 22 variants; nine as exact variants, and thirteen same consequence variants. Of those 22, 17 of the genes had gene expression data available. 13 genes had a p-value of < 0.05, with 12 genes showing FDR (0.05) significance, and one gene showing FDR (0.2) significance. Kaplan Meier graphs were created for the 13 remaining genes when patients were treated with cisplatin, and eight of the genes had a p-value < 0.05 (Figure 5.28, Appendix A35). Four genes showed that low gene expression resulted in patient death when patient is treated with cisplatin; C*OL22A1, EPB41, HUWE1* and *RGS9* (Figure 5.28A-D) It could be predicted that these genes harbour a LOF variant which may lead to cisplatin resistance.

COL22A1 encodes the collagen type XXII alpha 1 chain, which has a predominant role in extracellular matrix organisation. Veskimäe *et al.*, 2018 found that mRNA from COL222A1 in platinum sensitive patients was upregulated compared to platinum resistant patients. When *COL22A1* expression is low, there is poor patient response when treated with cisplatin with a statistical significance of p = 0.013 (Figure 5.28A). It could be predicted that the missense variant in *COL22A1* in HCC1806^{r}CDDP^{500} may be a LOF variant, agreeing with the literature.

EPB41 encodes the protein 4.1, which is known as a major structural element of erythrocyte membrane skeleton. When *EPB41* expression is low, there is poor patient response when treated with cisplatin with a statistical significance of p = 0.027 (Figure 5.28B). It could be predicted that the missense variant in *EPB41* in MDA-MB-468^{r}CDDP^{1000} may be a LOF variant, although its role in cisplatin resistance is unknown.

HUWE1 encodes the E3 ubiquitin-protein ligase HUWE1, which been implicated in the regulation of CDC6 after DNA damage by marking CDC6 for proteasomal degradation and also mediates the ubiquitination of DNA polymerase β (POLB) playing a role in BER (Hall *et al.*, 2007; Parsons *et al.*, 2009). When *HUWE1* expression is low, there is poor patient response when treated with cisplatin with a statistical significance of p < 0.0001 (Figure 5.28C). It could be predicted that the missense variant in *HUWE1* in HCC1806^{r}CDDP^{500} may be a LOF variant. However, in contradiction to the data shown here, deletion of *HUWE1* in colorectal cancer results in sensitivity to cisplatin (Myant *et al.*, 2017). The role of HUWE1 in cisplatin resistance or sensitivity is unknown, but may be context or tumour type specific.

RGS9 encodes the protein regulator of G-protein signalling 9, which is known to inhibit signal transduction through the increase of GTPase activity of the G protein α subunits, thereby increasing the inactive GDP-bound form. The canonical phototransduction pathway of which RGS9 belongs to has been found to be differentially expressed in cisplatin resistant ovarian cancer (Jeyshka *et al.*, 2020). Here it can be seen that when the *RGS9* expression is low, there is poor patient response when treated with cisplatin with a statistical significance of p = 0.012 (Figure 5.28D). It could be predicted that the splice site variant in *RGS9* in HCC1806^{r}CDDP^{500} may be a LOF variant, although its role in cisplatin resistance is unknown.

Four genes showed that high expression of the gene resulted in patient death when patient is treated with cisplatin; *FKBP7, KCND2, SLC2A12* and *OGN* (Figure 5.28E-H). It could be predicted that these genes harbour a GOF variant which could result in cisplatin resistance.

FKBP7 encodes a peptidyl-prolyl cis-trans isomerase, which are known to accelerate the folding of proteins during protein synthesis. FKBP7 has been seen to be overexpressed in taxane resistant prostate and ovarian cancer, but this has not been observed with the DNA-damaging agents, such as cisplatin (N. K. Sun *et al.*, 2014; Garrido *et al.*, 2019). Interestingly none of the cisplatin resistant cell lines demonstrated cross-resistance to the taxane drug; paclitaxel (Chapter 3). Here it can be seen that when the *FKBP7* expression is high, there is poor patient response when treated with cisplatin with a statistical significance of p < 0.0001 (Figure 5.28E). It could be predicted that the

missense variant in *RGS9* in HCC1806^{r}CDDP^{500} may be a GOF variant, although its role in cisplatin resistance is unknown.

KCND2 encodes the potassium voltage-gated channel subfamily D member 2, which is a voltage-gated potassium channel (Zhu *et al.*, 1999). When *KCND2* expression is high, there is poor patient response when treated with cisplatin with a statistical significance of p = 0.00011 (Figure 5.28F). It could be predicted that the missense variant in *KCND2* in HCC1806^{r}CDDP^{500} may be a GOF variant, although its role in cisplatin resistance is unknown.

SLC2A12 encodes the solute carrier family 2, facilitated glucose transporter member, an insulin-independent facilitative glucose transporter, also known as GLUT12 (Waller *et al.*, 2013). Cisplatin has been previously reported to supress glucose uptake, and that early reduction of glucose uptake after treatment with cisplatin is a marker of cisplatin sensitivity in ovarian cancer (Egawa-Takata *et al.*, 2010). When *SLC2A12* expression is high, there is poor patient response when treated with cisplatin with a statistical significance of p = 0.00071 (Figure 5.28G). It could be predicted that the missense variant in *SLC2A12* in MDA-MB-468^{r}CDDP^{1000} may be a GOF variant, with a role consistent with the literature.

OGN encodes mimecan, which has been found to have growth factor activity inducing bone formation in conjunction with TGFβ1 or TGFβ2 (Funderburgh *et al.*, 2001). When *OGN* expression is high, there is poor patient response when treated with cisplatin with a statistical significance of p = 0.0005 (Figure 5.28H). It could be predicted that the missense variant in *OGN* in HCC38^{r}CDDP^{3000} may be a GOF variant, although its role in cisplatin resistance is unknown.

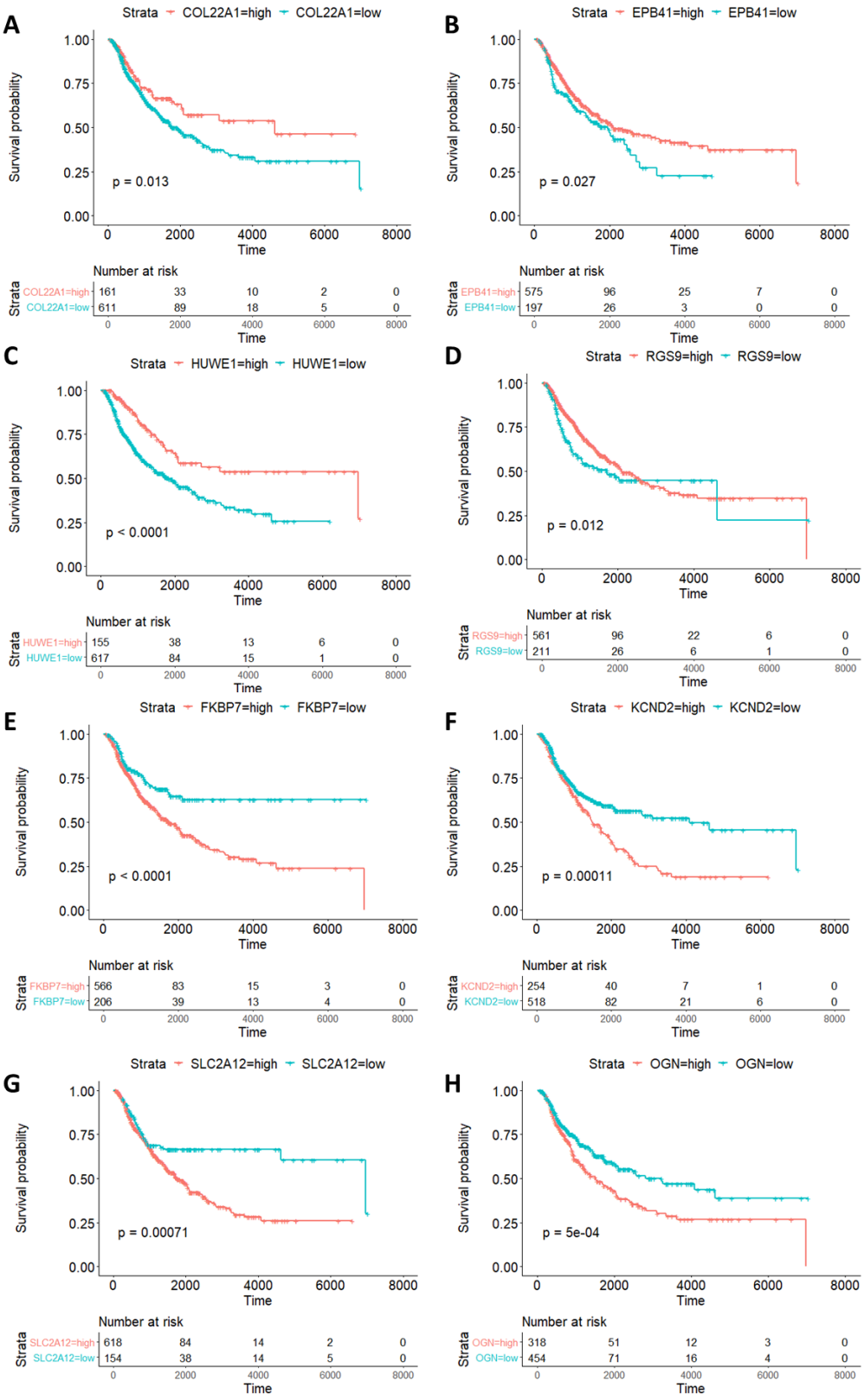


Figure 5.28 Patient survival analysis of mutated genes associated with cisplatin resistance
Variants identified as beneficially selected in the cisplatin resistant TNBC cell lines were compared to the TCGA data to determine exact or similar consequence variants. Filtered for statistical significance, $p < 0.05$ and false discovery rate < 0.2, the survival probability of the final mutated genes for when gene expression was high or low and the patient treated with cisplatin was calculated and represented in Kaplan-Meier graphs.

Analysis of the doxorubicin resistant cell lines identified nine variants; three as exact variants, and six same consequence variants. All nine variants had gene expression data available with a p-value of < 0.05 and showing an FDR (0.05) significance. Kaplan Meier graphs were created for all nine genes when patients were treated with doxorubicin, and all nine had a p-value < 0.05 (Figure 5.29, Appendix A35). Eight genes showed that low gene expression resulted in patient death when patient is treated with doxorubicin; *ADNP, C11orf80(TOP6BL), C5orf42, GXYLT1, RNF213, USH2A, NCOR1* and *TRPM7* (Figure 5.29A-H). It could be predicted that these genes harbour a LOF variant which may lead to doxorubicin resistance.

ADNP encodes the activity-dependent neuroprotector homeobox protein, which is predicted to be a potential transcription factor due to the zinc finger domain in its structure (UniProtKB, Q9H2P0). When *ADNP* expression is low, there is poor patient response when treated with doxorubicin with a statistical significance of p < 0.0001 (Figure 5.29A). It could be predicted that the missense variant in *ADNP* in HCC1806^{r}DOX$^{12.5}$ may be a LOF variant, although its role in doxorubicin resistance is unknown.

The *gene* C11orf80, also known as *TOP6BL,* encodes for the type 2 DNA topoisomerase VIB-like protein. In combination with SPO11, it mediates DNA cleavage that forms the dsDNA breaks that initiate meiotic recombination. This complex promotes relaxation of negative and positive supercoiled DNA and DNA decatenation through cleavage and ligation cycles (Robert *et al.*, 2016). When *C11orf80* expression is low, there is poor patient response when treated with doxorubicin with a statistical significance of p = 0.0047 (Figure 5.29B). It could be predicted that the inframe insertion variant in *C11orf80* in HCC1806^{r}DOX$^{12.5}$ may be a LOF variant, although its role in doxorubicin resistance is unknown.

The gene *C5orf42*, also known as *CPLANE1,* encodes for the callogenesis and planar polarity effector 1 protein, which is involved in the establishment of cell polarity required for directional cell migration (Damerla *et al.*, 2015). When *C5orf42* expression is low, there is poor patient response when treated with doxorubicin with a statistical significance of p < 0.0001 (Figure 5.29C). It could be predicted that the missense variant in *C5orf42* in HCC1806^{r}DOX$^{12.5}$ may be a LOF variant, although its role in doxorubicin resistance is unknown.

GXYLT1 encodes for glucoside xylosyltransferase 1, which is required to elongate O-linked glucose attached to EGF-like repeats in the extracellular domain of Notch proteins (Sethi *et al.*, 2010). When *GXYLT1* expression is low, there is poor patient response when treated with doxorubicin with a statistical significance of p < 0.0001 (Figure 5.29D). It could be predicted that the missense variant

in *GXYLT1* in HCC1806^{r}DOX$^{12.5}$ may be a LOF variant, although its role in doxorubicin resistance is unknown.

RNF213 encodes for an E3 ubiquitin-protein ligase involved in angiogenesis and the non-canonical Wnt signalling pathway (Liu *et al.*, 2011; Scholz *et al.*, 2016). When *RNF213* expression is low, there is poor patient response when treated with doxorubicin with a statistical significance of $p < 0.0001$ (Figure 5.29E). It could be predicted that the missense variant in *RNF213* in MDA-MB-468^{r}DOX^{50} may be a LOF variant, although its role in doxorubicin resistance is unknown.

USH2A encodes for the protein Usherin, which is known to be involved in hearing and vision as the USH2 complex (UniProtKB, O75445). When *USH2A* expression is low, there is poor patient response when treated with doxorubicin with a statistical significance of $p = 0.0026$ (Figure 5.29F). It could be predicted that the missense variant in *USH2A* in MDA-MB-468^{r}DOX^{50} may be a LOF variant, although its role in doxorubicin resistance is unknown.

NCOR1 encodes for nuclear receptor corepressor 1 protein which mediates transcriptional repression through certain nuclear receptors (Cui *et al.*, 2011). When *NCOR1* expression is low, there is poor patient response when treated with doxorubicin with a statistical significance of $p < 0.0001$ (Figure 5.29G). It could be predicted that the missense variant in *NCOR1* in HCC38^{r}DOX^{40} may be a LOF variant, although its role in doxorubicin resistance is unknown.

TRPM7 encodes for the transient receptor potential cation channel subfamily M member 7 protein, an essential ion channel known to be permeable to calcium and magnesium (Schmitz *et al.*, 2003). Interestingly Cazzaniga *et al.*, 2017 has shown that downregulation of TRPM7 results in a doxorubicin resistant phenotype in colon carcinoma. When *TRPM7* expression is low, there is poor patient response when treated with doxorubicin with a statistical significance of $p < 0.0001$ (Figure 5.29H). It could be predicted that the missense variant in *TRPM7* in HCC38^{r}DOX^{40} may be a LOF variant, which is consistent with the literature.

Three genes showed that high expression of the gene resulted in patient death when patient is treated with doxorubicin; *C20orf27, GBGT1,* and *KCND2* (Figure 5.29I-K).

C20orf27 encodes for an uncategorised protein in a predicted open reading frame. When *C20orf27* expression is high, there is poor patient response when treated with doxorubicin with a statistical significance of $p = 0.0005$ (Figure 5.29I). It could be predicted that the stop gain variant in *C20orf27* in HCC1806^{r}DOX$^{12.5}$ may be a GOF variant, although its role in doxorubicin resistance is unknown.

GBGT1 encodes the globoside α-1,3-N-acetylgalactosaminyltransferase 1 protein which is thought to catalyse the formation of glycolipid (UniProtKB, Q8N5D6). When *GBGT1* expression is high, there is poor patient response when treated with doxorubicin with a statistical significance of p = 0.034, although the final outcome is not that dissimilar (Figure 5.29J). It could be predicted that the missense variant in *GBGT1* in MDA-MB-468^{r}DOX^{50} may be a GOF variant, although its role in doxorubicin resistance is unknown.

KCND2 encodes for a voltage-gated potassium channel, and as previously seen was identified as a significant variant in HCC1806^{r}CDDP^{500}. Interestingly, Lyu *et al.*, 2006, found that topoisomerase 2β has a direct role in the expression of *KCND2.* When *KCND2* expression is high, there is also poor patient response when treated with doxorubicin with a statistical significance of p < 0.0001, although the final outcome is not that dissimilar (Figure 5.29K). It could be predicted that the missense variant in *KCND2* in HCC1806^{r}DOX$^{12.5}$ may be a GOF variant, although its role in doxorubicin resistance is unknown.

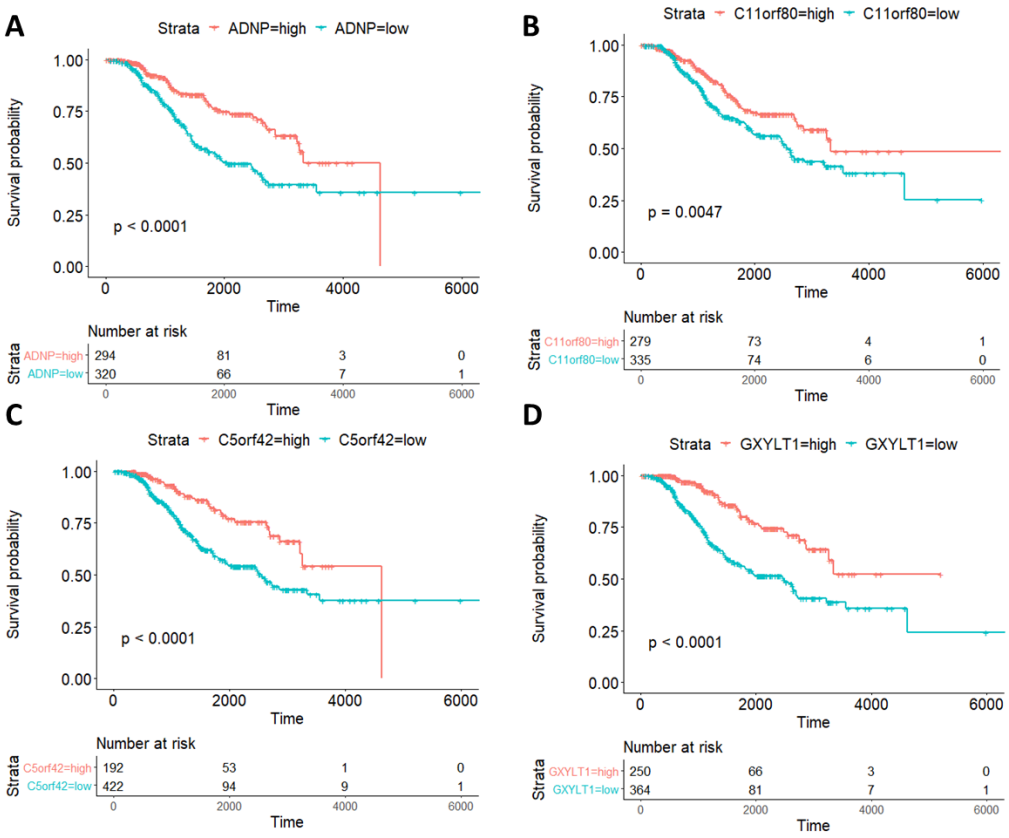


Figure 5.29 Patient survival analysis of mutated genes associated with doxorubicin resistance
Figure continued on next page

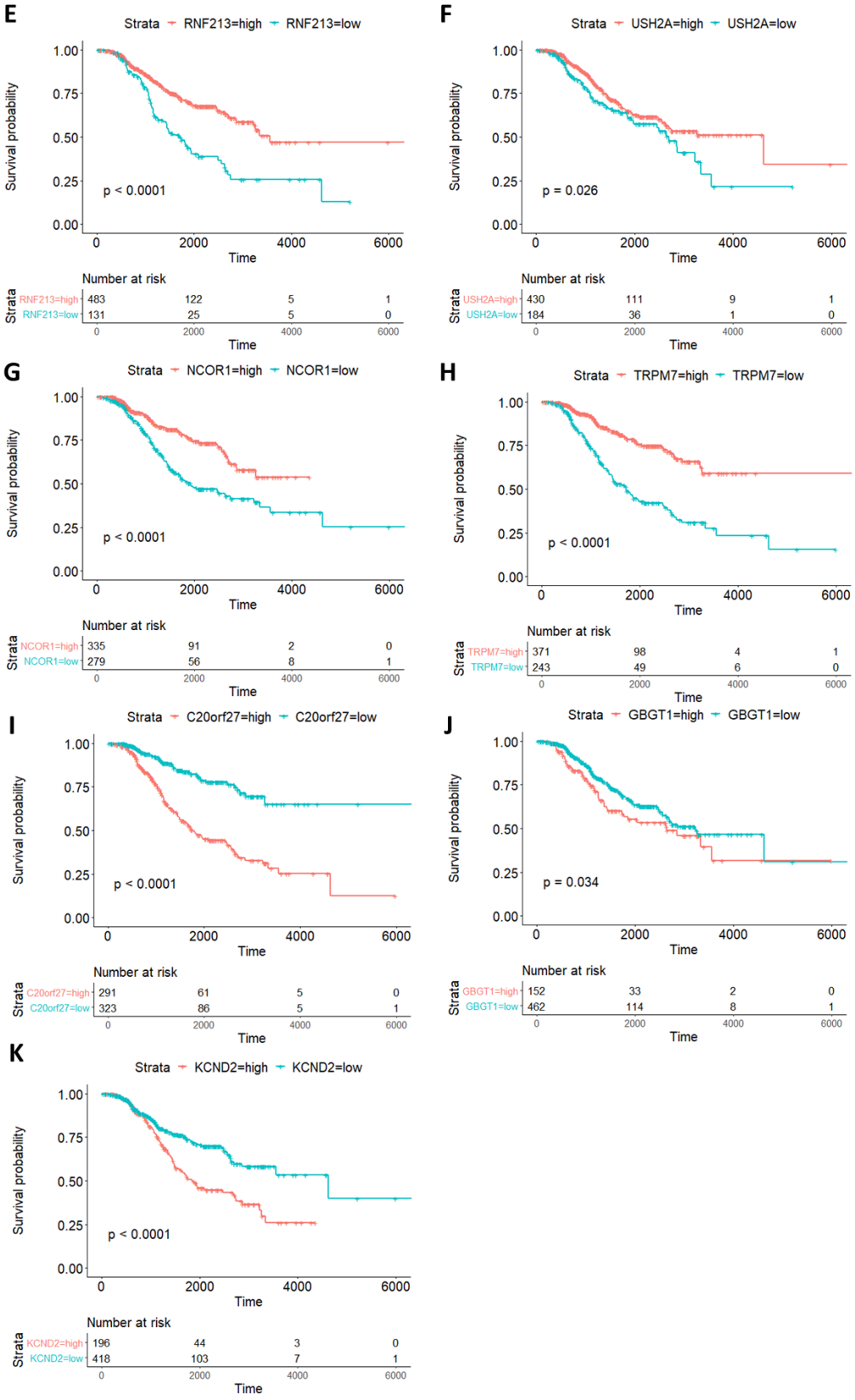


Figure 5.29 Patient survival analysis of mutated genes associated with doxorubicin resistance
Variants identified as beneficially selected in the doxorubicin resistant TNBC cell lines were compared to the TCGA data to determine exact or similar consequence variants. Filtered for statistical significance, $p < 0.05$ and false discovery rate < 0.2, the survival probability of the final mutated genes for when gene expression was high or low and the patient treated with doxorubicin was calculated and represented in Kaplan-Meier graphs.

Analysis of the gemcitabine resistant cell lines identified 16 variants; eight exact variant, and eight same consequence variants. Of the 16 variants, only 15 had gene expression data, and only five had a p-value of < 0.05 with three showing an FDR (0.005) significance and two showing an FDR (0.2) significance. Kaplan Meier graphs were created the five gene when patients were treated with gemcitabine, and all five had a p-value < 0.05 (Figure 5.30, Appendix A35). Four genes showed that low gene expression resulted in patient death when patient is treated with gemcitabine; *ADNP, ABCB10, ANK2* and *TBC1D9* (Figure 5.30A-D). It could be predicted that these genes harbour a LOF variant which may lead to gemcitabine resistance.

The gene *ADNP* was seen to have low gene expression which resulted in patient death. As previously mentioned, the gene *ADNP* is predicted to be a potential transcription factor and was shown earlier that when ADNP expression is low there is a poor outcome when patients are treated with doxorubicin. When *ADNP* expression is low, there is also poor patient response when treated with gemcitabine with a statistical significance of p = 0.022 (Figure 5.30A). It could be predicted that the missense variant in *ADNP* in HCC1806^{r}GEM^{20} may be a LOF variant, although its role in gemcitabine resistance in unknown.

The gene *ABCB10* encodes for the ATP-binding cassette sub-family B member 10 protein which has been predicted to mediate critical mitochondrial transport functions and protection of mitochondria against oxidative stress (Shintre *et al.*, 2013). When *ABCB10* expression is low, there is poor patient response when treated with gemcitabine with a statistical significance of p = 0.0094 (Figure 5.30B). It could be predicted that the missense variant in *ABCB10* in HCC38^{r}GEM^{20} may be a LOF variant, although its role in gemcitabine resistance is unknown.

The gene *ANK2* encodes for the Ankyrin-2 protein which has an essential role in the localisation and membrane stabilisation of ion transporters and ion channels (Cunha and Mohler, 2011). When *ANK2* expression is low, there is poor patient response when treated with gemcitabine with a statistical significance of p = 0.0057 (Figure 5.30C). It could be predicted that the missense variant in *ANK2* in HCC38^{r}GEM^{20} may be a LOF variant, although its role in gemcitabine resistance is unknown.

The gene *TBC1D9* encodes for the TBC1 domain family member 9 protein, which is predicted to act as a GTPase-activating protein for Rab family proteins (UniProtKB, Q6ZT07) (Nakamura *et al.*, 2015). When *TBC1D9* expression is low, there is poor patient response when treated with gemcitabine with a statistical significance of p = 0.023 (Figure 5.30D). It could be predicted that the missense

variant in *TBC1D9* in HCC38^{r}GEM[20] may be a LOF variant, although its role in gemcitabine resistance is unknown.

Only *FGF14* showed that high expression of the gene resulted in patient death when patient is treated with gemcitabine. *FGF14* encodes for the fibroblast growth factor 14, a member of the fibroblast growth factor (FGF) family, which is predicted to be involved in nervous system development and function (UniProtKB, Q92915). Although nothing is known specifically about FGF14 and gemcitabine resistance, FGF receptors (FGFR) have been implicated in chemo-resistance. Specifically, FGFR3 has been implicated in gemcitabine resistance in urothelial cancer (Y. Zhou *et al.*, 2020). When *FGF14* expression is high, there is poor patient response when treated with gemcitabine with a statistical significance of p = 0.00061, although the final patient outcome is not that different (Figure 5.30E). It could be predicted that the missense variant in *FGF14* in HCC38^{r}GEM[20] may be a GOF variant, although its role in gemcitabine resistance is unknown.

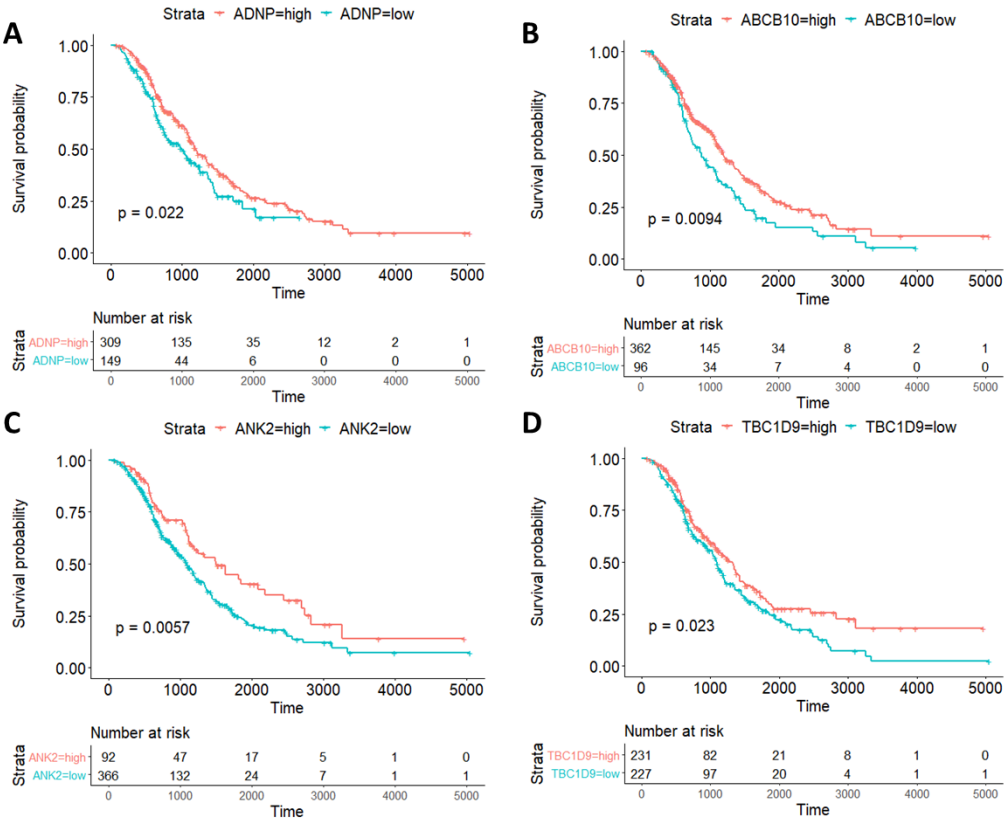


Figure 5.30 Patient survival analysis of mutated genes associated with gemcitabine resistance
Figure continued on next page

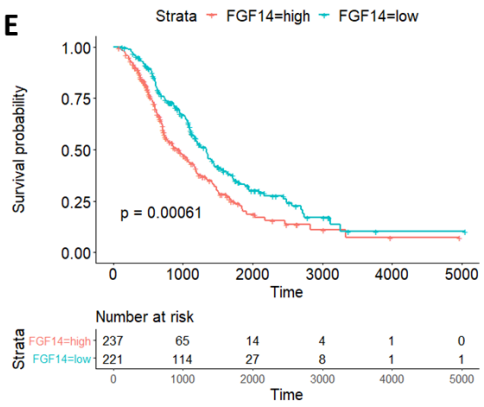


Figure 5.30 Patient survival analysis of mutated genes associated with gemcitabine resistance
Variants identified as beneficially selected in the gemcitabine resistant TNBC cell lines were compared to the TCGA data to determine exact or similar consequence variants. Filtered for statistical significance, $p < 0.05$ and false discovery rate < 0.2, the survival probability of the final mutated genes for when gene expression was high or low and the patient treated with gemcitabine was calculated and represented in Kaplan-Meier graphs.

Analysis of the paclitaxel resistant cell lines identified 12 variants; five as exact variants, and seven same consequence variants. Of those 12, 9 of the genes had gene expression data available. All nine variants had gene expression data available with a p-value of < 0.05 and showing an FDR (0.05) significance. Kaplan Meier graphs were created for the nine genes when patients were treated with paclitaxel, and eight of the genes had a p-value < 0.05 (Figure 5.31, Appendix A35). Seven genes showed that low gene expression resulted in patient death when patient is treated with paclitaxel; *ADNP, DNAJC13, FAT4, GXYLT1, MSK1 (RPS6KAS), NCOR1* and *PHFT* (Figure 5.31A-G). It could be predicted that these genes harbour a LOF variant which may lead to paclitaxel resistance.

ADNP is predicted to be a potential transcription factor and was shown earlier that when *ADNP* expression is low there is a poor outcome when patients are treated with doxorubicin and gemcitabine. Interestingly, NAP, a neuroprotective peptide, is derived from ADNP. NAP has been shown to modulate tubulin pools (Oz, Ivashko-Pachima and Gozes, 2012). When *ADNP* expression is low, there is poor patient response when treated with paclitaxel with a statistical significance of $p < 0.0001$ (Figure 5.31A). It could be predicted that the missense variant in *ADNP* in HCC1806^{r}PCL^{20} may be a LOF variant, although its role in paclitaxel resistance is unknown.

DNAJC13 encodes for DnaJ homolog subfamily C member 13 protein, also called RME-8, which has been shown to involved in membrane trafficking through early endosomes (Fujibayashi *et al.*, 2008). When *DNAJC13 e*xpression is low, there is poor patient response when treated with paclitaxel with a statistical significance of $p < 0.0001$ (Figure 5.31B). It could be predicted that the missense variant in *DNAJC13* in HCC38^{r}PCL$^{2.5}$ may be a LOF variant, although its role in paclitaxel resistance is unknown.

FAT4 encodes for protocadherin Fat 4 protein which are cell adhesion proteins. It has been found that FAT4 silenced cells show less sensitivity to paclitaxel, as we all as 5-Fluorouracil, cisplatin and oxaliplatin compared to the control cells (Ma *et al.*, 2016). HCC38^{r}PCL$^{2.5}$, which contains the FAT4 variant demonstrates cross-resistance to 5-Fluoruoracil, but not to cisplatin. When *FAT4 e*xpression is low, there is poor patient response when treated with paclitaxel with a statistical significance of p = 0.00018 (Figure 5.31C). It could be predicted that the missense variant in *FAT4* in HCC38^{r}PCL$^{2.5}$ may be a LOF variant, although its role in paclitaxel resistance is unknown.

GXYLT1 a protein which is required to elongate O-linked glucose to extracellular domain of Notch proteins and was shown earlier that when *GXYLT1* expression is low there is a poor outcome when patients are treated with doxorubicin. When *GXYLT1 e*xpression is low, there is poor patient response when treated with paclitaxel with a statistical significance of p < 0.0001 (Figure 5.31D). It could be predicted that the missense variant in *GXYLT1* in MDA-MB-468^{r}PCL^{20} may be a LOF variant, although its role in paclitaxel resistance is unknown.

MSK1, also known as *RPS6KAS,* encodes for ribosomal protein S6-inase α5, which is a serine/threonine protein kinase required for the phosphorylation and regulation of transcription factors (Ahn *et al.*, 2018). When *MSK1 e*xpression is low, there is poor patient response when treated with paclitaxel with a statistical significance of p < 0.0001 (Figure 5.31E), although the final patient outcome shows very little difference. It could be predicted that the missense variant in *MSK1* in MDA-MB-468^{r}PCL^{20} may be a LOF variant. In contradiction to the data here, Fujita *et al.*, 2010 showed that MSK1 knockdown reduced paclitaxel resistance in prostate epithelial cells. The role of MSK1 in paclitaxel resistance or sensitivity is unknown, but may be context or tumour type specific.

NCOR1 is predicted to be a potential transcription factor and was shown earlier that when *NCOR1* expression is low there is a poor outcome when patients are treated with doxorubicin. When *NCOR1 e*xpression is low, there is poor patient response when treated with paclitaxel with a statistical significance of p < 0.0001 (Figure 5.31F). It could be predicted that the missense variant in *NCOR1* in HCC1806^{r}PCL^{20} may be a LOF variant, although its role in paclitaxel resistance is unknown.

The gene *PHF2* encodes for the Lysine-specific demethylase PHF2 protein, which has a known role in demethylating both histones and non-histone proteins once activated by protein kinase A (PKA) (Baba *et al.*, 2011). It has been shown that increase of PKA induces mesenchymal-to-epithelial transition (MET) via through PHF2 which relieves H3K9me2/3 -mediated repression of epithelial

genes, and MET induced differentiation is accompanied by a loss of stem-like properties which results in a sensitivity to chemotherapeutic drugs such as doxorubicin and paclitaxel (Pattabiraman *et al.*, 2016). When *PHF2 e*xpression is low, there is poor patient response when treated with paclitaxel with a statistical significance of $p < 0.0001$ (Figure 5.31G). It could be predicted that the inframe insertion variant in *PHF2* in MDA-MB-468^{r}PCL^{20} may be a LOF variant and it was interesting to note the MDA-MB-468^{r}PCL^{20} demonstrated cross-resistance to doxorubicin. Here it could be predicted that a LOF of PHF2, prevents activation of MET induced differentiation, thereby preventing cellular sensitivity to both paclitaxel and doxorubicin.

Two genes showed that high expression of the gene resulted in patient death when patient is treated with paclitaxel; *EXT1* and *ITGB4* (Figure 5.31H-I). *EXT1* encodes for the exostosin-1 protein, which is a glycosyltransferase required for the biosynthesis of heparan-sulphate. When *EXT1 e*xpression is high, there is poor patient response when treated with paclitaxel with a statistical significance of $p < 0.0001$ (Figure 5.31H). It could be predicted that the frameshift variant in *EXT1* in HCC1806^{r}PCL^{20} may be a GOF variant, although its role in paclitaxel resistance is unknown.

ITGB4 encodes for the integrin β-4 protein, is known as a receptor for lamin, and important for IGF1 signalling (Fujita *et al.*, 2012). Interestingly, proteomic analysis conducted by Kawakami *et al.*, 2015 showed ITGB4 to be upregulated in exosomes derived from taxane-resistant prostate cancer. When *ITGB4 e*xpression is high, there is poor patient response when treated with paclitaxel with a statistical significance of $p < 0.0001$ (Figure 5.31I). It could be predicted that the frameshift variant in *ITGB4* in HCC38^{r}PCL$^{2.5}$ may be a GOF variant. Although its role in paclitaxel resistance is unknown, *ITGB4* has been previously implicated.

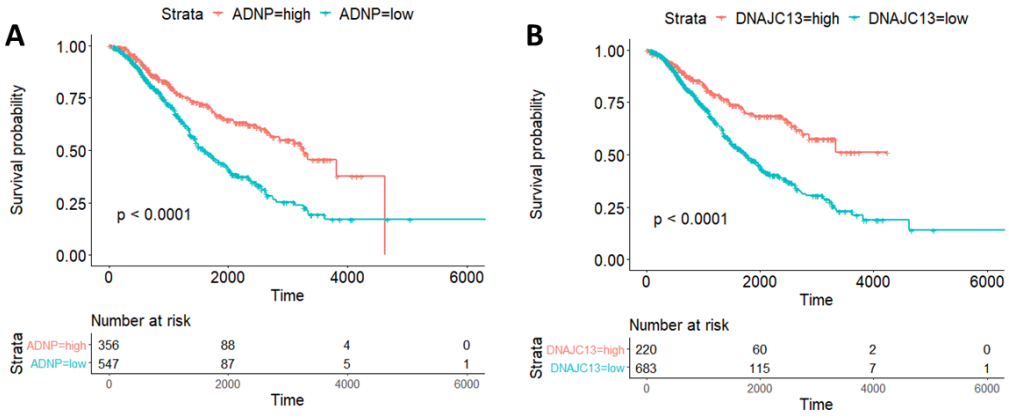


Figure 5.31 Patient survival analysis of mutated genes associated with paclitaxel resistance
Figure continued on next page

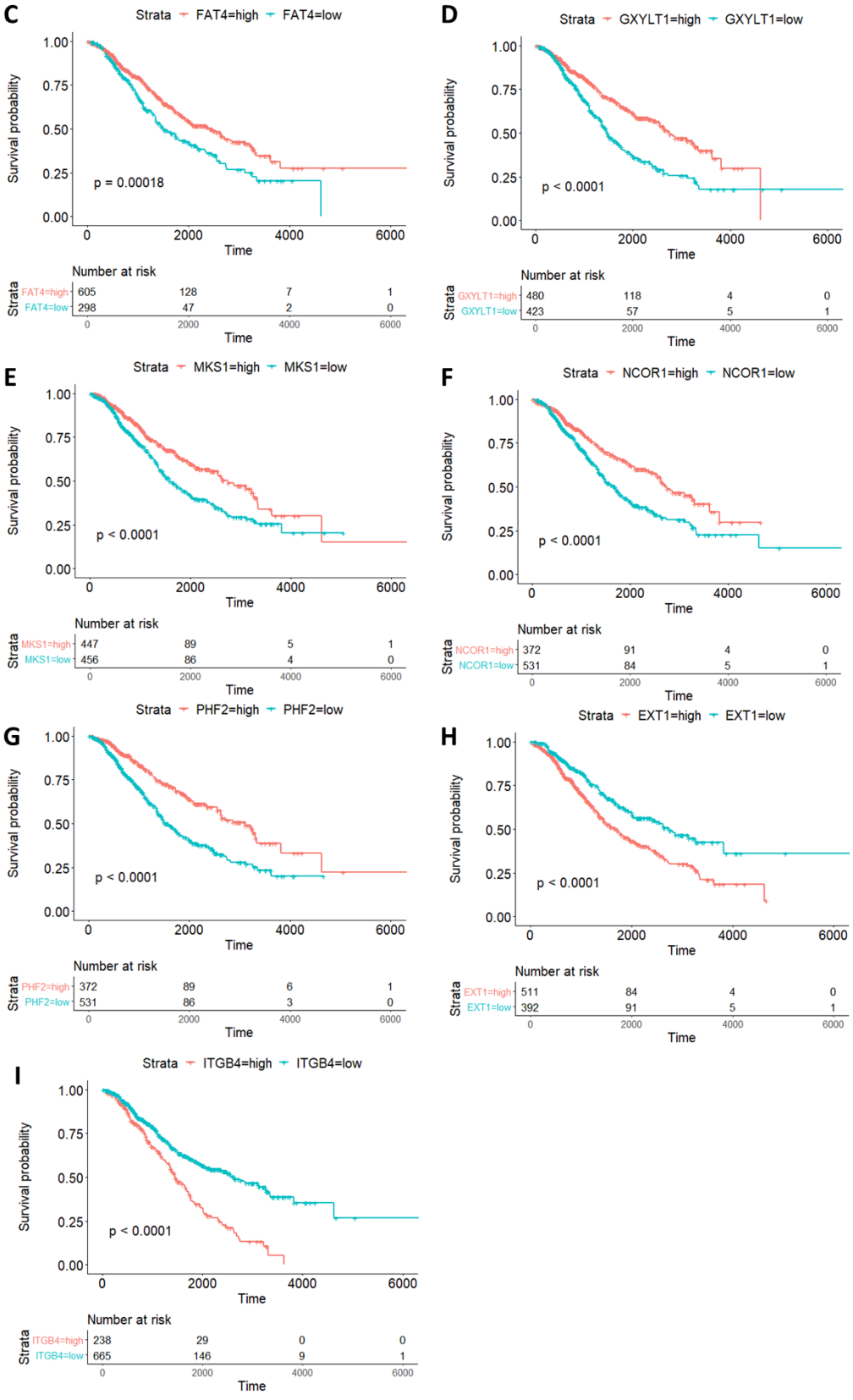


Figure 5.31 Patient survival analysis of mutated genes associated with paclitaxel resistance
Variants identified as beneficially selected in the paclitaxel resistant TNBC cell lines were compared to the TCGA data to determine exact or similar consequence variants. Filtered for statistical significance, $p < 0.05$ and false discovery rate < 0.2, the survival probability of the final mutated genes for when gene expression was high or low and the patient treated with paclitaxel was calculated and represented in Kaplan-Meier graphs.

Analysis of the 5-Fluorouracil resistant cell line, HCC1806^{r}5-F^{1500} identified eight variants; two as exact variants, and six same consequence variants. All eight of the genes had gene expression data available. Five genes had a p-value of < 0.05, with one gene showing FDR (0.05) significance, and four genes showing FDR (0.2) significance. Kaplan Meier graphs were created for the five remaining genes when patients were treated with 5-Fluorouracil, and all five of the genes had a p-value < 0.05 (Figure 5.32, Appendix A35). Two genes showed that low gene expression resulted in patient death when patient is treated with 5-Fluorouracil; *ADNP* and *C11orf80(TOP6BL)* (Figure 5.32A-B). It could be predicted that these genes harbour a LOF variant which may lead to 5-Fluorouracil resistance.

ADNP is predicted to encode for a potential transcription factor and was shown earlier that when *ADNP* expression is low there is a poor outcome when patients are treated with doxorubicin, gemcitabine and paclitaxel. When *ADNP e*xpression is low, there is poor patient response when treated with 5-Fluorouracil with a statistical significance of p = 0.00012 (Figure 5.32A). It could be predicted that the missense variant in *ADNP* in HCC1806^{r}5-F^{1500} may be a LOF variant although its role in 5-Fluorouracil resistance is unknown.

C11orf80 (*TOP6B) is known to encode for the* type 2 DNA topoisomerase VIB-like protein. Here it can be seen that when the *C11orf80 e*xpression is low, there is poor patient response when treated with 5-Fluorouracil with a statistical significance of p = 0.034 (Figure 5.32B). It could be predicted that the inframe insertion variant in *C11orf80* in HCC1806^{r}5-F^{1500} may be a LOF variant although its role in 5-Fluorouracil resistance is unknown.

Three genes showed that high expression of the gene resulted in patient death when patient is treated with 5-Fluorouracil; *CHST11, FLG* and *PIK3C2B* (Figure 5.32C-E). It could be predicted that these genes harbour a GOF variant which may lead to 5-Fluorouracil resistance. *CHST11* encodes for the carbohydrate sulfotransferase 11 protein, which catalyses the transfer of sulphate to acetylgalactosamine (GalNAc) residue of chondroitin (Okuda *et al.*, 2000). When *CHST11* expression is high, there is poor patient response when treated with 5-Fluorouracil with a statistical significance of p = 0.0012 (Figure 5.32C). It could be predicted that the missense variant in *CHST1* in HCC1806^{r}5-F^{1500} may be a gain of function, although its implication in 5-Fluorouracil resistance is unknown.

FLG encodes for the protein Filaggrin which aggregates keratin intermediate filaments (UniProtKB, P20930). When *FLG* expression is high, there is poor patient response when treated with 5-Fluorouracil with a statistical significance of p < 0.0001 (Figure 5.32D). It could be predicted that

the missense variant in *FLG* in HCC1806^{r}5-F^{1500} may be a gain of function, although its implication in 5-Fluorouracil resistance is unknown.

PIK3C2B encodes for a phosphatidylinositol 4-phosphate 3-kinase C2 domain-containing subunit β protein, which may be involved in EGF and PDGF signalling cascades (Arcaro *et al.*, 2000). When *PIK3C2B* expression is high, there is poor patient response when treated with 5-Fluorouracil with a statistical significance of p =0.029 (Figure 5.32E). It could be predicted that the missense variant in *PIK3C2B* in HCC1806^{r}5-F^{1500} may be a gain of function, although its implication in 5-Fluorouracil resistance is unknown.

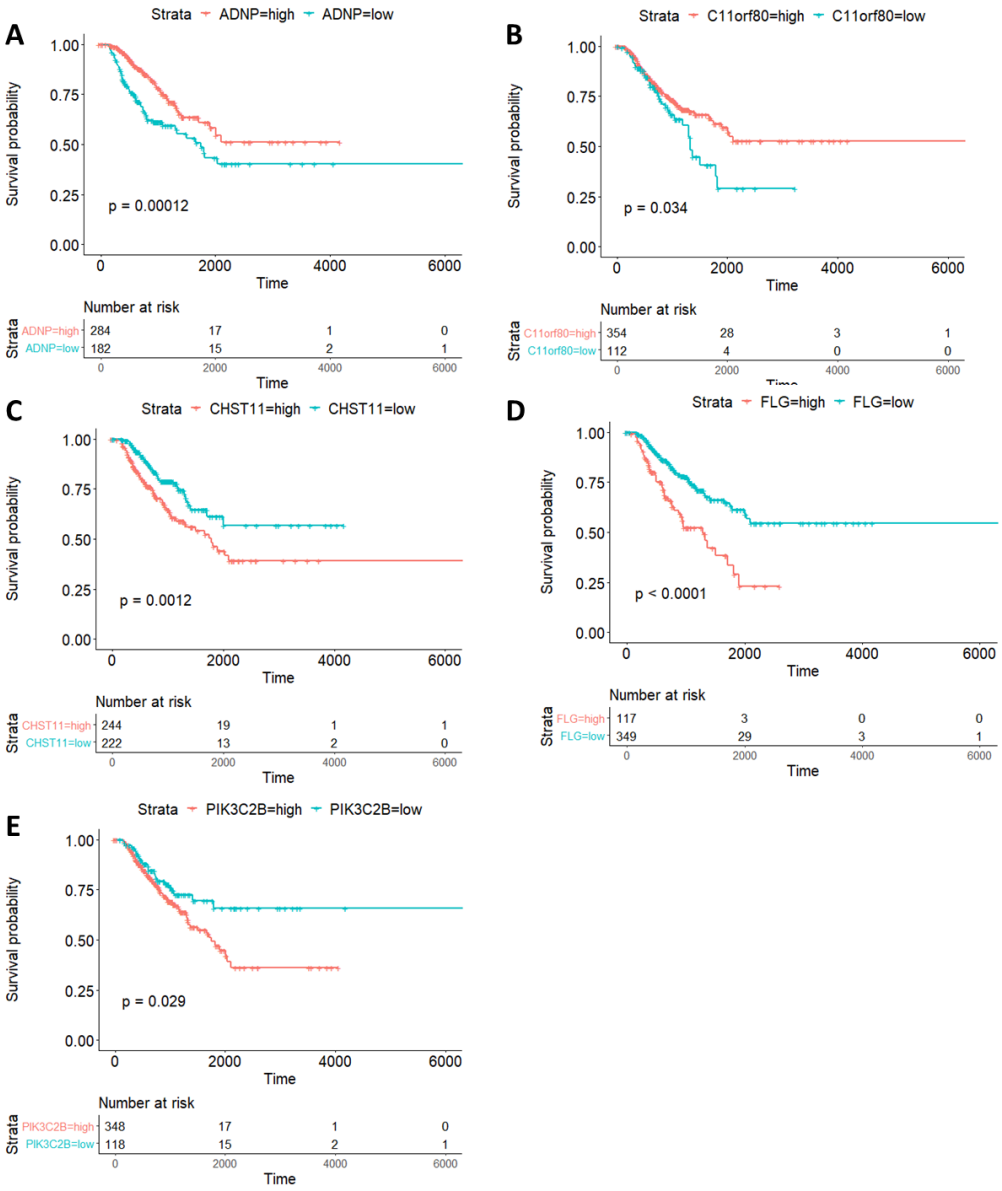


Figure 5.32 Patient survival analysis of mutated genes associated with 5-Fluorouracil resistance
Variants identified as beneficially selected in the 5-Fluorouracil resistant TNBC cell lines were compared to the TCGA data to determine exact or similar consequence variants. Filtered for statistical significance, $p < 0.05$ and false discovery rate < 0.2, the survival probability of the final mutated genes for when gene expression was high or low and the patient treated with 5-Fluorouracil was calculated and represented in Kaplan-Meier graphs.

5.3.3.2 Comparison of protein truncating variants to gene expression in the TCGA

As previously established earlier in this chapter, PTV's are the most damaging type of variants which often lead to a LOF of the protein. This next analysis aimed to consider all genes which contain PTVs in the chemo-resistant cell lines regardless of the VAF in the cell populations, and only removed those from the dataset which are seen in high confidence in the respective chemo-naive cell line. Variants were selected if they had a similar consequence to the variants in the TCGA, i.e a frameshift. Whilst these variants are not at the exact same chromosomal position, they may result in the loss of a similar length of protein. If no similar variant or consequence was seen in the TCGA dataset, the variant was removed from the list. Next the patient response to a drug when these genes were seen in high or low expression was determined using the TCGA dataset as before (section 5.3.3.1). Given that a PTV would most likely give a LOF protein, those which showed a high gene expression and high patient mortality were removed from the list. As in the previous analysis, no gene expression data is available in the TCGA for when a patient is treated with eribulin, and so the eribulin resistant cell lines cannot be considered. The final gene set was also checked to determine if they are found in the cancer gene census or are considered to be a hallmark. The full list of variants, along with the filtering by statistical significance is summarised in Appendix A36.

When considering the cisplatin resistant cell lines, 78 PTVs were identified. After filtering out PTVs found in the chemo-naive cell lines, 30 were found only in the cisplatin resistant cell lines, and 11 were found in both the cisplatin resistant cell lines and in low confidence in the chemo-naive cell lines. Of the 41 variants, 34 had available expression data for the affected gene and the patient's response when treated with cisplatin. Of the 34 variants, 29 were found to have similar consequences. When considering the expression data of the genes associated to the 29 variants, 29 had a p-value of < 0.05, with 16 showing an FDR (0.05) significance and 2 showing an FDR (0.2) significance. Kaplan-Meier graphs were created for the remaining 18 genes. Six showed that when patient had low expression of the gene, this resulted in patient death with a p-value < 0.05. None of the genes were identified to be in the cancer gene census, nor considered as a hallmark of cancer. The identified genes included; *BRD7, HSD17B3, IDO1, RGS9, KLF11* and *PRLR* (Figure 5.33A-F, Appendix A36).

The variant in *BRD7* was already identified in the MDA-MB-468^{r}CDDP^{1000} during the candidate approach analysis (section 5.2.2). When *BRD7* expression is low, there is poor patient response when treated with cisplatin with a statistical significance of $p < 0.0001$. (Figure 5.3A). This is consistent with the predicted LOF splice acceptor variant identified in *BRD7* in the cisplatin resistant cell line.

The gene *HSD17B3* encodes for the testosterone 17-β-dehydrogenase 3 protein, which allows for the reduction of androstenedione to (Engeli *et al.*, 2016). When *HSD17B3* expression is low, there is a poorer patient response when treated with cisplatin, with a statistical significance of p = 0.048 (Figure 5.33B). This is not a strong difference in patient response, and the implication of *HSD17B3* in cisplatin resistance is unknown.

The variant in *IDO1* was already identified in MDA-MB-468^{r}CDDP^{1000} during the candidate approach analysis (section 5.2.2). When *IDO1* expression is low, there is poor patient response when treated with cisplatin with a statistical significance of p < 0.0001 (Figure 5.33C). This is consistent with the predicted LOF splice acceptor variant identified in *IDO1* in the cisplatin resistant cell line.

The variant in *KLF11* was also already identified in the MDA-MB-468^{r}CDDP^{1000} during the candidate approach analysis (section 5.2.2). Here it can be seen that when the *KLF11* expression is low, there is poor patient response when treated with cisplatin with a statistical significance of p = 0.005 (Figure 5.33D). This is consistent with the predicted LOF stop gain variant identified in *KLF11* in the cisplatin resistant cell line

The final two genes which were identified in this analysis are *RGS9* and *PRLR*. The Kaplan-Meier graphs a statistical p-value equal to 0.012 and 0.011 respectively, although little difference is seen between the final patient response (Figure 5.33-E-F). The gene *RGS9* was already identified during the candidate approach analysis (section 5.2.2), and was seen to show that low expression was associated to poorer patient response upon treatment with cisplatin. *PRLR* encodes for the prolactin receptor protein, which is a receptor for the anterior pituitary hormone prolactin (Trott *et al.*, 2003). It was interesting to find that Asad *et al.*, 2019 has shown that activation of the PRLR pathway enhances proliferation, migration and cisplatin resistance in glioblastoma cells.

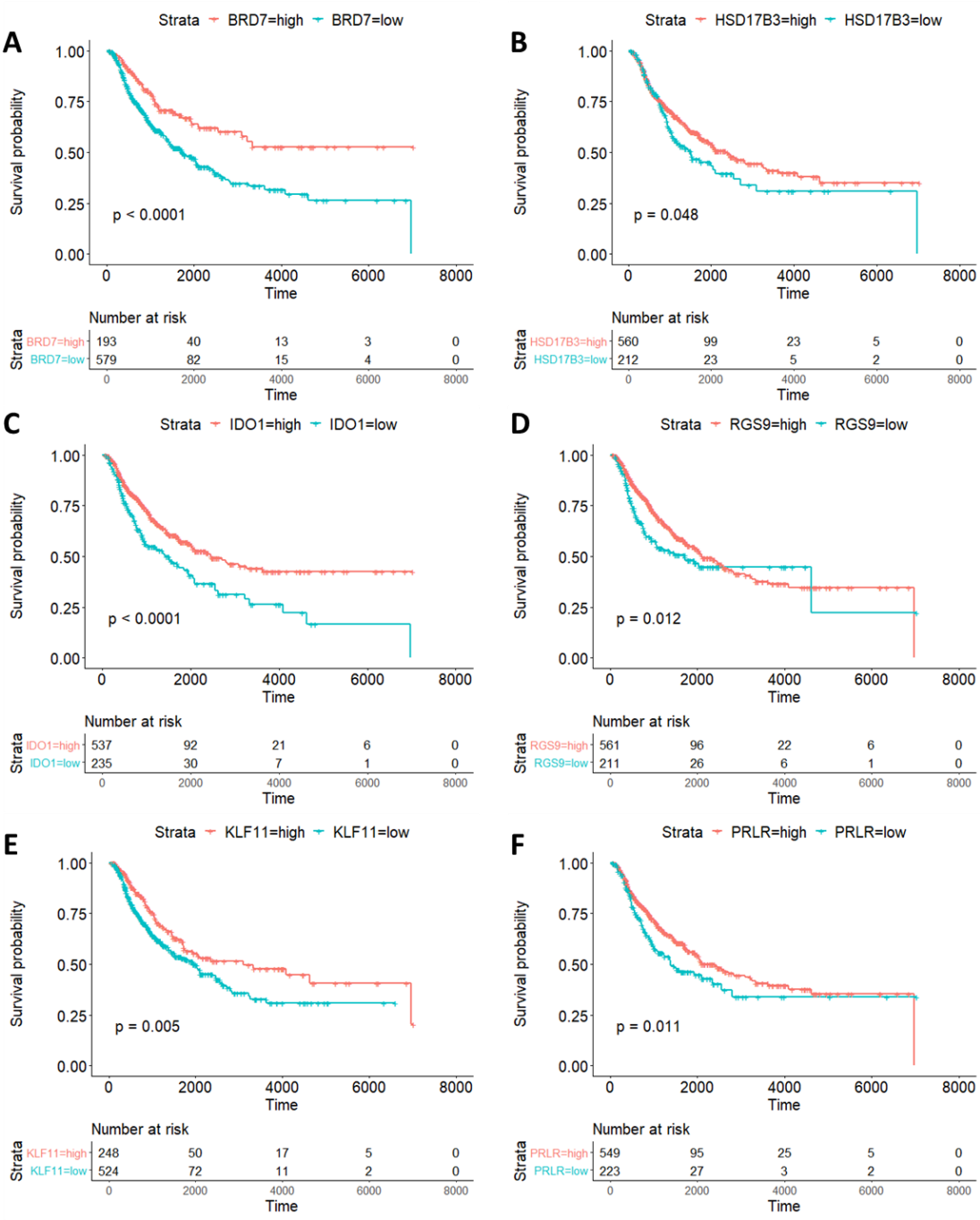


Figure 5.33 Patient survival analysis of genes containing protein truncated variants associated with cisplatin resistance
All protein truncating variants found in the cisplatin resistant TNBC cell lines were selected, with the exception of those seen in high confidence in the chemo-naive cell lines. Gene expression data for the gene was selected from the TCGA and filtered for statistical significance, $p < 0.05$ and false discovery rate < 0.2. Resulting mutated genes were selected if low gene expression showed higher patient mortality. Survival probability for the final mutated genes for when gene expression was high or low and the patient treated with cisplatin was calculated and represented in Kaplan-Meier graphs.

When considering the doxorubicin resistant cell lines, 59 PTVs were identified. After filtering out PTVs found in the chemo-naive cell lines, 16 were found only in the doxorubicin resistant cell lines, and 5 were found in both the doxorubicin resistant cell lines and in low confidence in the chemo-naive cell lines. Of the 16 variants, 14 had available expression data for the affected gene and the patient's response when treated with cisplatin. Of the 14 variants, all were found to have similar consequences. When considering the expression data of the genes associated to the 14 variants, 14 had a p-value of < 0.05, showing an FDR (0.05) significance. Kaplan-Meier graphs were created for the remaining 14 genes. Seven showed that when patient had low expression of the gene, and were treated with doxorubicin, this resulted in patient death with a p-value < 0.05. None of the genes were identified to be in the cancer gene census, nor considered as a hallmark of cancer. The identified genes included; *SLC22A23, CDON, CES2, ZKSCAN3, ABCA8, TOP2A,* and *ZNF442* (Figure 5.34, Appendix A36).

SLC22A23 encodes for the solute carrier family 22-member 23 protein, which is known for transmembrane transporter activity. In the same family, the transporter SLC22A16 has been found to mediate the uptake of doxorubicin in cancer cells, although the implication of SLC22A23 in doxorubicin resistance is unknown (Okabe *et al.*, 2005a).Here when *SLC22A23* expression is considered, there is a statistically significant p value of < 0.0001 difference in patient response when treated with doxorubicin, although very little difference between the final outcome (Figure 5.34A).

CDON encodes for the cell adhesion molecule-related/down regulated by oncogenes protein. CDON is a component of a cell-surface receptor complex that can mediate cell-cell interactions. When *CDON* expression is low, there is poor patient response when treated with doxorubicin with a statistical significance of p = 0.036. (Figure 5.34B), although its implication in doxorubicin resistance is unknown.

CES2 encodes for cocaine esterase, which is involved in the detoxification of xenobiotics as well as the activation of ester and amide prodrugs (Zhu *et al.*, 1999). Interestingly, the pro-drug, PPD, is cleaved by CES2 to obtain doxazolidine, which then rapidly hydrolyses to doxorubicin (Koch, no date; Post *et al.*, 2005; Barthel *et al.*, 2009). When *CES2* expression is low, there is poor patient response when treated with doxorubicin with a statistical significance of p = 0.0018. (Figure 5.34C), which is consistent with the LOF stop gain variant found in *CES2* in HCC1806^{r}DOX$^{12.5}$.

ZKSCAN3 encodes for the zinc finger protein with KRAB and SCAN domains 3 which is a transcriptional factor, which upon binding to a consensus sequence acts as a repressor for autophagy(Yang *et al.*, 2008). When *ZKSCAN3* expression is low, there is poor patient response

when treated with doxorubicin with a statistical significance of p = 0.0013, although the final outcome does not show much difference in patient survival (Figure 5.34D). The role of *ZKSCAN3* in doxorubicin resistance is unknown.

ABCA8 encodes for the ATP-binding cassette sub-family A member 8 protein, which is an ATP-dependent lipophilic drug transporter (Tsuruoka *et al.*, 2002). When *ABCA8* expression is low, there is poor patient response when treated with doxorubicin with a statistical significance of p = 0.0001 (Figure 5.34E). This is consistent with the hypothesis that the frameshift variant in ABCA8 in HCC1806^{r}DOX$^{12.5}$ is a LOF variant. However, this result is in contradiction to the work of Tsuruoka *et al.*, 2002, which stated that doxorubicin was not accumulated when investigating the transport capacity of ABCA8. The role of ABCA8 in doxorubicin resistance or sensitivity is unknown, but may be context or tumour type specific.

The frameshift variant found in *TOP2A* was already identified in the HCC1806^{r}DOX$^{12.5}$ during the candidate approach analysis (section 5.2.2). When *TOP2A* expression is low, there is poor patient response when treated with doxorubicin with a statistical significance of p = 0.00043. (Figure 5.34F). This is consistent with the predicted LOF frameshift variant identified in *TOP2A* in the doxorubicin resistant cell line.

ZNF442 encodes for the zinc finger protein 442, which is thought to be involved in transcriptional regulation (UniProtKB – Q9H7R0). When *ZNF442* expression is low, there is poor patient response when treated with doxorubicin with a statistical significance of p = 0.00073 (Figure 5.34G), although its implication in doxorubicin resistance is unknown.

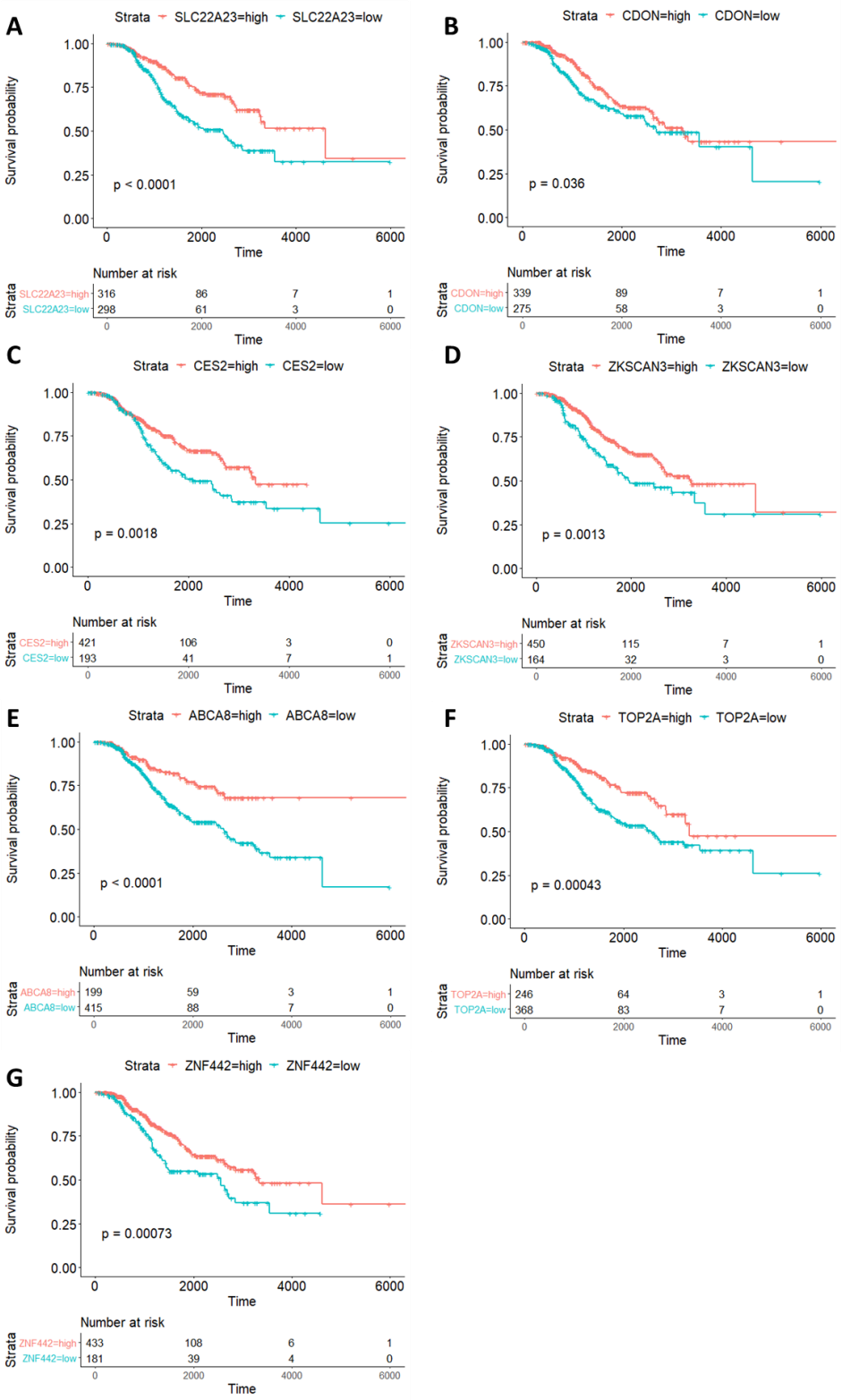


Figure 5.34 Patient survival analysis of genes containing protein truncated variants associated with doxorubicin resistance
All protein truncating variants found in the doxorubicin resistant TNBC cell lines were selected, with the exception of those seen in high confidence in the chemo-naive cell lines. Gene expression data for the gene was selected from the TCGA and filtered for statistical significance, $p < 0.05$ and false discovery rate < 0.2. Resulting mutated genes were selected if low gene expression showed higher patient mortality. Survival probability for the final mutated genes for when gene expression was high or low and the patient treated with doxorubicin was calculated and represented in Kaplan-Meier graphs.

When considering the gemcitabine resistant cell lines, 74 PTVs were identified. After filtering out PTVs found in the chemo-naive cell lines, 32 were found in the gemcitabine resistant cell lines. Of the 32 variants, 22 had available expression data for the affected gene and the patient's response when treated with gemcitabine. It was found that 21 variants were found to have similar consequences. When considering the expression data of the genes associated to the 21 variants, only four had a p-value of < 0.05, with nine showing an FDR (0.05) significance and the other showing an FDR (0.2) significance. Kaplan-Meier graphs were created for the remaining 13 genes. Kaplan-Meier graphs were created for the remaining 13 genes. Six showed that low expression of the gene resulted in patient death with a p-value of <0.05, although it was important to note that very little difference in the final outcome of low expressing and high expressing patient response was seen (Figure 5.35, Appendix A6). One of the genes were identified to be in the cancer gene census, but none are considered to be a hallmark of cancer. The identified genes included; *AGAP6, MSH2, MTCH2, SMC1B, SYNGR1,* and *TYK2*.

AGAP6 encodes for the Arf-GAP with GTPase, ANK repeat and PH domain-containing protein 6, which is predicted to be a putative GTPase-activating protein (UniProtKB, Q5VW22). When *AGAP6* expression is low, there is poor patient response when treated with gemcitabine with a statistical significance of p = 0.01 (Figure 5.35A), although its implication in gemcitabine resistance is unknown.

MSH2 encodes for the DNA mismatch repair protein Msh2, which is a major component of the DNA MMR pathway, and has been identified in the cancer gene census (Li, 2008). When *MSH2* expression is low, there is poor patient response when treated with gemcitabine with a statistical significance of p = 0.025 (Figure 5.35B). This is consistent with the predicted LOF splice variant found in HCC38^{r}GEM^{20}. However, this pattern is in contradiction with the work by Cloyd *et al.*, 2017, who showed that down-regulation of MSH2 enhances cytotoxicity induced by gemcitabine in human lung adenocarcinoma cells, and Dong *et al.*, 2009, who showed defects in MMR sensitises pancreatic cancer to gemcitabine. Also, as previously mentioned, defects in MMR have been associated with cisplatin and 5-Fluorouracil resistance. HCC38^{r}GEM^{20} does demonstrate cross-resistance to both cisplatin and 5-Fluororuacil (Figure 3.11, Figure 3.16).

The gene *MTCH2* encodes for the mitochondrial carrier homolog 2 protein, although the substrate in which it transports is not yet known (UniProtKB, Q9Y6C9). When *MTCH2* expression is low, there is poor patient response when treated with gemcitabine with a statistical significance of p = 0.018

(Figure 5.35C) which is consistent with the predicted LOF stop gain variant identified in HCC38^{r}GEM[20]. The implication of *MTCH2* in gemcitabine resistance is unknown.

The frameshift variant found in *SMC1B* was already identified in the HCC38^{r}GEM[20] during the candidate approach analysis (section 5.2.2), although was not discussed due to the low VAF = 0.1. *SMC1B* encodes for the structural maintenance of chromosomes protein 1B, which has been implicated in being a meiosis-specific component of cohesion complex (Mannini *et al.*, 2015). When *SMC1B* expression is low, there is poor patient response when treated with gemcitabine with a statistical significance of p = 0.019 (Figure 5.35D), although its implication in gemcitabine resistance is unknown.

SYNGR1 encodes for the synaptogyrin-1 protein which has been predicted to play a role in the regulation of exocytosis (UniProtHB, O43759). When *SYNGR1* expression is low, there is poor patient response when treated with gemcitabine with a statistical significance of p = 0.021 (Figure 5.35E), although its implication in gemcitabine resistance is unknown.

Finally, the gene *TYK2* encodes the non-receptor tyrosine-protein kinase TYK2, which has been found to phosphorylate interferon-α/β receptor chains during interferon signalling (Colamonici *et al.*, 1994). When *TYK2* expression is low, there is poor patient response when treated with gemcitabine with a statistical significance of p = 0.012 (Figure 5.35F), although its implication in gemcitabine resistance is unknown.

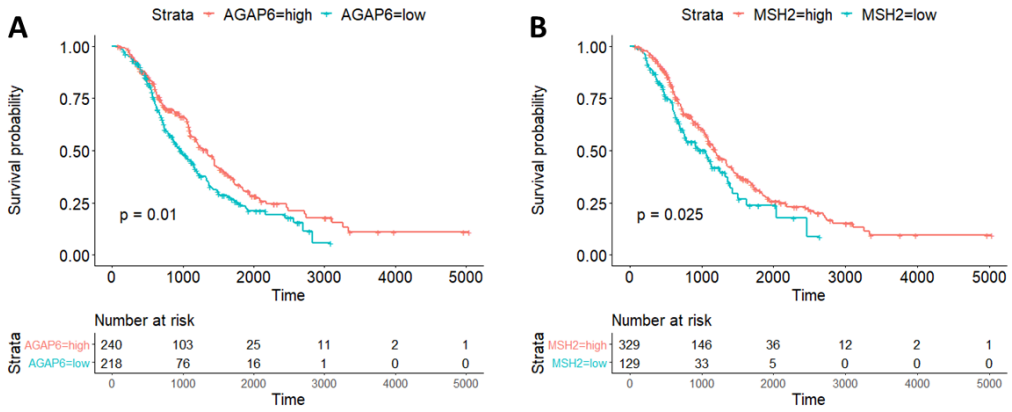


Figure 5.35 Patient survival analysis of genes containing protein truncated variants associated with gemcitabine resistance
Figure continued on next page

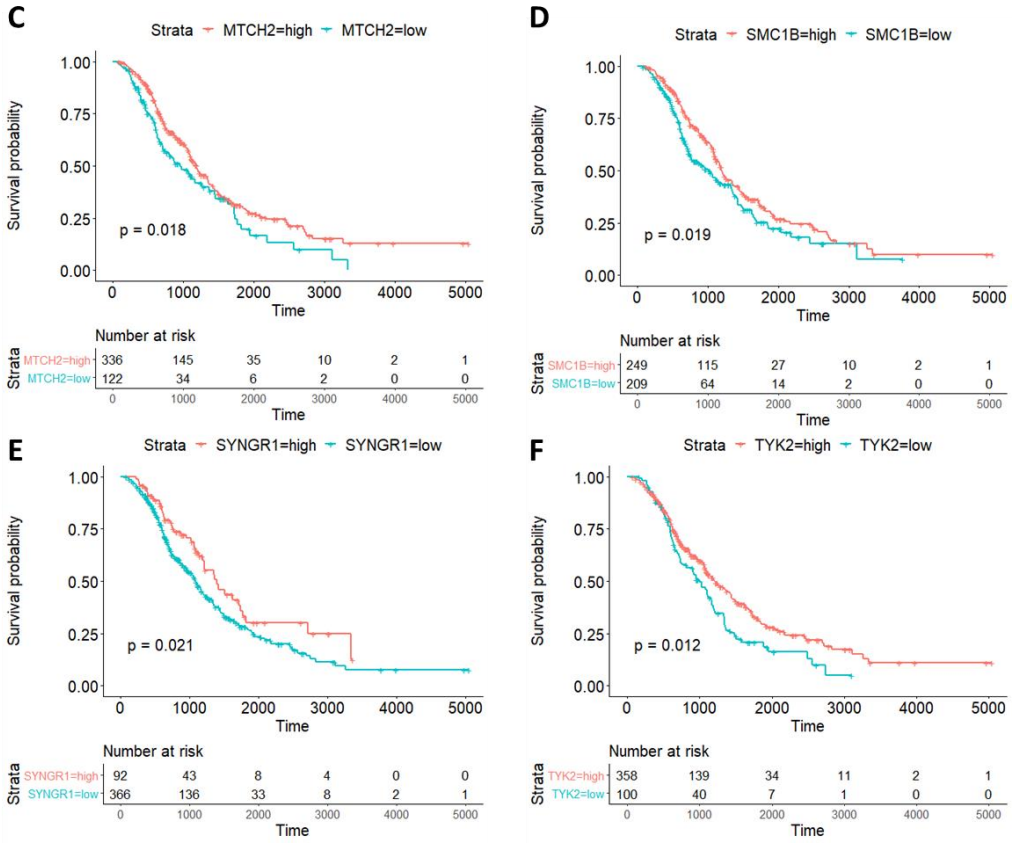


Figure 5.35 Patient survival analysis of genes containing protein truncated variants associated with gemcitabine resistance
All protein truncating variants found in the gemcitabine resistant TNBC cell lines were selected, with the exception of those seen in high confidence in the chemo-naive cell lines. Gene expression data for the gene was selected from the TCGA and filtered for statistical significance, $p < 0.05$ and false discovery rate < 0.2. Resulting mutated genes were selected if low gene expression showed higher patient mortality. Survival probability for the final mutated genes for when gene expression was high or low and the patient treated with gemcitabine was calculated and represented in Kaplan-Meier graphs.

When considering the paclitaxel resistant cell lines, 57 PTVs were identified. After filtering out PTVs found in the chemo-naive cell lines, 23 were found in the paclitaxel resistant cell lines. Of the 23 variants, 14 had available expression data for the affected gene and the patient's response when treated with paclitaxel. Of the 14 variants, 12 were found to have similar consequences. When considering the expression data of the genes associated to the 12 variants, all 12 had a p-value of < 0.05, and had an FDR (0.05) significance. Kaplan-Meier graphs were created for the remaining 12 genes. Seven showed that when patient had low expression of the gene, and were treated with paclitaxel, this resulted in patient death with a p-value < 0.05. None of the genes were identified to be in the cancer gene census, nor considered as a hallmark of cancer. The identified genes included; *ACIN1, SETX, SLC24A1, CUBN, DNAH5, INHBA,* and *KIAA0586* (Figure 5.36, Appendix A36).

The gene *ACIN1* encodes the apoptotic chromatic condensation inducer in the nucleus protein which has been found to a component of the splicing-dependent exon junction complex found at

splice junction on mRNAs (Schwerk *et al.*, 2003). When *ACIN1* expression is low, there is poor patient response when treated with paclitaxel with a statistical significance of p = 0.00015, which is consistent with the predicted LOF frameshift variant in MDA-MB-468^{r}PCL^{20} (Figure 5.36A). The role of ACIN1 in paclitaxel resistance is unknown.

The gene *SETX* encodes the RNA/DNA helicase senataxin protein which has been found to have roles in RNA metabolism, transcription regulation, as well as genomic integrity (Suraweera *et al.*, 2009). When *SETX* expression is low, there is poor patient response when treated with paclitaxel with a statistical significance of p < 0.0001, which is consistent with the predicted LOF frameshift variant in MDA-MB-468^{r}PCL^{20} (Figure 5.36B). The role of SETX in paclitaxel resistance is unknown.

SLC24A1 encodes for the sodium/potassium/calcium exchanger 1 which has been shown to be a component of the visual transduction cascade (McKiernan and Friedlander, 1999). When *SLC24A1* expression is low, there is poor patient response when treated with paclitaxel with a statistical significance of p < 0.0001, which is consistent with the predicted LOF frameshift variant in MDA-MB-468^{r}PCL^{20} (Figure 5.36C). The role of the variant in *SLC24A1* in paclitaxel resistance is unknown.

CUBN encodes for cubilin, which is an endocytic receptor, which have roles in lipoprotein, iron and vitamin metabolism by facilitating their uptake into the cell (Kozyraki *et al.*, 1999, 2001; Nykjaer *et al.*, 2001; Fyfe *et al.*, 2004). When *CUBN* expression is low, there is poor patient response when treated with paclitaxel with a statistical significance of p < 0.0001, which is consistent with the predicted LOF stop gain variant in HCC1806^{r}PCL^{20} (Figure 5.36D). The role of CUBN in paclitaxel resistance is unknown.

DNAH5 encodes the dynein heavy chain 5, which has been found to be the force generating protein in cilia, where it produces force towards the minus ends of microtubules (Oltean *et al.*, 2018). When *DNAH5* expression is low, there is poor patient response when treated with paclitaxel with a statistical significance of p < 0.0001, which is consistent with the predicted LOF stop gain variant in MDA-MB-468^{r}PCL^{20} (Figure 5.36E). The role of DNAH5 in paclitaxel resistance is unknown.

The stop gain variant found in *INHBA* was already identified in MDA-MB-468^{r}PCL^{20} during the candidate approach analysis (section 5.2.2). When *INHBA* expression is low, there is poor patient response when treated with paclitaxel with a statistical significance of p < 0.0001. (Figure 5.36F). This is consistent with the predicted LOF stop gain variant identified in *INHBA* in MDA-MB-468^{r}PCL^{20}, although the role of INHBA in paclitaxel resistance is unknown.

KIAA0586 encodes the protein TALPID3, which is known for ciliogenesis and hedgehog/SHH signalling. When *KIAA0586* expression is low, there is poor patient response when treated with paclitaxel with a statistical significance of $p < 0.0001$. (Figure 5.36G). This is consistent with the predicted LOF stop gain variant identified in *KIAA0586* in MDA-MB-468^{r}PCL^{20}, although the role of KIAA0586 in paclitaxel resistance is unknown.

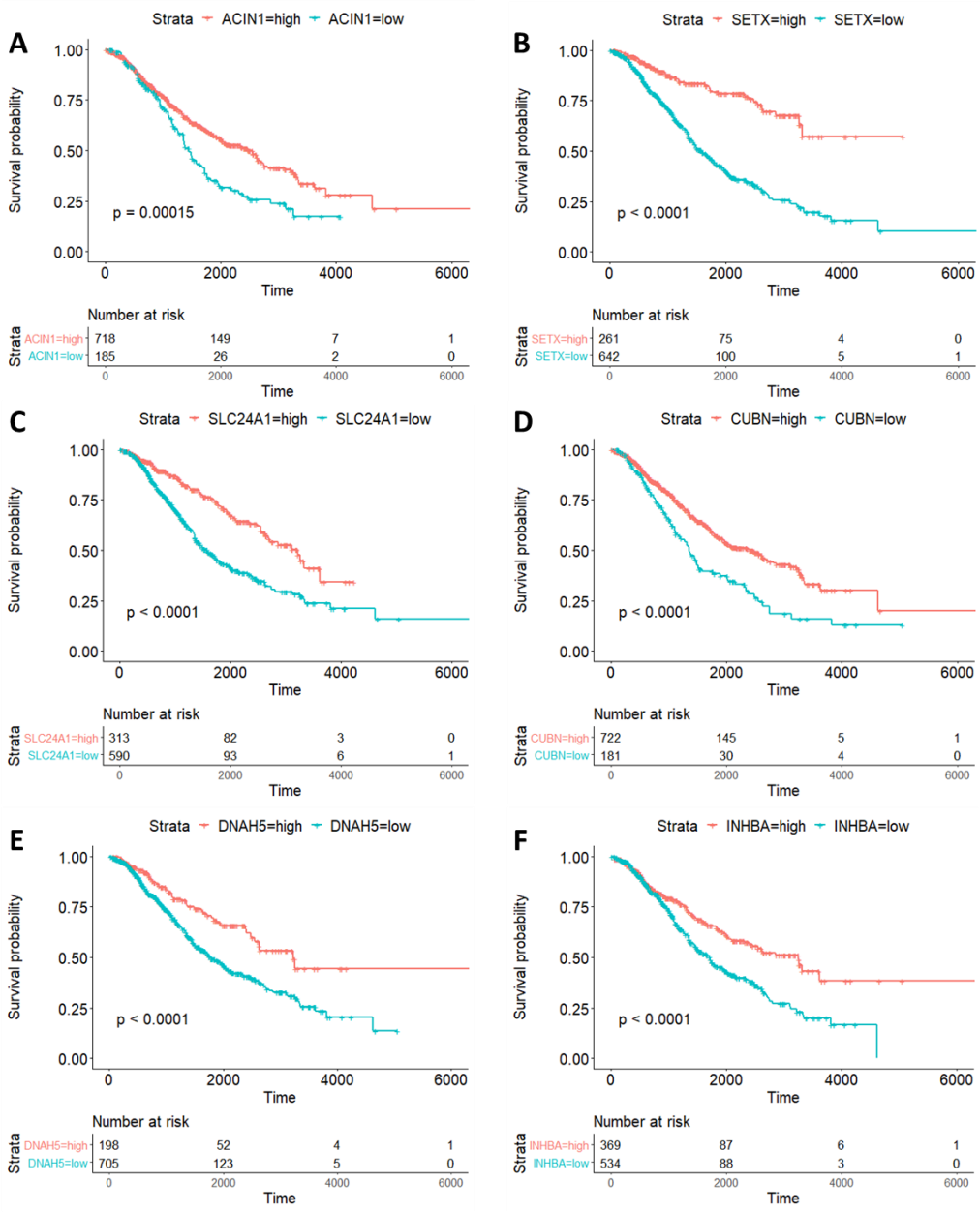


Figure 5.36 Patient survival analysis of genes containing protein truncated variants associated with paclitaxel resistance

Figure continued on next page

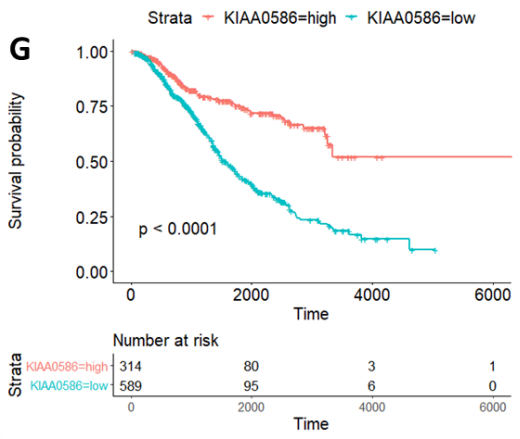


Figure 5.36 Patient survival analysis of genes containing protein truncated variants associated with paclitaxel resistance

All protein truncating variants found in the paclitaxel resistant TNBC cell lines were selected, with the exception of those seen in high confidence in the chemo-naive cell lines. Gene expression data for the gene was selected from the TCGA and filtered for statistical significance, $p < 0.05$ and false discovery rate < 0.2. Resulting mutated genes were selected if low gene expression showed higher patient mortality. Survival probability for the final mutated genes for when gene expression was high or low and the patient treated with paclitaxel was calculated and represented in Kaplan-Meier graphs.

When considering the 5-Fluorouracil resistant cell line, HCC1806^{r}5-F^{1500}, 31 PTVs were identified. After filtering out PTVs found in HCC1806, 17 was found in HCC1806^{r}5-F^{1500}. Of the 17 variants, 15 had available expression data for the affected gene and the patient's response when treated with 5-Fluorouracil. Of the 15 variants, 14 were found to have similar consequences. When considering the expression data of the genes associated to the 14 variants, 11 had a p-value of < 0.05, and all had an FDR (0.05). Kaplan-Meier graphs were created for the remaining 11 genes. Six showed that when patient had low expression of the gene, and were treated with 5-Fluorouracil, this resulted in patient death with a p-value < 0.05. None of the genes were identified to be in the cancer gene census, nor considered as a hallmark of cancer. The identified genes included; *CEBPZ, CNEP1R1, COG6, RPL14, EFCAB6,* and *SLC4A8* (Figure 5.37, Appendix A36).

CEBPZ encodes the CCAAT/enhancer-binding protein ζ, which has been found to stimulate transcription from the HSP70 promoter (UniProtKB, Q03701). When *CEBPZ* expression is low, there is poor patient response when treated with 5-Fluorouracil with a statistical significance of p = 0.0004, which is consistent with the predicted LOF splice variant in HCC1806^{r}5-F^{1500} (Figure 5.37A). The role of CEBPZ in 5-Fluorouracil resistance is unknown.

CNEP1R1 encodes the nuclear envelop phosphatase-regulatory subunit 1 protein which has been found to have a role in lipid metabolic processes. It has been found to form an active complex with CTDNEP1, which together can dephosphorylate and activate LPIN1 and LPIN2 which are required for the conversion of phosphatidic acid to diacylglycerol and control metabolism of fatty acids (Han

et al., 2012). When *CNEP1R1* expression is low, there is poor patient response when treated with 5-Fluorouracil with a statistical significance of p = 0.01, which is consistent with the predicted LOF splice variant in HCC1806^{r}5-F^{1500} (Figure 5.37B). The role of CNEP1R1 in 5-Fluorouracil resistance is unknown.

COG6 encodes for the conserved oligomeric golgi complex subunit 6 protein, which has been identified to be required for normal golgi function (UniProtKB, Q9Y2V7). When the *COG6* expression is low, there is poor patient response when treated with 5-Fluorouracil with a statistical significance of p = 0.0043, although the final patient outcome is not that different. This is consistent with the predicted LOF frameshift variant in HCC1806^{r}5-F^{1500} (Figure 5.37C). The role of COG6 in 5-Fluorouracil resistance is unknown.

RPL14 encodes for the 60S ribosomal protein L14, which is a component of the large ribosomal subunit (Odintsova *et al.*, 2003). A transcriptional profiling study was conducted with biopsy samples from metastatic gastric cancer patients, which had developed acquired resistance to combinational therapy of cisplatin and 5-Fluorouracil. Here, Kim *et al.*, 2011 found that the most highly represented functional category in the acquired resistance signature was Protein Synthesis, which included RPL14. When *RPL14* expression is low, there is poor patient response when treated with 5-Fluorouracil with a statistical significance of p < 0.0001, which is consistent with the predicted LOF frameshift variant in HCC1806^{r}5-F^{1500} (Figure 5.37D). Although the role of RPL14 in 5-Fluorouracil resistance is unknown, it has been previously implicated.

EFCAB6 encodes for the EF-hand calcium-binding domain-containing protein 6, which has been found to negatively regulate the androgen receptor through the recruiting of the histone deacetylase complex (Takahashi *et al.*, 2001). Mehta *et al.*, 2015, found that in the presence of 5α-dihydrotestosterone (DHT), a known androgen hormone, 5-Fluorouracil induced apoptosis is reversed. This chemo-preventive nature of the androgen DHT was reversed with the introduction of bicalutamide, an inhibitor of androgen action by binding to cytosol androgen receptors. When *EFCAB6* expression is low, there is poor patient response when treated with 5-Fluorouracil with a statistical significance of p = 0.0098, which is consistent with the predicted LOF stop gain variant in HCC1806^{r}5-F^{1500} (Figure 5.37E). It could be predicted that the LOF function variant in EFCAB6 can prevent its negative regulation on the androgen receptor, increasing androgen activity and preventing 5-Fluorouracil induced apoptosis.

SLC4A8 encodes for the electroneutral sodium bicarbonate exchanger 1, which is known to mediate the electroneutral sodium- and carbonate dependent exchange. It is thought to be involved in cell

pH regulation, and has been found to mediate lithium dependent HCO_3^- transport (Amlal, Burnham and Soleimani, 1999; Grichtchenko *et al.*, 2001). Although 5-Fluorouracil has not been implicated in direct transport with SLC4A8, Sijens, Baldwin and Ng, 1991, has shown that 5-Fluorouacil uptake is pH dependent, and the acidification seems to induce local retention of 5-Fluororuacil which suggests the existence of active transport. When *SLC4A8* expression is low, there is poor patient response when treated with 5-Fluorouracil with a statistical significance of p = 0.00045, which is consistent with the predicted LOF stop gain variant in HCC1806^{r}5-F^{1500} (Figure 5.37F). It could be predicted that the LOF function variant in SLC4A8 leads to decreased facilitated transport of 5-Fluorouracil into the cell, thereby preventing 5-Fluorouracil from inducing its cytotoxic damage.

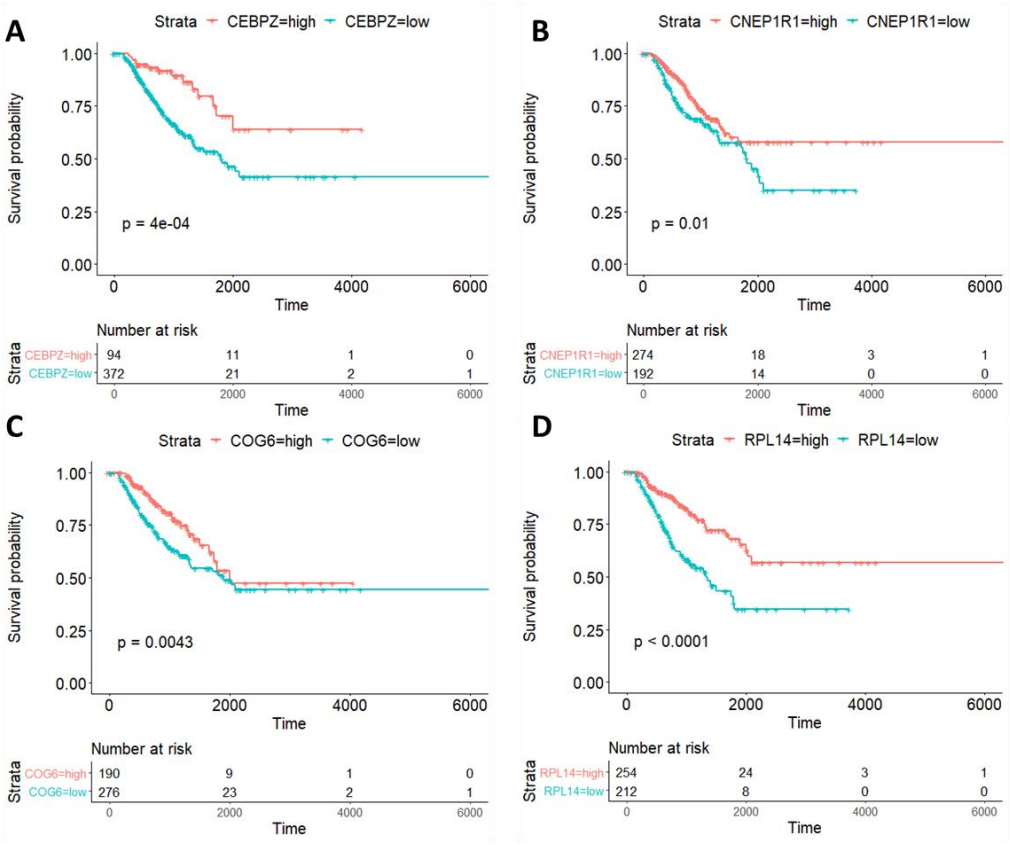


Figure 5.37 Patient survival analysis of genes containing protein truncated variants associated with 5-Fluorouracil resistance

Figure continued on next page

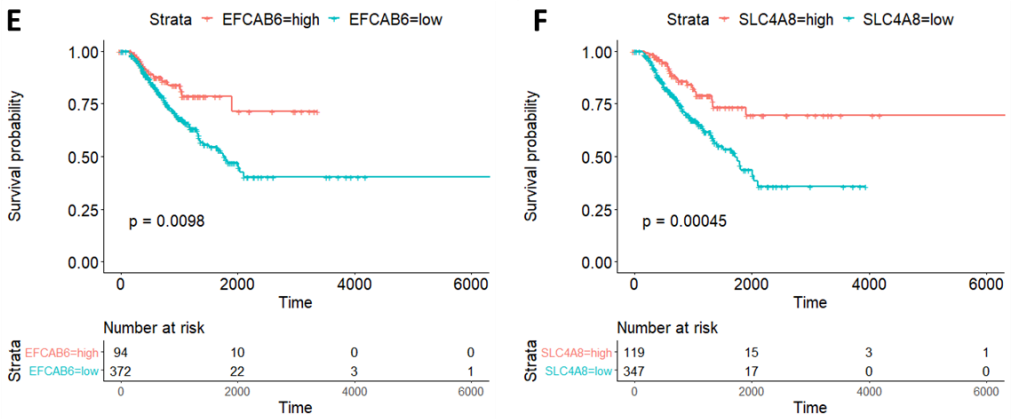


Figure 5.37 Patient survival analysis of genes containing protein truncated variants associated with 5-Fluorouracil resistance

All protein truncating variants found in the 5-Fluorouracil resistant TNBC cell lines were selected, with the exception of those seen in high confidence in the chemo-naive cell lines. Gene expression data for the gene was selected from the TCGA and filtered for statistical significance, $p < 0.05$ and false discovery rate < 0.2. Resulting mutated genes were selected if low gene expression showed higher patient mortality. Survival probability for the final mutated genes for when gene expression was high or low and the patient treated with 5-Fluorouracil was calculated and represented in Kaplan-Meier graphs.

5.3.3.3 Summary of TCGA analysis

Taking each of the TCGA analysis together (section 5.3.3.1 – 5.3.3.2), several genes have been identified which could be considered as potential biomarkers of resistance. Table 5.11 summarises each of the identified genes with consideration of the drug. Further validation is required to confirm if the variants in these genes result in the predicted drug resistance phenotype.

Table 5.14. Predicted biomarkers of chemo-resistance with clinical significance

Drug of resistance	Predicted clinically relevant biomarkers of resistance
Cisplatin	*COL22A1, EPB41, HUWE1, RGS9, FKBP7, KCND2, SLC2A12, OGN, BRD7, HSD17B3, IDO1, RGS9, KLF11, PRLR*
Doxorubicin	*ADNP, C11orf80(TOP6BL), C5orf42, GXYLT1, RNF213, USH2A, NCOR1, TRPM7, C20orf27, GBGT1, KCND2, SLC22A23, CDON, CES2, ZKSCAN3, ABCA8, TOP2A, ZNF442*
Gemcitabine	*ADNP, ABCB10, ANK2, TBC1D9, FGF14, AGAP6, MSH2, MTCH2, SMC1B, SYNGR1, TYK2*
Paclitaxel	*ADNP, DNAJC13, FAT4, GXYLT1, MSK1 (RPS6KAS), NCOR1, PHFT, EXT1, ITGB4, ACIN1, SETX, SLC24A1, CUBN, DNAH5, INHBA, KIAA0586.*
5-Fluorouracil	*ADNP, C11orf80(TOP6BL), CHST11, FLG, PIK3C2B, CEBPZ, CNEP1R1, COG6, RPL14, EFCAB6, SLC4A8.*

5.3 Discussion

In the previous chapters, the chemo-resistant TNBC cell lines have each demonstrated distinct resistance profiles to a panel of chemotherapeutic agents and to DDR targeted inhibitors. The aim of this chapter was to further characterise the chemo-resistant cell lines with consideration to their mutational profiles in order to hypothesise chemo-resistance mechanisms and identify clinically relevant candidate biomarkers driving resistance. To this end, WES was performed and analysed to identify genomic alterations in each of the chemo-naive and chemo-resistant TNBC cell lines in this panel.

5.3.1 Overview of exomic characterisation of chemo-naive and chemo-resistant TNBC cell lines

Upon receiving the exome sequencing files, the reads had to be aligned correctly against a reference genome and filtered in order to identify high confidence somatic variants. The two computational pipelines (Figure 5.2, 5.4) aimed to reduce the calling of false positive variants found from reads which harboured sequencing errors, mis-mapped short reads, PCR duplicates as well as mis-aligned INDELs to ensure veracity; the confidence or trust ability in the WES data (Robasky, Lewis and Church, 2014). However, a major shortcoming of WES is the uneven coverage of sequence reads over exome targets. This can contribute to low coverage regions, which therefore affects the downstream analysis and hinders accurate variant calling (Q. Wang *et al.*, 2017). Uneven coverage was observed in HCC38^{r}CDDP^{3000}, HCC38^{r}DOX^{40} and HCC38^{r}GEM^{20}, which impacted both the number of high confidence somatic variants that were called in these cell lines, and the analysis which considered the shift in VAF between the chemo-naive and chemo-resistant populations.

Analysis of the variants showed that there were often fewer variants present in the chemo-resistant cell lines compared to the respective chemo-naive cell line (Figure 5.6A, Appendix A5). Whilst it could be thought that new variants are introduced to the chemo-resistant cell lines through the process of developing a resistance phenotype, it seems paradoxical that less variants are observed overall. Importantly, the cell lines are not clonal, and are heterogenous in nature. By forcing the cell line to grow in a chemotherapeutic agent, the distribution of variants in the cell population will change, and some variants, such as those with a low VAF or not considered beneficial for the resistance mechanism, will be lost. Alternatively, the variant may still be present in the chemo-naive cell line, but at a level that no longer passes the stringent filtering requirements of three reads required to support the variant call. Furthermore, a variant may not be called as a result of an artefact of sequencing.

It was found that the number of variants called was significantly higher than the number of mutated genes (Figure5.6). This could suggest that multiple variants are present in the same gene, i.e. variant 'X' and 'Y' are present in gene 'Z'. In this scenario, whilst variant 'X' may be damaging alone, when present with variant 'Y', variant 'Y' could be compensatory for the damage occurring from 'X'. Alternatively, both 'X' and 'Y' variants can be damaging, causing disruption to the proteins structure and/or function.

However, given that the cells populations are heterogenous, it is more likely to indicate that a sub-population has variant 'X' in gene 'Z' whilst another sub-population has variant 'Y' in gene 'Z'. This is further supported by the fact that often there are differences in the VAF for each of the variants in the gene. This observation can indicate that a variant in gene 'Z' is beneficial for the progression of cancer, or drug resistance across the cell population, regardless of the position within the gene. The TNBC cell lines tested in this thesis were generated as polyclonal populations. To determine this if this is the case, subcloning can be performed to isolate resistant colonies and analyse the variants in gene in the clonal population. This is an established method for determining strong candidate resistance mechanisms or biomarkers, however not all resistance mechanisms can be isolated in this manner. For example, resistant mechanisms which have co-evolved together and rely on co-dependence (Burrell and Swanton, 2014).

It was also observed that there were more high confidence somatic non-synonymous variants called than there were high confidence somatic synonymous variants. This result was surprising as synonymous variants are normally more prevalent than non-synonymous variants in cancer genes (Chu and Wei, 2019). It could be predicted that synonymous variants are more prevalent as germline variants in these cell lines, or that it may be in part due to the lack of overall discovery power, hence the removal from the final call set (Lawrence *et al.*, 2014).

It is important to note, that whilst synonymous variants are identified in each of the TNBC cell lines, the prediction of the phenotypic outcome for gene/protein affected in the cell population harbouring these variants, are difficult to estimate. Previously, synonymous variants were intuitively thought to be functionally silent and evolutionary neutral, however, this is now known not to be the case. Synonymous variants can disrupt processes such as transcription, splicing, co-translational folding and mRNA stability (Fyfe *et al.*, 2004; Pagani, Raponi and Baralle, 2005; Pechmann and Frydman, 2013; Stergachis *et al.*, 2013; Presnyak *et al.*, 2015). Although there are some tools which claim to be synonymous effect-predictors, there are currently no publicly available, large collection of synonymous variants, which have experimentally validated effect annotations, that can be used to build a gold standard data test set for a successful effect-predictor

(Adzhubei, Jordan and Sunyaev, 2015). To this end, although synonymous variants are identified here, these were not taken into account in the larger analyses conducted in this chapter, including the candidate approach and comparison to the TCGA.

It was observed that each of the TNBC cell lines harboured several variants in mucin (MUC) genes. Large genes, such as TTN and MUC-genes, have a higher probability of harbouring variants due to their size (Shyr *et al.*, 2014). MUC-gene variants were observed in both chemo-naive and chemo-resistant cell lines, and in *de novo* and *not called* gene lists. When GO analysis was performed, the cell lines would be enriched with terms such as; O-glycan processing (GO:0016266), extracellular matrix structural constituent (GO:0005201), extracellular matrix constituent, lubricant activity (GO:0030197), extracellular matrix (GO:0031012), as a result of the different mutated MUC genes. These genes have been found to be commonly mutated or dysregulated in cancer, including breast cancer (Ringel and Löhr, 2003; Mukhopadhyay *et al.*, 2011). Therefore, the fluctuation of MUC-genes in the TNBC cell lines are not thought to be driving drug resistance, but are predicted to be related to cancer development and progression.

GO analysis of the called variants in the TNBC cell lines, also identified an enrichment in terms associated with mitochondrial ATP, NADH dehydrogenase, and electron transport chains. These terms were attributed to mitochondrial genes, such as *MT-ND1*. It has been found that genomic alterations in mitochondrial genes can be a result of the close proximity of ROS formed during metabolic events, increasing the risk of mitochondrial DNA perturbation and instability (Giampazolias and Tait, 2016; Hertweck and Dasgupta, 2017). It is thought that the combination of ROS and dysregulated DDR mechanisms, could result in a higher mitochondrial DNA mutation rate with an order of magnitude higher than the nuclear genome (Khrapko *et al.*, 1997). It is therefore unsurprising to find these genes commonly mutated across the cell lines, and like the MUC-genes, are not thought to be driving drug resistance but are implicated in cancer development and progression.

The KEGG BRITE analysis showed that genes were associated to the same terms in both beneficially selected or not beneficially selected groups. No real significant changes were seen when considering a single term. This could be a result of positive and negative regulators associated to pathways. For example, if a resistance mechanism requires pathway ‘X’ to be upregulated; 1. There could be LOF variants in negative regulators of pathway ‘X’ beneficially selected for, which result in an upregulation of the pathway 2. There could be a reduction of LOF variants which are positive regulators of pathway ‘X’ which are not beneficially selected for, which also results in upregulation

of the pathway. To this end, the KEGG BRITE analysis demonstrated that simplification of GO analysis is not useful here.

There are further limitations when considering functionally enriched GO terms. Firstly, the input gene sets were small as the mutated genes had been focused down to ones that are considered interesting in the population. This leads to low calculated p-values by the tool, which means that only umbrella GO terms are statistically significant, such as “extracellular matrix”, which does not provide much useful information.

Secondly, functional enrichment is based on known functions of the genes/proteins. Many proteins are promiscuous in their intracellular functions, and signalling pathways have been found to cross-talk to large extents. Although functional gene enrichment can provide an idea of affected functions, it is not extensive, and can be context dependent. Further to this, the functional enrichment does not give any indication of how significant a variant in a gene may be. For example, an input of genes A, B, C, D, E could lead to the identification of statistically significant enriched term X which is attributed to the genes A, B and C. However, A, B, C may not have damaging variants, whilst D and E, not assigned to any enriched terms, have significantly damaging variants.

An interesting find in this chapter was that very few variants, or mutated genes were found to be shared between chemo-resistant cell lines which were developed to have resistance to the same drug. Even less so when the *de novo* or *not called* variants are considered, and those identified were often variants in the MUC-genes. It can be speculated that this finding may be a consequence of two things:

1. The heterogeneity nature of the chemo-naive cell line from which the chemo-resistant cell lines are derived. The chemo-naive population harbours a collection of cells with distinct molecular signatures and have been shown to have differential levels of sensitivity to treatment with both chemotherapeutic agents and DDR targeted drugs (Figure 4.6). This different biological context can result in genetically distinct subpopulations when undergoing genetic changes when forced to grow in drug.

2. Several mechanisms of drug resistance can occur to any given drug. Multiple drug resistant mechanisms can exist in any given heterogenous population, which means there are multiple ways in which a cell can develop resistance to these agents (Fodale *et al.*, 2011).

5.3.2 Predicted mechanisms of chemo-resistance

5.3.2.1 Cancer stem cell like properties in MDA-MB-468^{r}PCL^{20}

MDA-MB-468^{r}PCL^{20} has been observed to show distinct differences compared to MDA-MB-468. It was observed in chapter 3 that the cell line underwent a morphological change, resulting in a cell population with elongated mesenchymal shapes (Figure 3.1). This chapter identified that whilst MDA-MB-468^{r}PCL^{20} and MDA-MB-468 had a similar number of variants called, there were very few *shared* between them, and MDA-MB-468^{r}PCL^{20} had many variants which are considered *not called* or *de novo* (Figure 5.8). Together, this data supports the idea that MDA-MB-468^{r}PCL^{20} has drifted from MDA-MB-468 as a result of a sub-population of the heterogenous cell population surviving after growth in a paclitaxel created genetic bottle neck.

As previously mentioned, changes in cancer cells driven through EMT have been identified as properties of cancer stem cells (CSCs) (Phi *et al.*, 2018). The CSC concept postulates that a sub-population of cells with stem-cell properties can serve as a critical driver of tumour progression, with the phenotype of CSCs to be attributed to epigenetic changes caused by the activation of the EMT programme (Shibue and Weinberg, 2017a). CSCs are known to express ABC transporters, including MDR1, which can result in drug efflux and drug resistance (Moitra, 2015; Sugano *et al.*, 2015). Evidence that MDR1 is upregulated in MDA-MB-468^{r}PCL^{20}, compared to MDA-MB-468, was seen in chapter 3 (Figure 3.17), and supported by the observed cross-resistance patterns to known MDR1 substrates in chapter 3 and 4. Further evidence supporting the hypothesis of changes in EMT was seen in this chapter with the identification of variants linked to this pathway.

A *de novo* stop gain variant was identified in *INHBA*, and analysis of this variant using the TCGA database predicted that a LOF variant may lead to paclitaxel resistance, as patients with low *INHBA* expression had a poorer outcome when treated with paclitaxel with $p < 0.0001$ (Figure 5.36F). INHBA is a subunit of both activin and inhibin, which are two closely related glycoproteins but with opposing biological effects. Both belong to the TGF-β family, which is responsible for the control of EMT (Kahata, Dadras and Moustakas, 2018). Additionally, an inframe insertion variant was identified in *PHF2*, and analysis of this variant using the TCGA database predicted that a LOF variant may lead to paclitaxel resistance, as patients with low *PHF2* expression had a poorer outcome when treated with paclitaxel with $p < 0.0001$ (Figure 5.31G). PHF2 has a role in demethylating both histones and non-histone proteins once activated by PKA (Baba *et al.*, 2011). It has been shown that increase of PKA induces MET (mesenchymal – epithelial transition) through PHF2 which relieves H3K9me2/3 -mediated repression of epithelial genes, and MET induced differentiation is

accompanied by a loss of stem-like properties which results in a sensitivity to chemotherapeutic drugs such as doxorubicin and paclitaxel (Pattabiraman *et al.*, 2016).

Both of these LOF variants demonstrate clinical significance to the response of paclitaxel, and both of the affected genes are implicated in EMT. Taking the data together, it could be hypothesised that changes in EMT, and the development of a CSC population, result in the observed resistance phenotype. Further experimental validation will be required to determine if a CSC population is observed in this cell line, and if the variants are contributing towards the resistance mechanism. Detection of CSC markers, such as CD133 and CD44, through western blot analysis, as well as the identification of Hoechst isolated side populations through flow cytometry can be used to determine the first, whilst gene knockdown studies in MDA-MB-468 and stable transfection studies in MDA-MB-468^{r}PCL^{20} , can confirm the latter. (Wang, Wang and Zhong, 2015).

5.3.2.2 Candidate cisplatin resistance mechanisms

Analysis of the cisplatin resistant cell lines identified both variants and mutational signatures which have been previously implicated in cisplatin resistance. MDA-MB-468^{r}CDDP^{1000} and HCC38^{r}CDDP^{3000} demonstrated a similarity to the mutational signature SBS31, which has been associated with prior chemotherapy treatment with platinum drugs (Figure 5.16). This highlights the usefulness of the chemo-resistant cell lines as appropriate models to investigate chemo-resistance.

A stop gain variant was identified in *SMARCAL1*, in MDA-MB-468^{r}CDDP^{1000}, whose role is known to be involved in fork reversal during DNA repair. It has been shown that depletion of SMARCAL1 restores fork protection in BRCA1/2 depleted cells and this increased drug resistance to cisplatin and PARP inhibitors (Cantor and Calvo, 2017; Kolinjivadi *et al.*, 2017; Taglialatela *et al.*, 2017). Interestingly, whilst MDA-MB-468^{r}CDDP^{1000} did not harbour any high confidence somatic variants in BRCA1/2, it did demonstrate cross-resistance to PARP inhibitors (FIGURE 4.4). This suggests that SMARCAL1 may be implicated in cisplatin and PARP inhibitor resistance, although the context of BRCA1/2 may be independent.

A splice acceptor variant in *BRD7* was identified in MDA-MB-468^{r}CDDP^{1000} and analysis of this variant using the TCGA database predicted that a LOF variant may lead to cisplatin resistance, as patients with low *BRD7* expression had a poorer outcome when treated with cisplatin with $p < 0.0001$ (Figure 5.33A). BRD7 interacts and negatively regulates YB-1 phosphorylation levels, and YB-1 has been implicated in response to drugs such as cisplatin, and also has been found to be overexpressed in cisplatin resistant cancer cell lines (Ohga *et al.*, 1996; Tomoko *et al.*, 1999; Niu *et*

al., 2020). It could be predicted that a LOF variant in *BRD7* prevents the negative regulation of YB-1, resulting in overexpressed of YB-1, resulting in cisplatin resistance.

Also identified in MDA-MB-468^{r}CDDP^{1000} was a *de novo* splice donor variant in *IDO1*. Analysis using the TCGA database predicted that a LOF variant may lead to cisplatin resistance, as patients with low *IDO1* expression had a poorer outcome when treated with cisplatin with $p < 0.0001$ (Figure 5.33C). However, whilst IDO1 has been previously implicated in cisplatin resistance, it in contradiction to the two observations made in this thesis. Nguyen *et al.*, 2020, has shown that increased expression of IDO1 results in cisplatin resistance in lung cancer. The TCGA analysis addressed pan-cancer samples, the cell lines used here, TNBC cell lines, whilst Nguyen *et al.*, 2020, used two cisplatin resistant NSCLC models. Therefore, whilst IDO1 has been observed to be implicated in cisplatin resistance in each case, this contradiction in predicted roles may be as a result of the context of the cancer type. Further investigation is required to determine the role of IDO1 in cisplatin resistance.

With these genes already shown to be implicated in cisplatin resistance in literature, it gives confidence that their dysregulation is involved in the cisplatin resistant mechanism within these cell lines. However, further evidence is required to predict their success as a predictive biomarker of cisplatin resistance. Gene knockdown studies in MDA-MB-468 and exogenous expression of a wild type version of the gene into MDA-MB-468^{r}CDDP^{1000} may be used to determine the resistance, or sensitive phenotype respectively.

5.3.2.3 Dysregulation of TOP2A driving doxorubicin resistance

When considering the doxorubicin resistant cell lines, it was interesting to find that they harboured variants, which either directly or indirectly affected the target of doxorubicin; TOP2A.

A predicted LOF frameshift variant was identified in *SOX2* in HCC1806^{r}DOX$^{12.5}$. SOX2, a transcription factor, has been shown to be a target for *TOP2A* (Lachmann *et al.*, 2010). It could be predicted that loss of SOX2 results in reduced transcription of *TOP2A*, thereby reducing the amount of available TOP2A for doxorubicin to bind to. A missense variant was identified in *CSNK1D* in HCC38^{r}DOX^{40}. CSNK1D has been shown to phosphorylate TOP2A at serine 1106 which enhances TOP2A activity and results in sensitivity to Topo II – targeted drugs in vivo (Chikamori *et al.*, 2003; Grozav *et al.*, 2009). It could be predicted that missense variant in *CSNK1D*, could lead to a downregulation in CSNK1D activity, or reduction of protein-protein-interaction, resulting in the hypophosphorylation of TOP2A and ultimately resistance to doxorubicin. A frameshift variant was also identified in *TOP2A*

itself, with a VAF equal to 0.8 in HCC1806^{r}DOX$^{12.5}$ cell line. Analysis using the TCGA database predicted that a LOF variant may lead to doxorubicin resistance, as patients with low *TOP2A* expression had a poorer outcome when treated with doxorubicin with $p < 0.0001$ (Figure 5.34A). It was predicted that the LOF variant in *TOP2A* in HCC1806^{r}DOX$^{12.5}$, was driving doxorubicin resistance.

Western blotting analysis confirmed decreased protein expression of TOP2A in HCC1806^{r}DOX$^{12.5}$ compared to HCC1806, and TOP2A targeted siRNA knockdown in HCC1806 resulted in a doxorubicin resistant phenotype (Appendix A32, Figure 5.19). These data show that doxorubicin resistance was achieved in HCC1806 through siRNA mediated TOP2A knockdown, suggesting that reduced TOP2A is driving the resistance mechanism in these cell lines context. Expression of wild type *TOP2A* into HCC1806^{r}DOX$^{12.5}$ to determine if sensitivity to doxorubicin is observed would provide further evidence that doxorubicin resistance is a result of the LOF variant in *TOP2A*.

Taken together, these data suggest that LOF variants, or reduced expression of *TOP2A* could act as a predictive biomarker of doxorubicin resistance. Furthermore, genetic perturbations in genes, which encode proteins that regulate TOP2A could also be candidate biomarkers for prediction of doxorubicin sensitivity. Use of these biomarkers in the clinic could inform clinicians of when resistance has developed to doxorubicin, and when a change in treatment is required.

5.3.3 Identification of clinically relevant biomarkers of chemo-resistance

It was reasoned that in order to identify relevant biomarkers of chemo-resistance, the identified variants need to be translational to clinical data. To this end, the publicly available clinical data from the TCGA was used to identify clinically relevant biomarkers of chemo-resistance.

Two analyses were performed using data from; 1. Beneficially selected gene set and 2. Protein truncating variants, in order to identify LOF and GOF biomarkers of chemo-resistance. Whilst data was not available in the TCGA for eribulin, variant and gene expression data was available for the rest of the chemotherapeutic agents addressed in this thesis. By comparing the variant data sets from this thesis and the TCGA, and combining statistically significant gene expression data to produce Kaplan Meir graphs, 70 candidate biomarkers of chemo-resistance were identified.

The identified candidate biomarkers have been implicated in chemo-resistance, not only in the cell line models tested in this thesis, but have now been predicted to be implicated in resistance in primary cancer cell lines. The candidate biomarkers must undergo validation to determine if they

can effectively predict drug resistance in the clinic. If successfully validated, these biomarkers may be useful in identifying when drug resistance has occurred in a patient, and be indicative of when a change of therapy is required.

There are some limitations of comparing the variants identified in the chemo-resistant cell lines to the data extracted from the TCGA. Patient data was not available for all of the chemotherapeutic agents investigated in this thesis, which meant that the clinical relevance of variants identified in eribulin resistant cell lines could not be determined. Furthermore, the analysis was conducted using pan-cancer data due to the limited availability of breast cancer/TNBC specific data, which means that the calculated mortality is not disease specific. This may have resulted in the removal of data which was not statistically significant when considering the pan-cancer data sets, but may have been important for resistance mechanisms in a TNBC disease context. Alternatively, data may have been included which may not be appropriate when considering the context of TNBC. Additionally, whilst analysis of the Kaplan-Meier graphs showed statistically significant differences between the high and low gene expression data, the final patient outcome may have not shown much of a difference. The longer the time-line, the higher the chance of an increase in mortality as patients can die of other complications or diseases.

In summary, the data presented in this chapter has identified unique exome phenotypes in each of the chemo-resistant TNBC cell lines. This was investigated through the analysis of exome sequencing data from each of the chemo-naive and chemo-resistant TNBC cell lines. This data has shown that chemo-resistance has not been developed in the same way in cell lines developed to have chemo-resistance to the same chemotherapeutic agent. This chapter identifies chemo-resistant mechanisms, and predicts potentially 70 clinically relevant biomarkers of resistance which warrant further investigation to validate these candidate targets.

Chapter 6

General Discussion

6. General Discussion

6.1 Introduction

TNBC is an aggressive metastatic breast cancer commonly identified through the lack of the oestrogen receptor, progesterone receptor and human epidermal growth receptor 2 (Ryu *et al.*, 2011). Although initial patient response is good with chemotherapy treatment, patients often relapse as a result of acquired drug resistance (Carey *et al.*, 2007). It is evident that new treatment options are required for TNBC patients, as well as early identification of emergence of drug resistance to inform clinicians that a change of therapy is required.

Inhibitors targeting the DNA damage response and repair (DDR) pathway are entering the clinic. These drugs are under investigation both as a monotherapy, or in combination with a DNA damaging agent (Topatana *et al.*, 2020). The agents are used in cancer types with high replication stress, or in cancers that harbour genetic mutations, which will result in a synthetic lethal relationship when combined with the drug. Often, mechanisms of chemoresistance arise through the dysregulation of DDR pathways, which provides a synthetically lethal opportunity to combine with DDR targeted drugs. However, the use of DDR inhibitors as a second line therapy option after the emergence of chemoresistance has not been investigated. Furthermore, recognition that chemoresistance has occurred is important to inform clinicians that a second line therapy is required. Identification of genetic changes in the tumour cells can provide biomarkers to indicate therapy refractory TNBC.

Previous studies have shown that cell line models can be successfully used in order to identify clinically relevant biomarkers or mechanisms of drug resistance (Garraway and Jänne, 2012b). Studies using cell line models can consider patterns of cross-resistance, or acquired vulnerability, and analyse genetic changes to identify new therapeutic strategies for refractory TNBC. To this end, both chemo-naive and chemo-resistant TNBC cell lines were selected from the Resistant Cancer Cell Line (RCCL) collection and underwent cross-resistance profiling and exomic sequencing (Michaelis, Wass and Cinatl, 2019). The work presented in this thesis identified strong patterns of cross-resistance and acquired vulnerability to a panel of chemotherapeutic agents and DDR inhibitors. Furthermore, candidate biomarkers were identified for each chemotherapeutic agent analysed here, which may prove useful in identifying when a TNBC patient has developed chemo-resistance. These findings may assist clinicians in determining a therapeutic strategy, personalising the treatment of TNBC patients. This chapter discusses the key findings of this thesis, and the wider implications of these findings on the treatment of chemo-resistant TNBC, as well as suggestions for

future work that may help improve understanding of drug resistance in an heterogenous cancer type.

6.2 Summary of main findings and future work

6.2.1 Cell line heterogeneity drives development of different acquired drug resistance mechanisms

In this thesis, it was observed that chemo-resistant cell lines which have acquired resistance to the same chemotherapeutic agent demonstrated different cross-resistance patterns during drug profiling and also developed different *de novo* and *gained* variants. It is well documented that acquired drug resistance is driven, in part, by intra-tumour heterogeneity which gives rise to phenotypic diversity (Lim and Ma, 2019). Intra-tumour heterogeneity can exist at genetic, transcriptomic and proteomic levels. Several factors can influence heterogeneity including genomic instability, drug exposure, microenvironment interactions and extrinsic factors including pH, hypoxia and paracrine signalling (Burrell and Swanton, 2014). It is clear that the heterogenous nature of the chemo-naive cells line had an effect on the way in which acquired resistance to the chemotherapeutic agents was developed. Furthermore, due to the polyclonal nature of the cell lines, chemoresistance may be driven by more than one resistance mechanism within that cell population. This is also true of tumours, where it has been shown polyclonal resistance within a tumour is a result of multiple clonal populations. For example, it has been reported that 17/24 colorectal cancer patients with acquired resistance to panitumumab or cetuximab developed *de novo* mutations within the EGFR pathway, whilst the other therapy refractory patients did not (Bettegowda *et al.*, 2014). The observation that the TNBC chemo-resistant cell lines have developed resistance in multiple-ways is important to emphasise as tumour heterogeneity must be taken into account when determining an appropriate therapy for a patient.

6.2.2 Hypothesised drug resistance mechanisms

6.2.2.1 Cancer stem cell like properties driving paclitaxel resistance

This thesis proposes that an observed resistance mechanism identified in MDA-MB-468^{r}PCL^{20} is a result of the population of cancer cells being driven through an epithelial-to-mesenchymal transition (EMT), resulting in the cancer cell population having properties of cancer stem cells (CSCs). Evidence was seen in changes in morphology, MDR1 expression and variant distribution in MDA-MB-468^{r}PCL^{20} compared to MDA-MB-468, as well as the identification of high confidence damaging somatic variant in the genes *INHBA* and *PHF2* whose roles are implicated in EMT (section

5.3.2). The latter variants were also supported by data extracted from the TCGA, suggesting that LOF of these variants could result in paclitaxel resistance.

It has been observed in both pre-clinical and clinical samples that chemotherapy and/or radiotherapy eliminates the bulk of non-CSCs, whilst leaving behind a sub-population of CSCs which are resistant to the therapy (Shibue and Weinberg, 2017b). This resistance is thought to be a result of several factors including; increased expression of ABC transporters (including MDR1), elevated expression of antiapoptotic proteins and the slow proliferation rate of stem cells (Shibue and Weinberg, 2017b). The identification of the EMT phenotype in MDA-MB-468^{r}PCL^{20} supports the evidence in literature, however further investigation is required. Analysis of CSC markers, such as CD133 and CD44, as well as antiapoptotic proteins in MDA-MB-468^{r}PCL^{20} can provide further supporting evidence. However, it should be noted that the scientific community are still struggling to identify reliable markers to be used for the confirmation of the CSC phenotype (Shibue and Weinberg, 2017b).

This observation in MDA-MB-468^{r}PCL^{20} highlights the significance of monitoring for changes in EMT and distribution of non-CSCs and CSCs in patient tumour samples in the clinic, which can be used to inform clinicians on the development of a resistance phenotype. Targeting the EMT programme to eliminate CSCs offers a promising avenue for the treatment of drug resistant cancer, but a further understanding of the mechanistic link between EMT and CSCs is required. There have been several approaches proposed, which target EMT for therapeutic benefit, including; prevention of EMT induction, reversing the process of EMT and selective targeting of cells that have undergone EMT (Shibue and Weinberg, 2017b). However, there is still very little consensus on how to recreate a tumour microenvironment to induce EMT *in vitro*, preventing the testing of agents which can treat EMT for therapeutic benefit. Further research into EMT is required in order to target CSCs for reversal of chemotherapeutic resistance.

6.2.2.2 High MDR1 expression is implicated in MK-8776 resistance

This thesis proposes that high expression of MDR1 is associated with observed resistance to MK-8776. The evidence of this comes from the fact that three of the MDR1 expressing cell lines demonstrate cross-resistance to MK-8776, and MDA-MB-468^{r}PCL^{20}, only demonstrates cross-resistance to MK-8776 out of the four CHK1 inhibitors tested here (Figure 4.1D, Figure 3.17).

As previously explained, it has been shown that MK-8776 can restore sensitivity to chemotherapeutics that are substrates of MDR1 in cells overexpressing MDR1, by binding to high

expressing MDR1, stimulating the ATPase activity of the transporter, which has been hypothesised to competitively limit the uptake of substrates of MDR1 leading to the inhibition of efflux function (Cui *et al.*, 2019). However, literature has also shown that substrates of MDR1 can stimulate the ATPase activity of MDR1 transporter, which would support cross-resistance observation in this thesis. However, tariquidar, a third-generation MDR1 inhibitor, has been found to stimulate the MDR1 ATPase activity, but not be transported, which leaves the question if this is the same behaviour being observed with MK-8776 (Loo *et al.*, 2012).

Whilst MK-8776 is no longer in clinical trials, elucidation of this resistance mechanism may provide useful information on the stimulation of ATPase activity, and highlights the need for patient MDR1 status to be determined prior to treatment with MDR1 substrates. The role of MDR1 in MK-8776 resistance can be further investigated through a combinational assay of MK-8776 and a MDR1 inhibitor, such as zosuquidar, to determine if the cells are sensitised to MK-8776 activity, or through the use of a UIC2 shift assay (Park *et al.*, 2003).

6.2.3 Consideration of appropriate second line therapies after emergence of chemo-resistance

Analysis of the cross-resistance profiling identified some significant patterns, which, whilst they require further investigation, have the potential to provide clinical benefit. Importantly, it was found that the majority of the chemo-resistant cell lines demonstrated cross-resistance to doxorubicin and the two PARP inhibitors; olaparib and rucaparib (Figure 4.14). These data suggest that after chemo-resistance has emerged to any of these drugs, none of them would be an appropriate second line therapy for the other, which was further supported by data identified when considering the absolute measure of drug activity using the Δ method (section 4.2.5). This is a significant observation, as four PARP inhibitors; olaparib, niraparib, rucaparib and talazoparib, are now FDA approved for the treatment of breast, pancreatic, prostate and ovarian cancer (Topatana *et al.*, 2020). Understanding how resistance occurs is important for clinical decisions.

With PARP inhibitors being used in the clinic, it has been reported that cancers develop resistance which can be acquired or inherent and is multifactorial (Topatana *et al.*, 2020). One of the most widely accepted mechanisms is the restoration of the HRR pathway through secondary reversion mutations (Bitler *et al.*, 2017). Studies have also shown that BRCA-mutated tumours can develop resistance to PARP inhibitors, and also DNA damaging agents, such as platinum-based chemotherapies, through reversion or secondary BRCA mutations, which restores the HRR pathway (Norquist *et al.*, 2011; Barber *et al.*, 2013; Bitler *et al.*, 2017). Furthermore, independent of BRCA1/2

reversion mutations, increased stabilisation of replication forks can give rise to both PARP inhibitor and cisplatin resistance (Chaudhuri *et al.*, 2016; Topatana *et al.*, 2020). These resistance mechanisms can therefore result in both PARP and platinum-based chemotherapy resistance, such as cisplatin, and it could be speculated that this is driving the cross-resistance phenotype in all three of the cisplatin resistant TNBC cell lines in the panel. These date supports the idea that PARP inhibitors should not be given following the development of platinum-based therapy.

As previously explained, another common resistance mechanism is drug efflux via upregulation of transporters such as MDR1. Both olaparib and rucaparib are known substrates of MDR1, and so it was not surprising to find cross-resistance of these drugs were seen in the four cell lines (MDA-MB-468^{r}ERI^{50}, MDA-MB-468^{r}PCL^{20}, HCC1806^{r}ERI^{50} and HCC1806^{r}PCL^{20}) which demonstrated elevated MDR1 expression (section 3.2.4). Several chemotherapeutic agents (including doxorubicin, eribulin and paclitaxel) are substrates for MDR1, which suggests that prior to secondary treatment with olaparib and rucaparib, the status of MDR1 in the patient's tumour should be considered. Targeting of MDR1 in the context of PARP inhibitor resistance could be an effective approach to manage the resistant disease, but further research is required. Furthermore, whilst MDR1 is a common resistance mechanism to doxorubicin, olaparib and rucaparib, neither of the doxorubicin resistant TNBC cell lines tested here demonstrated increased MDR1 expression. This hints at a further undiscovered common resistance mechanism, independent of MDR1 expression, which has not been established in literature.

Importantly, it was found that very few of the chemo-resistant cell lines demonstrated cross-resistance to the compounds; CCT241533, B02, and SBE13, which inhibit; CHK2, RAD51 and PLK1 respectively. Given the previous observation that the majority of cell lines demonstrated cross-resistance to PARP inhibitors, which could be a result of increased HRR capacity or a dependence on HRR, it could result in sensitivity to inhibitors targeting proteins of the HRR pathway, such as B02 (RAD51 recombinase) and CCT241533 (CHK2). RAD51 has been shown to promote chemoresistance by facilitating HRR for the efficient repair of chemotherapy-induced DNA breaks (Wiegmans *et al.*, 2014; Harris, Rabellino and Khanna, 2018). It could be hypothesised that if the cell lines are reliant on HRR due to a "BRCAness" phenotype, or an increased HRR capacity, that inhibition of RAD51 will result in cellular death. These data suggest that inhibition of RAD51 and CHK2 may be an appropriate second line therapy after chemo-resistance has emerged. It is important to note that whilst B02 and CCT241533 are highly specific for their protein targets, these compounds have not entered clinical trials (Anderson *et al.*, 2011b; Huang *et al.*, 2012).

Only one cell line (MDA-MB-468^{r}PCL^{20}) was identified to demonstrate cross-resistance to the polo kinase inhibitor, SBE13, whilst seven demonstrated cross-resistance to the polo kinase inhibitor BI2536. SBE13 is a more selective drug targeting PLK1 than BI2536, which has been shown to also inhibit PLK2 and PLK3 at low concentrations (Bhullar *et al.*, 2018). It has been found that the two drugs also have a different response in primary cells, with SBE13 demonstrating a transient G0/G1 arrest whilst inhibition with BI2536 results in a stable G2/M arrest (Eckerdt, 2011). Whilst inhibition using SBE13 may be an appropriate second line therapy option after chemoresistance has emerged, the question remains whether inhibition of PLK1 is a therapeutic strategy. There is a lot of debate in the scientific community as to whether PLKs are good drug targets. Whilst a large number of PLK inhibitors have entered preclinical and clinical development, the majority exhibit dose-limiting toxicity, which ultimately narrows the therapeutic window for treatment (Lee *et al.*, 2015).

It was also found that, aside from the two cell lines developed to have resistance to gemcitabine, only one cell line (HCC38^{r}CDDP^{3000}) demonstrated cross-resistance to gemcitabine. This suggests that gemcitabine may be an appropriate drug to use after chemo-resistance has developed. As gemcitabine is already a clinically registered drug, this could be a good drug of choice, that is already readily available, for clinicians to use as a second line treatment.

Further investigation into appropriate second line therapies was conducted by considering the differential effects of the drug activity profiles using the Δ method, previously established by the National Cancer Institute (Bracht *et al.*, 2006). This analysis identified interesting drug profile correlations which could be speculated upon the inhibitor's usefulness as appropriate second line therapies to each other. Strong positive correlations were found between inhibitors which target CHK1, ATR, WEE1 and CHK2. This would suggest that as resistance increases to one of the inhibitors of the stated target, that resistance is also seen to the other. Importantly, this suggests that the inhibitors would not be an appropriate second line therapy after resistance has occurred to the other. Similar observations have been found recently in literature, where prexasertib resistant ovarian cancer cells demonstrated cross-resistance to the CHK1 inhibitor AZD7762 and the ATR inhibitor AZD6738 (Nair *et al.*, 2020).

Negative correlations were seen with the aforementioned inhibitors against inhibitors targeting ATM, RAD51 and PLK1/2, as well as cisplatin, indicating that treatment with these will be beneficial. With drugs targeting the ATR-CHK1 and ATM-CHK2 axis in investigation and in clinical trials, further investigation to this should be considered, as this could provide possibilities to clinically manage drug resistant TNBC. However, the question as to whether these drug profile correlations exist

irrespective of the chemo-resistant model is yet to be conclusively answered. While the research carried out and presented in this thesis demonstrates some strong correlations, further work is required to confirm this observation.

6.2.5 Identification of candidate biomarkers of chemo-resistance

Through analysis of the cell line panel exome data, variants which are considered to be beneficially selected for in the chemo-resistant cell lines were identified. These variants were then compared to both variant and gene expression data from The Cancer Genome Atlas (TCGA) database (Weinstein, 2013). Kaplan-Meir graphs were produced to identify the consequence of high or low gene expression data on patient's survival upon treatment with a stated chemotherapeutic agent. Statistical significance and false discovery rate were taken into account, and candidate biomarkers were identified for each chemotherapeutic agent considered. Importantly, these identified candidate biomarkers that have clinical relevance as the data extracted from the TCGA are patient derived samples from the clinic. This not only highlights the usefulness of cell line models for initial studies, but implicates the clinical relevance of these candidate biomarkers.

Experimental validation of a candidate biomarker can be conducted in the cell line models to determine if the variant is driving the resistance phenotype. This can be approached in two ways; firstly, through exogenous expression of the wild type gene in the chemo-resistant cell line to determine an observable reduction in chemoresistance, and secondly, through gene knockdown or deletion (through shRNA, siRNA or CRISPR) in the respective chemo-naive cell line to determine if a resistance phenotype is observed. The latter experiment was conducted in HCC1806 to determine if siRNA knockdown of TOP2A results in doxorubicin resistance. A resistance phenotype to doxorubicin was observed through successful knockdown of TOP2A in HCC1806, which suggests that the frameshift variant in HCC1806^{r}DOX$^{12.5}$ may be driving the resistance to doxorubicin (section 5.2.2.2). This is a key observation, as resistance to inhibitors of topoisomerase II have been reported in a number of tumour model systems and is also prevalent in clinically refractory tumours (Ganapathi and Ganapathi, 2013). This observation suggests that LOF variants in *TOP2A* may result in doxorubicin resistance in patients, and could be indicative of intrinsic or acquired resistance to doxorubicin.

This thesis has identified 14, 18, 11, 16 and 11 candidate biomarkers of resistance to cisplatin, doxorubicin, gemcitabine, paclitaxel and 5-Fluorouracil respectively. This opens up new avenues of investigation, not only to experimentally validate the candidate biomarkers, but to analyse their

usefulness in the clinic. If successfully validated, these candidate biomarkers have the potential to inform clinicians that a change in therapy is required, furthering the development of precision oncology.

6.3 Future studies

The majority of the future work that has been considered in this chapter has been predominantly focused on answering questions that have arisen throughout this work, which have not been satisfactorily addressed. This section will focus on asking new questions regarding the analysis of the emergence of drug resistance in the clinic and how the understanding of this can be beneficial for clinical decision making.

Cancer, and in particular TNBC, demonstrate a strong heterogeneity phenotype whereby the tumours have distinct molecular signatures, differential levels of sensitivity and non-uniform genetically distinct tumour cell populations (Dagogo-Jack and Shaw, 2018). This tumour heterogeneity can result in the emergence of complex drug resistance, and to this end, future studies need further address the problem of the heterogeneity nature of cancer. Whilst tumour biopsies can provide genetic information, it is only a snapshot of the tumour at a fixed time. The tumour evolutionary history can provide information useful for the progression of the tumour growth and treatment that may be required (Dagogo-Jack and Shaw, 2018). A serial longitudinal tissue biopsy can track the tumours development under therapy progression, and identify the emergence of sub-clones which may drive drug resistance in the patient, resulting in refractory therapy. Multi-region, single cell signalling can provide evidence of clones and liquid biopsies can be used to trace drug resistance emergence in patience early (Lim and Ma, 2019).

Through analysis of the exome sequencing data, certain biases were taken in order to focus the research conducted. This leaves open questions as to how the non-analysed variants may be contributing towards chemoresistance. For example, the work conducted in this thesis focused on somatic variants, using the GNOMAD database to remove variants present at the frequency of ≥ 0.001% in the GNOMAD database (Karczewski *et al.*, 2019). However, germline variants can affect cancer prognosis, response to treatment and therapy outcome. Chatrath *et al.*, 2020 recently identified 79 germline variants in individual cancers, and 112 germline variants in groups of cancers which can be used for prognosis. Furthermore, it can be the combination of a germline and somatic variant which results in the observed resistance phenotype, rather than just a somatic variant alone. There is not yet a comprehensive analysis of germline genome sequencing data that can be used

for the treatment of cancer. Understanding the impact of the germline genome, alone, or in combination with somatic variants may provide insight into new treatment options (Lin *et al.*, 2019).

Further to this, whilst this thesis focuses predominantly on non-synonymous variants, synonymous variants are now known to disrupt process such as transcription, splicing, co-translational folding and mRNA stability (Fyfe *et al.*, 2004; Pagani, Raponi and Baralle, 2005; Pechmann and Frydman, 2013; Stergachis *et al.*, 2013; Presnyak *et al.*, 2015). New predictor tools are required to analyse synonymous variants and their link to cancer progression and drug resistance. Additionally, analysis of variants outside of the protein coding region, can also be associated as somatic drivers of cancer, and drivers of drug resistance. Recently, a large scale project through the PCAWG focused on the analysis of non-coding somatic variants in cancer whole genomes, which showed that whilst point mutations and structural variants are less frequent in non-coding genes, they are drivers of cancers (Rheinbay *et al.*, 2020). More drivers in non-coding regions will be identified when more cancer genomes become available. Whole genome sequencing analysis outside of protein-coding regions may therefore identify both drivers of cancer initiation and progression, as well as cancer drug resistance.

One outstanding question of this thesis is: are the candidate biomarkers identified in this thesis relevant to TNBC? Currently, the number of datasets of TNBC patient samples available for analysis are low compared to other cancer disease types. This means that comparison to datasets, such as those available in the TCGA and GDSC, are often done as pan-cancer or breast cancer, as it was in this thesis. Whilst some interesting data may be identified, this is not disease specific. Analysis of disease specific data sets will provide more useful information for the treatment of TNBC. To this end, an increase in the collection and analysis of TNBC patient samples needs to be achieved.

In order to determine if the candidate biomarkers are found in refractory TNBC patients, investigation into clinical samples must be conducted. Whilst it is more common place for databases and clinical publications to have both matched normal tissue and tumour tissue for the investigation of cancer drivers, it is less common to find matched pre- and post-drug therapy tumour tissues. The curation of databases with analysis of pre- and post-drug therapy tumour tissues will allow for identification of changes in genomic alterations and help understand the role of inherited genomic variation in shaping the onset of drug resistance and drivers of drug resistance mechanisms. Furthermore, the creation of tissue banks with patient tumour samples pre- and post-drug therapy can allow for validation assays to confirm predicted candidate biomarkers of drug resistance.

Importantly, with the expansion of these databases, and large-scale genomics projects, an important question remains as to what sequencing hundreds of tumours may reveal. It is currently not clear if identified cancer-critical somatic alterations are found recurrently in specific genes, or if the combination of recurrent and other variants defines the cancer genome as a whole, and therefore the therapy required (Mardis and Wilson, 2009).

6.4 Clinical implications of these findings

The data presented in this work has assessed DDR inhibitors as an appropriate second line therapy option after the emergence of chemo-resistance in TNBC, and identified both potential resistance mechanisms and candidate biomarkers of chemoresistance. To this end, the data in this study may have several clinical implications including; 1) Identification of the emergence of chemoresistance and 2) Alteration of patient treatment strategies.

It can be predicted that cross-resistance to PARP inhibitors after the emergence of chemo-resistance in the clinic is probable. This may be especially notable after platinum-based therapy, such as cisplatin, as common resistance mechanisms to these agents, such as restoration or upregulation of the HRR pathway, have been previously identified in literature. Therefore, the data from this thesis suggests that PARP inhibitors would be an inappropriate second line therapy following cisplatin resistance, and vice versa. This observation is particularly important to note in the treatment of TNBC, as platinum-agents are often given as the first line therapy (after surgery). Furthermore, it is predicted that cross-resistance to doxorubicin may also be observed after the emergence of chemoresistance in the clinic. This work identifies that potential use of drugs targeting CHK2, Rad51 recombinase and PLK may be useful in the clinic after the emergence of chemoresistance in TNBC patients. Whilst more fit for purpose drugs are required to be developed, which inhibit these targets, the potential use of them in the clinic opens new possibilities to clinically manage chemo-resistant TNBC. These findings, therefore, may provide an insight into effective and ineffective secondary treatment strategies.

This work supports the idea that ABC transporters and EMT should be monitored routinely in the clinic in order to identify the emergence of chemo-resistance early. These resistance mechanisms have been found to result in a multi-drug resistance phenotype in both pre-clinical, as seen in this work, and also in clinical samples, as seen the literature. Early detection of markers associated to these resistance mechanisms would be beneficial in informing clinicians when a new therapeutic strategy needs to be adopted.

Finally, this work has identified 70 candidate biomarkers, which are predicted to be indicative of chemoresistance. Whilst these have been identified in pre-clinical models, the evidence is supported by data of patient tumour samples from the TCGA database. Of particular note, the loss of function, or down regulation of *TOP2A* is predicted as a candidate biomarker for doxorubicin resistance. These biomarkers could be indicative of when acquired resistance to chemotherapeutic agents has occurred, informing clinicians of the relapse of the patient and the requirement of a therapy change. By examining tumour samples prior to treatment and after relapse, these biomarkers can contribute towards evaluating the most appropriate second line treatment.

6.5 Concluding remarks

TNBC is an aggressive, heterogenous, metastatic cancer, lacking known druggable targets. As a consequence of this, acquired drug resistance emerges, which results in patient relapse and therapy failure. It is evident that an appropriate second line therapy is required following chemo-resistance, and that there is also a need for clinically relevant biomarkers to identify when this change of therapy is required.

In this thesis, a panel of six chemotherapeutic agents and 16 DDR inhibitors are considered as appropriate second line therapy options against a panel of three chemo-naive and 15 chemo-resistant TNBC cell lines. This work identified that cross-resistance to the PARP inhibitors olaparib and rucaparib, as well as the chemotherapeutic agent doxorubicin, was seen in the chemo-resistant TNBC cell lines. This suggests that these drugs would be a poor choice as a second line treatment after the emergence of chemo-resistance. However, little cross-resistance was seen to inhibitors which target CHK2, RAD51 recombinase and PLK1, suggesting that after the emergence of chemo-resistance, these would be an appropriate second line therapy option. Therefore, these findings provide a beneficial insight into effective and ineffective second treatment strategies for refractory TNBC.

Using the TNBC cell lines exome sequencing data, as well as exome sequencing and gene expression data of patient samples extracted from the TCGA, 70 clinically relevant candidate biomarkers were identified, which may be used to predict when chemo-resistance has occurred. Importantly, a loss of function frameshift in *TOP2A* was predicted to result in doxorubicin resistance. Validation experiments in the cell lines supported this observation, suggesting that loss of function variants, or down regulation of *TOP2A* is a biomarker of doxorubicin resistance. As such, this work provides

a promising basis for future studies required to validate the 70 candidate biomarkers in both cell lines and patient tumour samples.

In conclusion, this thesis provides an insight into the use of chemotherapeutic agents and DDR inhibitors as an appropriate second line therapy option after the emergence of chemo-resistance. It also has identified 70 clinically relevant candidate biomarkers which are predicted to be indicative of chemo-resistance, and provides new avenues of research for further exploration of these findings. Therefore, the work presented in this thesis could help advance understanding of chemoresistance in the clinic, and potentially improve the outcomes of TNBC patients that develop chemoresistance.

Appendix

Appendix list

A1. Drug chemical structures. Chemical structures of each of the chemotherapeutic agents and DNA damage response and repair inhibitor used.

A2. Cross-resistance profiling to chemotherapeutic agents. GI_{50} values and resistance factor for each of the chemo-naive and chemo-resistant cell lines when treated with the chemotherapeutic agents.

A3. Cross-resistance profiling to DNA damage response and repair inhibitors. GI_{50} values and resistance factor for each of the chemo-naive and chemo-resistant cell lines when treated with the DNA damage response and repair inhibitors.

A4. Basic FASTQC report. Total number of reads, sequence length and GC % for each cell line across the multiple sequencing lanes for each chemo-naive and chemo-resistant TNBC cell line

A5. Filtering of called variants. Numbers of called variants at each step through the variant filtering pipeline to obtain the final high confidence somatic variants for each chemo-naive and chemo-resistant TNBC cell line.

A6. High confidence somatic variants called in MDA-MB-468. Chromosome position, base change from the reference genome to variant alternative and gene affected for all high confidence somatic variants called in MDA-MB-468.

A7. High confidence somatic variants called in MDA-MB-468^{r}CDDP^{1000}. Chromosome position, base change from the reference genome to variant alternative and gene affected for all high confidence somatic variants called in MDA-MB-468^{r}CDDP^{1000}.

A8. High confidence somatic variants called in MDA-MB-468^{r}DOX^{50}. Chromosome position, base change from the reference genome to variant alternative and gene affected for all high confidence somatic variants called in MDA-MB-468^{r}DOX^{50}.

A9. High confidence somatic variants called in MDA-MB-468^{r}ERI^{50}. Chromosome position, base change from the reference genome to variant alternative and gene affected for all high confidence somatic variants called in MDA-MB-468^{r}ERI^{50}.

A10. High confidence somatic variants called in MDA-MB-468^{r}PCL^{20}. Chromosome position, base change from the reference genome to variant alternative and gene affected for all high confidence somatic variants called in MDA-MB-468^{r}PCL^{20}.

A11. High confidence somatic variants called in HCC38. Chromosome position, base change from the reference genome to variant alternative and gene affected for all high confidence somatic variants called in HCC38.

A12. High confidence somatic variants called in HCC38^{r}CDDP^{3000}. Chromosome position, base change from the reference genome to variant alternative and gene affected for all high confidence somatic variants called in HCC38^{r}CDDP^{3000}.

A13. High confidence somatic variants called in HCC38^{r}DOX^{40}. Chromosome position, base change from the reference genome to variant alternative and gene affected for all high confidence somatic variants called in HCC38^{r}DOX^{40}.

A14. High confidence somatic variants called in HCC38^{r}ERI^{10}. Chromosome position, base change from the reference genome to variant alternative and gene affected for all high confidence somatic variants called in HCC38^{r}ERI^{10}.

A15. High confidence somatic variants called in HCC38^{r}GEM^{20}. Chromosome position, base change from the reference genome to variant alternative and gene affected for all high confidence somatic variants called in HCC38^{r}GEM^{20}.

A16. High confidence somatic variants called in HCC38^{r}PCL$^{2.5}$. Chromosome position, base change from the reference genome to variant alternative and gene affected for all high confidence somatic variants called in HCC38^{r}PCL$^{2.5}$.

A17. High confidence somatic variants called in HCC1806. Chromosome position, base change from the reference genome to variant alternative and gene affected for all high confidence somatic variants called in HCC1806.

A18. High confidence somatic variants called in HCC1806^{r}CDDP^{500}. Chromosome position, base change from the reference genome to variant alternative and gene affected for all high confidence somatic variants called in HCC1806^{r}CDDP^{500}.

A19. High confidence somatic variants called in HCC1806^{r}DOX$^{12.5}$. Chromosome position, base change from the reference genome to variant alternative and gene affected for all high confidence somatic variants called in HCC1806^{r}DOX$^{12.5}$.

A20. High confidence somatic variants called in HCC1806^{r}ERI^{50}. Chromosome position, base change from the reference genome to variant alternative and gene affected for all high confidence somatic variants called in HCC1806^{r}ERI^{50}.

A21. High confidence somatic variants called in HCC1806^{r}GEM^{20}. Chromosome position, base change from the reference genome to variant alternative and gene affected for all high confidence somatic variants called in HCC1806^{r}GEM^{20}.

A22. High confidence somatic variants called in HCC1806^{r}PCL^{20}. Chromosome position, base change from the reference genome to variant alternative and gene affected for all high confidence somatic variants called in HCC1806^{r}PCL^{20}.

A23. High confidence somatic variants called in HCC1806^{r}5-F^{1500}. Chromosome position, base change from the reference genome to variant alternative and gene affected for all high confidence somatic variants called in HCC1806^{r}5-F^{1500}.

A24. Consequence of variants. The number of the variants for the different consequences for each the chemo-naive and chemo-resistant cell lines

A25. Comparison of variants in chemo-resistant cell line compared to the chemo-naive cell line they are derived. Comparison of variants considered to be *de novo*, *gained*, *not called*, *lost* or *shared* between the chemo-resistant cell line and the chemo-naive cell they are derived from.

A26. Exact variants and mutated genes shared between the cisplatin resistant cell lines. Exact variants and mutated genes shared between each of the three cisplatin resistant cell lines.

A27. Exact variants and mutated genes shared between the doxorubicin resistant cell lines. Exact variants and mutated genes shared between each of the three doxorubicin resistant cell lines.

A28. Exact variants and mutated genes shared between the eribulin resistant cell lines. Exact variants and mutated genes shared between each of the three eribulin resistant cell lines.

A29. Exact variants and mutated genes shared between the gemcitabine resistant cell lines. Exact variants and mutated genes shared between each of the two gemcitabine resistant cell lines.

A30. Exact variants and mutated genes shared between the paclitaxel resistant cell lines. Exact variants and mutated genes shared between each of the three paclitaxel resistant cell lines.

A31. Identified *de novo* variants identified in chemo-resistant cell lines in the candidate approach. Each of the *de novo* variants for each category in the candidate approach. Contains the variant, the overlapping GO terms and the VAF for each of the chemo-resistant cell lines.

A32. Western blot analysis and siRNA knockdown optimisation of TOP2A. Analysis of TOP2A expression in HCC1806 and HCC1806^{r}DOX$^{12.5}$. Optimisation of reverse transfected TOP2A targeted siRNA in HCC1806.

A33. Mutated genes beneficially selected in the chemo-resistant cell lines. All mutated genes which are found to be beneficially selected in the chemo-resistant cell lines. Includes, *de novo* and *gained* variants, as well as *shared* variants with a two-fold increase in VAF.

A34. Mutated genes not beneficially selected in the chemo-resistant cell lines. All mutated genes which are found to be not beneficially selected in the chemo-resistant cell lines. Includes, *not called* and *lost* variants, as well as *shared* variants with a two-fold decrease in VAF.

A35. Comparison of beneficially selected variants to variants in the TCGA. Matched exact or same consequence variants identified as beneficially selected in the chemo-resistant cell lines to variants in the TCGA database. Includes gene expression data, statistical significance and Kaplan Meier analysis.

A36. Comparison of protein truncating variants to the TCGA. Matched exact or same consequence variants identified protein truncating in the chemo-resistant cell lines to variants in the TCGA database. Includes gene expression data, statistical significance and Kaplan Meier analysis.

References

References

Abdelfatah, S. *et al.* (2019) 'MCC1019, a selective inhibitor of the Polo-box domain of Polo-like kinase 1 as novel, potent anticancer candidate', *Acta Pharmaceutica Sinica B*. Chinese Academy of Medical Sciences, 9(5), pp. 1021–1034. doi: 10.1016/j.apsb.2019.02.001.

Adzhubei, I., Jordan, D. M. and Sunyaev, S. R. (2015) *Predicting Functional Effect of Human Missense Mutations Using PolyPhen-2*, *Current Protocols in Human Genetics*. doi: 10.1002/0471142905.hg0720s76.Predicting.

Afify, S. and Seno, M. (2019) 'Conversion of Stem Cells to Cancer Stem Cells: Undercurrent of Cancer Initiation', *Cancers*. MDPI AG, 11(3), p. 345. doi: 10.3390/cancers11030345.

Aherne, G. W. *et al.* (1996) 'Immunoreactive dUMP and TTP pools as an index of thymidylate synthase inhibition; effect of tomudex (ZD1694) and a nonpolyglutamated quinazoline antifolate (CB30900) in L1210 mouse leukaemia cells', *Biochemical Pharmacology*. Elsevier Inc., 51(10), pp. 1293–1301. doi: 10.1016/0006-2952(96)00035-4.

Ahmadzada, T. *et al.* (2018) 'An Update on Predictive Biomarkers for Treatment Selection in Non-Small Cell Lung Cancer', *Journal of Clinical Medicine*, 7(6), p. 153. doi: 10.3390/jcm7060153.

Ahn, J. *et al.* (2018) 'MSK1 functions as a transcriptional coactivator of p53 in the regulation of p21 gene expression', *Experimental and Molecular Medicine*. Nature Publishing Group, 50(10), pp. 1–12. doi: 10.1038/s12276-018-0160-8.

Al-Haj, L., Blackshear, P. J. and Khabar, K. S. A. (2012) 'Regulation of p21/CIP1/WAF-1 mediated cell-cycle arrest by RNase L and tristetraprolin, and involvement of AU-rich elements', *Nucleic Acids Research*, 40(16), pp. 7739–7752. doi: 10.1093/nar/gks545.

Alexander, B. M. *et al.* (2017) 'Individualized screening trial of innovative glioblastoma therapy (INSIGhT).', *Journal of Clinical Oncology*. American Society of Clinical Oncology (ASCO), 35(15_suppl), pp. TPS2079–TPS2079. doi: 10.1200/jco.2017.35.15_suppl.tps2079.

Alexandrov, L. B. *et al.* (2020) 'The repertoire of mutational signatures in human cancer', *Nature*, 578(7793), pp. 94–101. doi: 10.1038/s41586-020-1943-3.

Altmann, K. H. (2001) 'Microtubule-stabilizing agents: A growing class of important anticancer drugs', *Current Opinion in Chemical Biology*. Elsevier Ltd, pp. 424–431. doi: 10.1016/S1367-5931(00)00225-8.

Amlal, H., Burnham, C. E. and Soleimani, M. (1999) 'Characterization of Na+/HCO3/- cotransporter isoform NBC-3', *American Journal of Physiology - Renal Physiology*. Am J Physiol, 276(6 45-6). doi: 10.1152/ajprenal.1999.276.6.F903.

Anderson, V. E. *et al.* (2011a) 'CCT241533 is a potent and selective inhibitor of CHK2 that potentiates the cytotoxicity of PARP inhibitors', *Cancer Research*. American Association for Cancer Research, 71(2), pp. 463–472. doi: 10.1158/0008-5472.CAN-10-1252.

Anderson, V. E. *et al.* (2011b) 'CCT241533 is a potent and selective inhibitor of CHK2 that potentiates the cytotoxicity of PARP inhibitors', *Cancer Research*, 71(2), pp. 463–472. doi: 10.1158/0008-5472.CAN-10-1252.

Andrews S (2018) *FastQC A Quality control tool for high throughput sequence data*, *Babraham Bioinfo*.

Arcaro, A. *et al.* (2000) 'Class II Phosphoinositide 3-Kinases Are Downstream Targets of Activated Polypeptide Growth Factor Receptors', *Molecular and Cellular Biology*. American Society for Microbiology, 20(11), pp. 3817–3830. doi: 10.1128/mcb.20.11.3817-3830.2000.

Arruebo, M. *et al.* (2011) 'Assessment of the Evolution of Cancer Treatment Therapies', *Cancers*, 3, pp. 3279–3330. doi: 10.3390/cancers3033279.

Asad, A. S. *et al.* (2019) 'Prolactin and its receptor as therapeutic targets in glioblastoma multiforme', *Scientific Reports*, 9(1). doi: 10.1038/s41598-019-55860-x.

Asghar, U. S. *et al.* (2017) 'Single-cell dynamics determines response to CDK4/6 inhibition in triple-negative breast cancer', *Clinical Cancer Research*. American Association for Cancer Research Inc., 23(18), pp. 5561–5572. doi: 10.1158/1078-0432.CCR-17-0369.

Atkinson, A. J. *et al.* (2001) 'Biomarkers and surrogate endpoints: Preferred definitions and conceptual framework', *Clinical Pharmacology and Therapeutics*. Clin Pharmacol Ther, pp. 89–95. doi: 10.1067/mcp.2001.113989.

Baba, A. *et al.* (2011) 'PKA-dependent regulation of the histone lysine demethylase complex PHF2-ARID5B', *Nature Cell Biology*. Nat Cell Biol, 13(6), pp. 668–675. doi: 10.1038/ncb2228.

Baker, S. G. (2015) 'A Cancer Theory Kerfuffle Can Lead to New Lines of Research', *JNCI J Natl Cancer Inst*, pp. 1–8. doi: 10.1093/jnci/dju405.

Bamford, S. *et al.* (2004) 'The COSMIC (Catalogue of Somatic Mutations in Cancer) database and website', *British Journal of Cancer*, 2, pp. 355–358. doi: 10.1038/sj.bjc.6601894.

Banerji, U. *et al.* (2019) 'A phase I/II first-in-human trial of oral SRA737 (a Chk1 inhibitor) given in combination with low-dose gemcitabine in subjects with advanced cancer.', *Journal of Clinical Oncology*. American Society of Clinical Oncology (ASCO), 37(15_suppl), pp. 3095–3095. doi: 10.1200/jco.2019.37.15_suppl.3095.

Barber, L. J. *et al.* (2013) 'Secondary mutations in BRCA2 associated with clinical resistance to a PARP inhibitor', *Journal of Pathology*. J Pathol, 229(3), pp. 422–429. doi: 10.1002/path.4140.

Barr, A. R. *et al.* (2017) 'DNA damage during S-phase mediates the proliferation-quiescence decision in the subsequent G1 via p21 expression', *Nature Communications*, 8. doi: 10.1038/ncomms14728.

Barretina, J. *et al.* (2012) 'The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity', *Nature*, 483(7391), pp. 603–607. doi: 10.1038/nature11003.

Barthel, B. L. *et al.* (2009) 'Preclinical efficacy of a carboxylesterase 2-activated prodrug of

doxazolidine', *Journal of Medicinal Chemistry*, 52(23), pp. 7678–7688. doi: 10.1021/jm900694z.

Bash-Imam, Z. *et al.* (2017) 'Translational reprogramming of colorectal cancer cells induced by 5-fluorouracil through a miRNA-dependent mechanism', *Oncotarget*, 8(28), pp. 46219–46233. doi: 10.18632/oncotarget.17597.

Bass, T. E. *et al.* (2016) 'ETAA1 acts at stalled replication forks to maintain genome integrity', *Nature Cell Biology*. Nature Publishing Group, 18(11), pp. 1185–1195. doi: 10.1038/ncb3415.

Bavetsias, V. and Linardopoulos, S. (2015a) 'Aurora kinase inhibitors: Current status and outlook', *Frontiers in Oncology*, 5(DEC). doi: 10.3389/fonc.2015.00278.

Bavetsias, V. and Linardopoulos, S. (2015b) 'Aurora kinase inhibitors: Current status and outlook', *Frontiers in Oncology*, 5(DEC). doi: 10.3389/fonc.2015.00278.

Beauchamp, E. M. *et al.* (2014) 'Acquired resistance to dasatinib in lung cancer cell lines conferred by DDR2 gatekeeper mutation and NF1 loss', *Molecular Cancer Therapeutics*. Mol Cancer Ther, 13(2), pp. 475–482. doi: 10.1158/1535-7163.MCT-13-0817.

Beaufort, C. M. *et al.* (2014) 'Ovarian cancer cell line panel (OCCP): Clinical importance of in vitro morphological subtypes', *PLoS ONE*. Public Library of Science, 9(9). doi: 10.1371/journal.pone.0103988.

Bebbington, D. *et al.* (2009) 'The discovery of the potent aurora inhibitor MK-0457 (VX-680)', *Bioorganic and Medicinal Chemistry Letters*. Elsevier Ltd, 19(13), pp. 3586–3592. doi: 10.1016/j.bmcl.2009.04.136.

Bech-Hansen, N. T., Till, J. E. and Ling, V. (1976) 'Pleiotropic phenotype of colchicine-resistant CHO cells: Cross-resistance and collateral sensitivity', *Journal of Cellular Physiology*, 88(1), pp. 23–31. doi: 10.1002/jcp.1040880104.

Bekker-Jensen, S. *et al.* (2010) 'HERC2 coordinates ubiquitin-dependent assembly of DNA repair factors on damaged chromosomes', *Nature Cell Biology*, 12(1), pp. 80–86. doi: 10.1038/ncb2008.

Bergin, A. R. T. and Loi, S. (2019) 'Triple-negative breast cancer: Recent treatment advances', *F1000Research*, 8. doi: 10.12688/f1000research.18888.1.

Bergman, A. M. *et al.* (2003) 'Increased sensitivity to gemcitabine of P-glycoprotein and multidrug resistance-associated protein-overexpressing human cancer cell lines', *British Journal of Cancer*, 88(12), pp. 1963–1970. doi: 10.1038/sj.bjc.6601011.

Bétermier, M., Bertrand, P. and Lopez, B. S. (2014) 'Is Non-Homologous End-Joining Really an Inherently Error-Prone Process?', *PLoS Genetics*. Edited by S. Jinks-Robertson. Public Library of Science, 10(1), p. e1004086. doi: 10.1371/journal.pgen.1004086.

Bettegowda, C. *et al.* (2014) 'Detection of circulating tumor DNA in early- and late-stage human malignancies', *Science Translational Medicine*. American Association for the Advancement of Science, 6(224). doi: 10.1126/scitranslmed.3007094.

Beucher, A. *et al.* (2009) 'ATM and Artemis promote homologous recombination of radiation-induced DNA double-strand breaks in G2', *The EMBO Journal*. John Wiley & Sons, Ltd, 28(21), pp. 3413–3427. doi: 10.1038/emboj.2009.276.

Bhullar, K. S. *et al.* (2018) 'Kinase-targeted cancer therapies: Progress, challenges and future directions', *Molecular Cancer*, 17(1). doi: 10.1186/s12943-018-0804-2.

Bitler, B. G. *et al.* (2017) 'PARP inhibitors: Clinical utility and possibilities of overcoming resistance', *Gynecologic Oncology*. Academic Press Inc., pp. 695–704. doi: 10.1016/j.ygyno.2017.10.003.

Blackford, A. N. and Jackson, S. P. (2017) 'ATM, ATR, and DNA-PK: The Trinity at the Heart of the DNA Damage Response', *Molecular Cell*. Elsevier Inc., 66(6), pp. 801–817. doi: 10.1016/j.molcel.2017.05.015.

Blackford, A. N. and Stucki, M. (2020) 'How Cells Respond to DNA Breaks in Mitosis', *Trends in Biochemical Sciences*. Elsevier Ltd, 45(4), pp. 321–331. doi: 10.1016/j.tibs.2019.12.010.

Blagosklonny, M. V. (2005) 'Molecular theory of Cancer', *Cancer Biology and Therapy*. Landes Bioscience, pp. 621–627. doi: 10.4161/cbt.4.6.1818.

Bolger, A. M., Lohse, M. and Usadel, B. (2014) 'Trimmomatic: A flexible trimmer for Illumina sequence data', *Bioinformatics*, 30(15), pp. 2114–2120. doi: 10.1093/bioinformatics/btu170.

Borisa, A. C. and Bhatt, H. G. (2017) 'A comprehensive review on Aurora kinase: Small molecule inhibitors and clinical trial studies', *European Journal of Medicinal Chemistry*. Elsevier Masson SAS, pp. 1–19. doi: 10.1016/j.ejmech.2017.08.045.

Bracht, K. *et al.* (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', *Anti-Cancer Drugs*, 17(1), pp. 41–51. doi: 10.1097/01.cad.0000190280.60005.05.

Bracken, C. P. *et al.* (2009) 'Regulation of Cyclin D1 RNA stability by SNIP1', *Cancer Research*, 68(18), pp. 7621–7628. doi: 10.1158/0008-5472.CAN-08-1217.Regulation.

Brakora, K. A. *et al.* (2004) 'Utility of osteopontin as a biomarker in recurrent epithelial ovarian cancer', *Gynecologic Oncology*. Gynecol Oncol, 93(2), pp. 361–365. doi: 10.1016/j.ygyno.2004.01.050.

Brandsma, I. *et al.* (2017) 'Directing the use of DDR kinase inhibitors in cancer treatment', *Expert Opinion on Investigational Drugs*, 26(12), pp. 1341–1355. doi: 10.1080/13543784.2017.1389895.

Branzei, D. and Foiani, M. (2008) 'Regulation of DNA repair throughout the cell cycle', *Nature Reviews Molecular Cell Biology*, pp. 297–308. doi: 10.1038/nrm2351.

Bray, F. *et al.* (2018) 'Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries', *CA: A Cancer Journal for Clinicians*. Wiley, 68(6), pp. 394–424. doi: 10.3322/caac.21492.

Brown, C. W. *et al.* (2000) 'Insertion of Inhbb into the Inhba locus rescues the Inhba-null phenotype and reveals new activin functions', *Nature Genetics*, 25(4), pp. 453–457. doi: 10.1038/78161.

Brown, E. J. and Baltimore, D. (2003) 'Essential and dispensable roles of ATR in cell cycle arrest and genome maintenance', *Genes and Development*. Genes Dev, 17(5), pp. 615–628. doi: 10.1101/gad.1067403.

Budke, B. *et al.* (2016) 'Recent Developments Using Small Molecules to Target RAD51: How to Best Modulate RAD51 for Anticancer Therapy?', *ChemMedChem*, 11(22), pp. 2468–2473. doi: 10.1002/cmdc.201600426.

Burgess, D. J. *et al.* (2008) *Topoisomerase levels determine chemotherapy response in vitro and in vivo*, *Proceedings of the National Academy of Sciences of the United States of America*. doi: 10.1073/pnas.0803513105.

Burma, S. and Chen, D. J. (2004) 'Role of DNA-PK in the cellular response to DNA double-strand breaks', *DNA Repair*. DNA Repair (Amst), pp. 909–918. doi: 10.1016/j.dnarep.2004.03.021.

Burnichon, N. *et al.* (2010) 'SDHA is a tumor suppressor gene causing paraganglioma', *Human Molecular Genetics*, 19(15), pp. 3011–3020. doi: 10.1093/hmg/ddq206.

Burrell, R. A. and Swanton, C. (2014) 'Tumour heterogeneity and the evolution of polyclonal drug resistance', *Molecular Oncology*. John Wiley and Sons Ltd, pp. 1095–1111. doi: 10.1016/j.molonc.2014.06.005.

Burrows, M. and Wheeler, D. (1994) 'A block-sorting lossless data compression algorithm', *Algorithm, Data Compression*, (124), p. 18. doi: 10.1.1.37.6774.

Cai, L. *et al.* (2020) 'Prolonged Time to Adjuvant Chemotherapy Initiation Was Associated with Worse Disease Outcome in Triple Negative Breast Cancer Patients', *Scientific Reports*, 10(1), pp. 1–11. doi: 10.1038/s41598-020-64005-4.

Caldwell, J. J. *et al.* (2011) 'Structure-based design of potent and selective 2-(quinazolin-2-yl)phenol inhibitors of checkpoint kinase 2', *Journal of Medicinal Chemistry*. J Med Chem, 54(2), pp. 580–590. doi: 10.1021/jm101150b.

Campbell, J. *et al.* (2016) 'Large-Scale Profiling of Kinase Dependencies in Cancer Cell Lines', *Cell Reports*. The Authors, 14(10), pp. 2490–2501. doi: 10.1016/j.celrep.2016.02.023.

Campbell, P. J. *et al.* (2020) 'Pan-cancer analysis of whole genomes', *Nature*, 578(7793), pp. 82–93. doi: 10.1038/s41586-020-1969-6.

Cancer Research UK (2016) *Worldwide cancer statistics | Cancer Research UK*, *Cancer Research UK*. Available at: http://www.cancerresearchuk.org/health-professional/cancer-statistics/worldwide-cancer (Accessed: 8 June 2020).

Cantley, L. C. and Neel, B. G. (1999) 'New insights into tumor suppression: PTEN suppresses tumor formation by restraining the phosphoinositide 3-kinase/AKT pathway', *Proceedings of the National*

Academy of Sciences of the United States of America. Proc Natl Acad Sci U S A, pp. 4240–4245. doi: 10.1073/pnas.96.8.4240.

Cantor, S. B. . and Calvo, J. A. (2017) 'Fork Protection and Therapy Resistance in Hereditary Breast Cancer', *Cold Spring Harbor Symp Quant Biol*, 82, pp. 339–348. doi: 10.1016/j.physbeh.2017.03.040.

Carethers, J. M. *et al.* (1999) 'Mismatch repair proficiency and in vitro response to 5-fluorouracil', *Gastroenterology*, 117(1), pp. 123–131. doi: 10.1016/S0016-5085(99)70558-5.

Carey, L. A. *et al.* (2007) 'The triple negative paradox: Primary tumor chemosensitivity of breast cancer subtypes', *Clinical Cancer Research*, 13(8), pp. 2329–2334. doi: 10.1158/1078-0432.CCR-06-1109.

Carpenter, A. J. and Porter, A. C. G. (2004) 'Construction, characterization, and complementation of a conditional-lethal DNA topoisomerase IIα mutant human cell line', *Molecular Biology of the Cell*. Mol Biol Cell, 15(12), pp. 5700–5711. doi: 10.1091/mbc.E04-08-0732.

Catanzaro, D. *et al.* (2015) 'Inhibition of glucose-6-phosphate dehydrogenase sensitizes cisplatin-resistant cells to death', *Oncotarget*. Impact Journals LLC, 6(30), pp. 30102–30114. doi: 10.18632/oncotarget.4945.

Catanzaro, D. *et al.* (2018) 'Cisplatin liposome and 6-amino nicotinamide combination to overcome drug resistance in ovarian cancer cells', *Oncotarget*. Impact Journals LLC, 9(24), pp. 16847–16860. doi: 10.18632/oncotarget.24708.

Cazenave, L. A. *et al.* (1989) 'Glutathione S-transferase and drug resistance.', *Cancer treatment and research*. Cancer Treat Res, pp. 171–187. doi: 10.1007/978-1-4613-1601-5_11.

Cazzaniga, A. *et al.* (2017) 'The different expression of TRPM7 and MagT1 impacts on the proliferation of colon carcinoma cells sensitive or resistant to doxorubicin', *Scientific Reports*. Nature Publishing Group, 7. doi: 10.1038/srep40538.

Chae, Y. K. *et al.* (2017) *Challenges and future of biomarker tests in the era of precision oncology: Can we rely on immunohistochemistry (IHC) or fluorescence in situ hybridization (FISH) to select the optimal patients for matched therapy?*, *Oncotarget*. Available at: www.impactjournals.com/oncotarget (Accessed: 14 October 2020).

Chanvorachote, P. *et al.* (2006) 'Nitric oxide regulates cell sensitivity to cisplatin-induced apoptosis through S-nitrosylation and inhibition of Bcl-2 ubiquitination', *Cancer Research*, 66(12), pp. 6353–6360. doi: 10.1158/0008-5472.CAN-05-4533.

Chapuy, B. *et al.* (2009) 'ABC transporter A3 facilitates lysosomal sequestration of imatinib and modulates susceptibility of chronic myeloid leukemia cell lines to this drug', *Haematologica*. Haematologica, 94(11), pp. 1528–1536. doi: 10.3324/haematol.2009.008631.

Chatrath, A. *et al.* (2020) 'The pan-cancer landscape of prognostic germline variants in 10,582 patients', *Genome Medicine*. BioMed Central Ltd., 12(1), p. 15. doi: 10.1186/s13073-020-0718-7.

Chatterjee, N. and Walker, G. C. (2017) 'Mechanisms of DNA damage, repair, and mutagenesis', *Environmental and Molecular Mutagenesis*, 58(5), pp. 235–263. doi: 10.1002/em.22087.

Chatterjee, S. K. and Zetter, B. R. (2005) 'Cancer biomarkers: knowing the present and predicting the future', *Future Oncology*, 1(1), pp. 37–50. doi: 10.1517/1479-6694.1.1.37.

Chaudhuri, A. R. *et al.* (2016) 'Replication fork stability confers chemoresistance in BRCA-deficient cells', *Nature*. Nature Publishing Group, 535(7612), pp. 382–387. doi: 10.1038/nature18325.

Chen, H. H. W. and Kuo, M. T. (2010) 'Role of Glutathione in the Regulation of Cisplatin Resistance in Cancer Chemotherapy', *Metal-Based Drugs*. Hindawi Publishing Corporation, 2010. doi: 10.1155/2010/430939.

Chen, M. and Zhao, H. (2019) 'Next-generation sequencing in liquid biopsy: cancer screening and early detection', *Human genomics*. NLM (Medline), p. 34. doi: 10.1186/s40246-019-0220-8.

Chen, Z. *et al.* (2016) 'Mammalian drug efflux transporters of the ATP binding cassette (ABC) family in multidrug resistance: A review of the past decade', *Cancer Letters*, 370(1), pp. 153–164. doi: 10.1016/j.canlet.2015.10.010.

Chen, Z., Yang, H. and Pavletich, N. P. (2008) 'Mechanism of homologous recombination from the RecA-ssDNA/dsDNA structures', *Nature*. Nature Publishing Group, 453(7194), pp. 489–494. doi: 10.1038/nature06971.

Chikamori, K. *et al.* (2003) 'Phosphorylation of serine 1106 in the catalytic domain of topoisomerase IIα regulates enzymatic activity and drug sensitivity', *Journal of Biological Chemistry*, 278(15), pp. 12696–12702. doi: 10.1074/jbc.M300837200.

Chk2 Inhibitor Programme | Commercial Partnerships | Cancer Research UK (no date). Available at: http://commercial.cancerresearchuk.org/chk2-inhibitor-programme (Accessed: 22 June 2020).

Choi, Y. H. and Yu, A.-M. (2014) *ABC Transporters in Multidrug Resistance and Pharmacokinetics, and Strategies for Drug Development HHS Public Access*, *Curr Pharm Des*.

Chopra, P. *et al.* (2010) 'Polo-like kinase inhibitors: An emerging opportunity for cancer therapeutics', *Expert Opinion on Investigational Drugs*, 19(1), pp. 27–43. doi: 10.1517/13543780903483191.

Christowitz, C. *et al.* (2019) 'Mechanisms of doxorubicin-induced drug resistance and drug resistant tumour growth in a murine breast tumour model', *BMC Cancer*, 19(1). doi: 10.1186/s12885-019-5939-z.

Chu, D. and Wei, L. (2019) 'Nonsynonymous, synonymous and nonsense mutations in human cancer-related genes undergo stronger purifying selections than expectation', *BMC Cancer*. BioMed Central Ltd., 19(1), p. 359. doi: 10.1186/s12885-019-5572-x.

Church, D. M. *et al.* (2011) 'Modernizing reference genome assemblies', *PLoS Biology*. PLoS Biol, 9(7). doi: 10.1371/journal.pbio.1001091.

Cinatl, J. *et al.* (1999) 'Bovine seminal ribonuclease selectively kills human multidrug-resistant neuroblastoma cells via induction of apoptosis.', *International journal of oncology*. Spandidos Publications, 15(5), pp. 1001–9. doi: 10.3892/ijo.15.5.1001.

Cloyd, J. M. *et al.* (2017) 'Clinical and genetic implications of DNA mismatch repair deficiency in patients with pancreatic ductal adenocarcinoma', *JAMA Surgery*. American Medical Association, 152(11), pp. 1086–1088. doi: 10.1001/jamasurg.2017.2631.

Cohen, H. Y. *et al.* (2004) 'Acetylation of the C terminus of Ku70 by CBP and PCAF controls Bax-mediated apoptosis', *Molecular Cell*. Mol Cell, 13(5), pp. 627–638. doi: 10.1016/S1097-2765(04)00094-2.

Colamonici, O. *et al.* (1994) 'Direct binding to and tyrosine phosphorylation of the alpha subunit of the type I interferon receptor by p135tyk2 tyrosine kinase.', *Molecular and Cellular Biology*. American Society for Microbiology, 14(12), pp. 8133–8142. doi: 10.1128/mcb.14.12.8133.

Colaprico, A. *et al.* (2016) 'TCGAbiolinks: An R/Bioconductor package for integrative analysis of TCGA data', *Nucleic Acids Research*, 44(8), p. e71. doi: 10.1093/nar/gkv1507.

Conti, F. *et al.* (2015) 'Biomarkers for the early diagnosis of bacterial infection and the surveillance of hepatocellular carcinoma in cirrhosis.', *Biomarkers in medicine*. Future Medicine Ltd., 9(12), pp. 1343–51. doi: 10.2217/bmm.15.100.

Cortazar, P. and Geyer, C. E. (2015) 'Pathological Complete Response in Neoadjuvant Treatment of Breast Cancer', *Annals of Surgical Oncology*. Springer New York LLC, 22(5), pp. 1441–1446. doi: 10.1245/s10434-015-4404-8.

Cortes, J. *et al.* (2011) 'Eribulin monotherapy versus treatment of physician's choice in patients with metastatic breast cancer (EMBRACE): A phase 3 open-label randomised study', *The Lancet*. Lancet, 377(9769), pp. 914–923. doi: 10.1016/S0140-6736(11)60070-6.

Cortez, D. *et al.* (2001) 'ATR and ATRIP: Partners in checkpoint signaling', *Science*. American Association for the Advancement of Science, 294(5547), pp. 1713–1716. doi: 10.1126/science.1065521.

Cree, I. A. and Charlton, P. (2017) 'Molecular chess? Hallmarks of anti-cancer drug resistance', *BMC Cancer*, 17(1). doi: 10.1186/s12885-016-2999-1.

Cui, J. *et al.* (2011) 'FBI-1 functions as a novel AR co-repressor in prostate cancer cells', *Cellular and Molecular Life Sciences*. Cell Mol Life Sci, 68(6), pp. 1091–1103. doi: 10.1007/s00018-010-0511-7.

Cui, Q. *et al.* (2019) 'Chk1 inhibitor MK-8776 restores the sensitivity of chemotherapeutics in p-glycoprotein overexpressing cancer cells', *International Journal of Molecular Sciences*, 20(17), pp. 1–12. doi: 10.3390/ijms20174095.

Cunha, S. R. and Mohler, P. J. (2011) 'Ankyrin-based cellular pathways for cardiac ion channel and transporter targeting and regulation', *Seminars in Cell and Developmental Biology*, 22(2), pp. 166–

170. doi: 10.1016/j.semcdb.2010.09.013.

Van Cutsem, E. *et al.* (2009) 'Cetuximab and chemotherapy as initial treatment for metastatic colorectal cancer', *New England Journal of Medicine*. Massachussetts Medical Society, 360(14), pp. 1408–1417. doi: 10.1056/NEJMoa0805019.

Dabholkar, M. *et al.* (1992) 'ERCC1 and ERCC2 expression in malignant tissues from ovarian cancer patients', *Journal of the National Cancer Institute*, 84(19), pp. 1512–1517. doi: 10.1093/jnci/84.19.1512.

Dagogo-Jack, I. and Shaw, A. T. (2018) 'Tumour heterogeneity and resistance to cancer therapies', *Nature Reviews Clinical Oncology*, 15(2), pp. 81–94. doi: 10.1038/nrclinonc.2017.166.

Damerla, R. R. *et al.* (2015) 'Novel Jbts17 mutant mouse model of joubert syndrome with cilia transition zone defects and cerebellar and other ciliopathy related anomalies', *Human Molecular Genetics*, 24(14), pp. 3994–4005. doi: 10.1093/hmg/ddv137.

Dasari, S. and Bernard Tchounwou, P. (2014) 'Cisplatin in cancer therapy: Molecular mechanisms of action', *European Journal of Pharmacology*, 740, pp. 364–378. doi: 10.1016/j.ejphar.2014.07.025.

DeLellis, R., Wolfe, H. and Rule, A. (1987) 'Carcinoembryonic antigen as a tissue marker in medullary thyroid carcinoma', *New England Journal of Medicine*, 299(19), p. 1082.

Delou *et al.* (2019) 'Highlights in Resistance Mechanism Pathways for Combination Therapy', *Cells*, 8(9), p. 1013. doi: 10.3390/cells8091013.

Delou, J. M. A. *et al.* (2019) 'Highlights in Resistance Mechanism Pathways for Combination Therapy', *Cells*. NLM (Medline), p. 1013. doi: 10.3390/cells8091013.

Deng, C. X. (2006) 'BRCA1: Cell cycle checkpoint, genetic instability, DNA damage response and cancer evolution', *Nucleic Acids Research*, 34(5), pp. 1416–1426. doi: 10.1093/nar/gkl010.

Dent, R. *et al.* (2007) 'Triple-negative breast cancer: Clinical features and patterns of recurrence', *Clinical Cancer Research*. Clin Cancer Res, 13(15), pp. 4429–4434. doi: 10.1158/1078-0432.CCR-06-3045.

Derry, W. B., Wilson, L. and Jordan, M. A. (1995) 'Substoichiometric Binding of Taxol Suppresses Microtubule Dynamics', *Biochemistry*. Biochemistry, 34(7), pp. 2203–2211. doi: 10.1021/bi00007a014.

Dev, H. *et al.* (2018) 'Shieldin complex promotes DNA end-joining and counters homologous recombination in BRCA1-null cells', *Nature Cell Biology*, 20(8), pp. 954–965. doi: 10.1038/s41556-018-0140-1.

Dhar, S. *et al.* (2011) 'Targeted delivery of a cisplatin prodrug for safer and more effective prostate cancer therapy in vivo', *Proceedings of the National Academy of Sciences of the United States of America*. Proc Natl Acad Sci U S A, 108(5), pp. 1850–1855. doi: 10.1073/pnas.1011379108.

Diasio, R. B. and Harris, B. E. (1989) 'Clinical Pharmacology of 5-Fluorouracil', *Clinical*

Pharmacokinetics. Clin Pharmacokinet, pp. 215–237. doi: 10.2165/00003088-198916040-00002.

Dong, B. wei *et al.* (2018) 'Transactivation of PTGS2 by PAX5 signaling potentiates cisplatin resistance in muscle-invasive bladder cancer cells', *Biochemical and Biophysical Research Communications*. Elsevier Ltd, 503(4), pp. 2293–2300. doi: 10.1016/j.bbrc.2018.06.151.

Dong, X. *et al.* (2009) 'Significant Associations of Mismatch Repair Gene Polymorphisms With Clinical Outcome of Pancreatic Cancer', *J Clin Oncol*, 27, pp. 1592–1599. doi: 10.1200/JCO.2008.20.1111.

Doroshow, J. H. (1986) 'Role of hydrogen peroxide and hydroxyl radical formation in the killing of Ehrlich tumor cells by anticancer quinones', *Proceedings of the National Academy of Sciences of the United States of America*. Proc Natl Acad Sci U S A, 83(12), pp. 4514–4518. doi: 10.1073/pnas.83.12.4514.

Drilon, A. *et al.* (2018) 'Efficacy of larotrectinib in TRK fusion-positive cancers in adults and children', *New England Journal of Medicine*. Massachussetts Medical Society, 378(8), pp. 731–739. doi: 10.1056/NEJMoa1714448.

Du, B. and Shim, J. S. (2016) 'Targeting epithelial-mesenchymal transition (EMT) to overcome drug resistance in cancer', *Molecules*. MDPI AG. doi: 10.3390/molecules21070965.

Dumontet, C. and Jordan, M. A. (2010) 'Microtubule-binding agents: A dynamic field of cancer therapeutics', *Nature Reviews Drug Discovery*. Nature Publishing Group, pp. 790–803. doi: 10.1038/nrd3253.

Dziadkowiec, K. N. *et al.* (2016) 'PARP inhibitors: review of mechanisms of action and BRCA1/2 mutation targeting.', *Menopause review*, 15(4), pp. 215–219. doi: 10.5114/pm.2016.65667.

Eckerdt, F. (2011) 'Polo-like kinase 1 inhibitors SBE13 and BI 2536 induce different responses in primary cells', *Cell Cycle*. Taylor and Francis Inc., pp. 1031–1030. doi: 10.4161/cc.10.7.15213.

Egawa-Takata, T. *et al.* (2010) 'Early reduction of glucose uptake after cisplatin treatment is a marker of cisplatin sensitivity in ovarian cancer', *Cancer Science*. Blackwell Publishing Ltd, 101(10), pp. 2171–2178. doi: 10.1111/j.1349-7006.2010.01670.x.

Elmore, S. (2007) 'Apoptosis: A Review of Programmed Cell Death', *Toxicologic Pathology*. Toxicol Pathol, pp. 495–516. doi: 10.1080/01926230701320337.

Engeland, K. (2018) 'Cell cycle arrest through indirect transcriptional repression by p53: I have a DREAM', *Cell Death and Differentiation*. Nature Publishing Group, 25(1), pp. 114–132. doi: 10.1038/cdd.2017.172.

Engeli, R. T. *et al.* (2016) 'Biochemical analyses and molecular modeling explain the functional loss of 17β-hydroxysteroid dehydrogenase 3 mutant G133R in three Tunisian patients with 46, XY Disorders of Sex Development', *Journal of Steroid Biochemistry and Molecular Biology*. Elsevier Ltd, 155(Pt A), pp. 147–154. doi: 10.1016/j.jsbmb.2015.10.023.

Evrard, A. *et al.* (1999) 'Increased cytotoxicity and bystander effect of 5-fluorouracil and 5'-deoxy-5-fluorouridine in human colorectal cancer cells transfected with thymidine phosphorylase', *British Journal of Cancer*. Br J Cancer, 80(11), pp. 1726–1733. doi: 10.1038/sj.bjc.6690589.

Falasca, M. and Linton, K. J. (2012) 'Investigational ABC transporter inhibitors', *Expert Opinion on Investigational Drugs*. Taylor & Francis, pp. 657–666. doi: 10.1517/13543784.2012.679339.

Falzone, L., Salomone, S. and Libra, M. (2018) 'Evolution of cancer pharmacological treatments at the turn of the third millennium', *Frontiers in Pharmacology*. Frontiers Media S.A., p. 1300. doi: 10.3389/fphar.2018.01300.

Fang, G., Hongtao, Y. and Kirschner, M. W. (1998) 'Direct binding of CDC20 protein family members activates the anaphase-promoting complex in mitosis and G1', *Molecular Cell*, 2(2), pp. 163–171. doi: 10.1016/S1097-2765(00)80126-4.

Fassett, J. T. *et al.* (2009) 'Adenosine regulation of microtubule dynamics in cardiac hypertrophy', *American Journal of Physiology - Heart and Circulatory Physiology*, 297(2). doi: 10.1152/ajpheart.00462.2009.

FDA alerts health care professionals and oncology clinical investigators about efficacy and potential safety concerns with atezolizumab in combination with paclitaxel for treatment of breast cancer | FDA (no date). Available at: https://www.fda.gov/drugs/drug-safety-and-availability/fda-alerts-health-care-professionals-and-oncology-clinical-investigators-about-efficacy-and (Accessed: 13 October 2020).

FDA Approves First Targeted Therapeutic Based on Tumor Biomarker, Not Tumor Origin - American Association for Cancer Research (AACR) (no date). Available at: https://www.aacr.org/professionals/blog/fda-approves-first-targeted-therapeutic-based-on-tumor-biomarker-not-tumor-origin/ (Accessed: 29 May 2020).

Finn, R. S. *et al.* (2009) 'PD 0332991, a selective cyclin D kinase 4/6 inhibitor, preferentially inhibits proliferation of luminal estrogen receptor-positive human breast cancer cell lines in vitro', *Breast Cancer Research*. Breast Cancer Res, 11(5). doi: 10.1186/bcr2419.

Fletcher, O. and Houlston, R. S. (2010) 'Architecture of inherited susceptibility to common cancer', *Nature Reviews Cancer*. Nat Rev Cancer, pp. 353–361. doi: 10.1038/nrc2840.

Fodale, V. *et al.* (2011) 'Mechanism of cell adaptation: When and how do cancer cells develop chemoresistance?', *Cancer Journal*, 17(2), pp. 89–95. doi: 10.1097/PPO.0b013e318212dd3d.

Foote, K. M. *et al.* (2018) 'Discovery and Characterization of AZD6738, a Potent Inhibitor of Ataxia Telangiectasia Mutated and Rad3 Related (ATR) Kinase with Application as an Anticancer Agent', *Journal of Medicinal Chemistry*. American Chemical Society, 61(22), pp. 9889–9907. doi: 10.1021/acs.jmedchem.8b01187.

Forbes, S. A. *et al.* (2017) 'COSMIC: Somatic cancer genetics at high-resolution', *Nucleic Acids Research*, 45(D1), pp. D777–D783. doi: 10.1093/nar/gkw1121.

Freedman, G. M. *et al.* (2009) 'Locoregional recurrence of triple-negative breast cancer after breast-conserving surgery and radiation', in *Cancer*. NIH Public Access, pp. 946–951. doi: 10.1002/cncr.24094.

Fridman, J. S. and Lowe, S. W. (2003) 'Control of apoptosis by p53', *Oncogene*. Nature Publishing Group, pp. 9030–9040. doi: 10.1038/sj.onc.1207116.

Fry, D. W. *et al.* (2004) 'Specific inhibition of cyclin-dependent kinase 4/6 by PD 0332991 and associated antitumor activity in human tumor xenografts', *Molecular Cancer Therapeutics*, 3(11), pp. 1427–1437.

Fujibayashi, A. *et al.* (2008) 'Human RME-8 is involved in membrane trafficking through early endosomes', *Cell Structure and Function*. Cell Struct Funct, 33(1), pp. 35–50. doi: 10.1247/csf.07045.

Fujita, M. *et al.* (2012) 'Cross-talk between integrin α6β4 and Insulin-like Growth Factor-1 Receptor (IGF1R) through direct α6β4 binding to IGF1 and subsequent α6β4-IGF1-IGF1R ternary complex formation in anchorage-independent conditions', *Journal of Biological Chemistry*. J Biol Chem, 287(15), pp. 12491–12500. doi: 10.1074/jbc.M111.304170.

Fujita, Y. *et al.* (2010) 'MiR-148a attenuates paclitaxel resistance of hormone-refractory, drug-resistant prostate cancer PC3 cells by regulating MSK1 expression', *Journal of Biological Chemistry*. J Biol Chem, 285(25), pp. 19076–19084. doi: 10.1074/jbc.M109.079525.

Funderburgh, J. L. *et al.* (2001) 'Proteoglycan Expression during Transforming Growth Factor β-induced Keratocyte-Myofibroblast Transdifferentiation', *Journal of Biological Chemistry*. J Biol Chem, 276(47), pp. 44173–44178. doi: 10.1074/jbc.M107596200.

Fyfe, J. C. *et al.* (2004) 'The functional cobalamin (vitamin B12)-intrinsic factor receptor is a novel complex of cubilin and amnionless', *Blood*. Blood, 103(5), pp. 1573–1579. doi: 10.1182/blood-2003-08-2852.

Gallagher, D. *et al.* (2020) 'A Rad51-Independent Pathway Promotes Single-Strand Template Repair in Gene Editing', *bioRxiv*. Cold Spring Harbor Laboratory, p. 2020.02.24.962423. doi: 10.1101/2020.02.24.962423.

Ganapathi, R. N. and Ganapathi, M. K. (2013) 'Mechanisms regulating resistance to inhibitors of topoisomerase II', *Frontiers in Pharmacology*. Frontiers Media S.A., 4, p. 89. doi: 10.3389/fphar.2013.00089.

Gao, H. *et al.* (2019) 'Omeprazole protects against cisplatin-induced nephrotoxicity by alleviating oxidative stress, inflammation, and transporter-mediated cisplatin accumulation in rats and HK-2 cells', *Chemico-Biological Interactions*. Elsevier Ireland Ltd, 297, pp. 130–140. doi: 10.1016/j.cbi.2018.11.008.

Gao, Z. *et al.* (2012) 'Determination of protein interactome of transcription factor Sox2 in embryonic stem cells engineered for inducible expression of four reprogramming factors', *Journal of Biological*

Chemistry, 287(14), pp. 11384–11397. doi: 10.1074/jbc.M111.320143.

Gareau, J. R., Reverter, D. and Lima, C. D. (2012) 'Determinants of small ubiquitin-like modifier 1 (SUMO1) protein specificity, E3 ligase, and SUMO-RanGAP1 binding activities of nucleoporin RanBP2', *Journal of Biological Chemistry*, 287(7), pp. 4740–4751. doi: 10.1074/jbc.M111.321141.

Garnett, M. J. *et al.* (2012a) 'Systematic identification of genomic markers of drug sensitivity in cancer cells', *Nature*. Nature, 483(7391), pp. 570–575. doi: 10.1038/nature11005.

Garnett, M. J. *et al.* (2012b) 'Systematic identification of genomic markers of drug sensitivity in cancer cells', *Nature*. Nature Publishing Group, 483(7391), pp. 570–575. doi: 10.1038/nature11005.

Garraway, L. A. and Jänne, P. A. (2012a) 'Circumventing cancer drug resistance in the era of personalized medicine', *Cancer Discovery*, 2(3), pp. 214–226. doi: 10.1158/2159-8290.CD-12-0012.

Garraway, L. A. and Jänne, P. A. (2012b) 'Circumventing cancer drug resistance in the era of personalized medicine', *Cancer Discovery*. American Association for Cancer Research, pp. 214–226. doi: 10.1158/2159-8290.CD-12-0012.

Garrett, M. D. and Collins, I. (2011) 'Anticancer therapy with checkpoint inhibitors: What, where and when?', *Trends in Pharmacological Sciences*. Elsevier Ltd, 32(5), pp. 308–316. doi: 10.1016/j.tips.2011.02.014.

Garrido, M. F. *et al.* (2019) 'Regulation of eIF4F Translation Initiation Complex by the Peptidyl Prolyl Isomerase FKBP7 in Taxane-resistant Prostate Cancer', *Clinical Cancer Research*, 25(2), pp. 710–723. doi: 10.1158/1078-0432.CCR-18-0704.

Gartel, A. L. (2006) 'Is p21 an oncogene?', *Molecular cancer therapeutics*. American Association for Cancer Research, 5(6), pp. 1385–6. doi: 10.1158/1535-7163.MCT-06-0163.

Ghandi, M. *et al.* (2019) 'Next-generation characterization of the Cancer Cell Line Encyclopedia', *Nature*. doi: 10.1038/s41586-019-1186-3.

Giampazolias, E. and Tait, S. W. G. (2016) 'Mitochondria and the hallmarks of cancer', *FEBS Journal*. Blackwell Publishing Ltd, pp. 803–814. doi: 10.1111/febs.13603.

Gohda, J. *et al.* (2018) 'BI-2536 and BI-6727, dual Polo-like kinase/bromodomain inhibitors, effectively reactivate latent HIV-1', *Scientific Reports*. Springer US, 8(1), pp. 1–13. doi: 10.1038/s41598-018-21942-5.

Goldstein, M. and Kastan, M. B. (2015) 'The DNA damage response: Implications for tumor responses to radiation and chemotherapy', *Annual Review of Medicine*. Annual Reviews Inc., 66, pp. 129–143. doi: 10.1146/annurev-med-081313-121208.

Goossens, N. *et al.* (2015) 'Cancer biomarker discovery and validation', *Translational Cancer Research*, 4(3), pp. 256–269. doi: 10.3978/j.issn.2218-676X.2015.06.04.

Gorecki, L. *et al.* (2020) 'Discovery of ATR kinase inhibitor berzosertib (VX-970, M6620): Clinical candidate for cancer therapy', *Pharmacology and Therapeutics*. Elsevier Inc., p. 107518. doi:

10.1016/j.pharmthera.2020.107518.

Gottesman, M. M. (2002) 'Mechanisms of Cancer Drug Resistance', *Annual Review of Medicine*, 53, pp. 615–627.

Gottlieb, T. M. and Jackson, S. P. (1993) 'The DNA-dependent protein kinase: Requirement for DNA ends and association with Ku antigen', *Cell*. Cell, 72(1), pp. 131–142. doi: 10.1016/0092-8674(93)90057-W.

Graham, T. G. W., Walter, J. C. and Loparo, J. J. (2016) 'Two-Stage Synapsis of DNA Ends during Non-homologous End Joining', *Molecular Cell*. Elsevier Inc., 61(6), pp. 850–858. doi: 10.1016/j.molcel.2016.02.010.

Grandér, D. and Grandér, G. (1998) 'How do mutated oncogenes and tumor suppressor genes cause cancer?', *Medical Oncology*, 15(1), pp. 20–26. doi: 10.1007/BF02787340.

Greten, F. R. *et al.* (2002) 'Stat3 and NF-κB activation prevents apoptosis in pancreatic carcinogenesis', *Gastroenterology*. W.B. Saunders, 123(6), pp. 2052–2063. doi: 10.1053/gast.2002.37075.

Grichtchenko, I. *et al.* (2001) 'Cloning, Characterization, and Chromosomal Mapping of a Human Electroneutral Na+-driven Cl-HCO3 Exchanger', *Journal of Biological Chemistry*. J Biol Chem, 276(11), pp. 8358–8363. doi: 10.1074/jbc.C000716200.

Grossman, R. L. *et al.* (2016) 'Toward a shared vision for cancer genomic data', *New England Journal of Medicine*. Massachussetts Medical Society, pp. 1109–1112. doi: 10.1056/NEJMp1607591.

Group, I. C. G. C. (2010) 'International network of cancer genome projects', *Nature*, 15(464), pp. 993–8. doi: 10.1016/j.physbeh.2017.03.040.

Grozav, A. G. *et al.* (2009) 'Casein kinase I delta;/ε phosphorylates topoisomerase IIα at serine-1106 and modulates DNA cleavage activity', *Nucleic Acids Research*, 37(2), pp. 382–392. doi: 10.1093/nar/gkn934.

Gu, M. *et al.* (2010) 'Cloning and characterization of a new BRCA1 variant: A role for BRCT domains in apoptosis', *Cancer Letters*. Cancer Lett, 295(2), pp. 205–213. doi: 10.1016/j.canlet.2010.03.002.

Guertin, A. D. *et al.* (2013) 'Preclinical evaluation of the WEE1 inhibitor MK-1775 as single-agent anticancer therapy', *Molecular Cancer Therapeutics*, 12(8), pp. 1442–1452. doi: 10.1158/1535-7163.MCT-13-0025.

Haahr, P. *et al.* (2016) 'Activation of the ATR kinase by the RPA-binding protein ETAA1', *Nature Cell Biology*. Nature Publishing Group, 18(11), pp. 1196–1207. doi: 10.1038/ncb3422.

Hall, J. R. *et al.* (2007) 'Cdc6 stability is regulated by the Huwe1 ubiquitin ligase after DNA damage', *Molecular Biology of the Cell*, 18(September), pp. 3250–3263. doi: 10.1091/mbc.E07.

Hall, M. D., Handley, M. D. and Gottesman, M. M. (2009) 'Is Resistance Useless? Multidrug Resistance and Collateral Sensitvity', *Cancer*, 30(10), pp. 546–556. doi:

10.1016/j.tips.2009.07.003.Is.

Han, S. *et al.* (2012) 'Nuclear envelope phosphatase 1-regulatory subunit 1 (formerly TMEM188) is the metazoan Spo7p ortholog and functions in the lipin activation pathway', *Journal of Biological Chemistry*. J Biol Chem, 287(5), pp. 3123–3137. doi: 10.1074/jbc.M111.324350.

Hanahan, D. and Weinberg, R. A. (2000) *The hallmarks of cancer*, *Cell*. doi: 10.1016/S0092-8674(00)81683-9.

Hanahan, D. and Weinberg, R. A. (2011) 'Hallmarks of cancer: the next generation.', *Cell*. Elsevier, 144(5), pp. 646–74. doi: 10.1016/j.cell.2011.02.013.

Harkin, D. P. *et al.* (1999) 'Induction of GADD45 and JNK/SAPK-dependent apoptosis following inducible expression of BRCA1', *Cell*. Cell Press, 97(5), pp. 575–586. doi: 10.1016/S0092-8674(00)80769-2.

Harris, J. L., Rabellino, A. and Khanna, K. K. (2018) 'RAD51 paralogs promote genomic integrity and chemoresistance in cancer by facilitating homologous recombination', *Annals of Translational Medicine*. AME Publishing Company, 6(S2), pp. S122–S122. doi: 10.21037/atm.2018.12.30.

Harvey, J. T. *et al.* (2002) 'An analysis of the forces required to drag sheep over various surfaces', *Applied Ergonomics*. Elsevier Ltd, 33(6), pp. 523–531. doi: 10.1016/S0003-6870(02)00071-6.

Hayashi, F. *et al.* (2001) 'The innate immune response to bacterial flagellin is mediated by Toll-like receptor 5', *Nature*. Nature, 410(6832), pp. 1099–1103. doi: 10.1038/35074106.

He, X. *et al.* (2017) 'IFN-I 3 regulates human dental pulp stem cells behavior via NF-I B and MAPK signaling', *Scientific Reports*. Nature Publishing Group, 7, pp. 1–13. doi: 10.1038/srep40681.

Heidelberger, C. *et al.* (1957) 'Fluorinated pyrimidines, a new class of tumour-inhibitory compounds', *Nature*. Nature Publishing Group, 179(4561), pp. 663–666. doi: 10.1038/179663a0.

Heidrich, I. *et al.* (2020) 'Liquid biopsies: Potential and challenges', *International Journal of Cancer*. Wiley-Liss Inc., p. ijc.33217. doi: 10.1002/ijc.33217.

Henry, N. L. and Hayes, D. F. (2012) 'Cancer biomarkers', *Molecular Oncology*. Elsevier B.V, 6(2), pp. 140–146. doi: 10.1016/j.molonc.2012.01.010.

Hertweck, K. L. and Dasgupta, S. (2017) 'The landscape of mtDNA modifications in cancer: A tale of two cities', *Frontiers in Oncology*. Frontiers Media S.A., p. 262. doi: 10.3389/fonc.2017.00262.

Herwig, R. *et al.* (2011) 'Targeted high throughput sequencing in clinical cancer Settings: formaldehyde fixed-paraffin embedded (FFPE) tumor tissues, input amount and tumor heterogeneity', *BMC Medical Genomics*. BioMed Central Ltd, 4(1), p. 68. doi: 10.1186/1755-8794-4-68.

Heyer, W.-D., Ehmsen, K. T. and Liu, J. (2010) 'Regulation of Homologous Recombination in Eukaryotes', *Annual Review of Genetics*. Annual Reviews, 44(1), pp. 113–139. doi: 10.1146/annurev-genet-051710-150955.

Hoffner, G., Kahlem, P. and Djian, P. (2002) 'Perinuclear localization of huntingtin as a consequence of its binding to microtubules through an interaction with β-tubulin: Relevance to Huntington's disease', *Journal of Cell Science*, 115(5), pp. 941–948.

Holohan, C. *et al.* (2013) 'Cancer drug resistance: An evolving paradigm', *Nature Reviews Cancer*, 13(10), pp. 714–726. doi: 10.1038/nrc3599.

Houghton, J. A. and Houghton, P. J. (1983) 'Elucidation of pathways of 5-fluorouracil metabolism in xenografts of human colorectal adenocarcinoma', *European Journal of Cancer and Clinical Oncology*. Eur J Cancer Clin Oncol, 19(6), pp. 807–815. doi: 10.1016/0277-5379(83)90013-5.

Housman, G. *et al.* (2014) 'Drug Resistance in Cancer: An Overview', *Cancers*, 6, pp. 1769–1792. doi: 10.3390/cancers6031769.

Hsu, P. D., Lander, E. S. and Zhang, F. (2014) 'Development and applications of CRISPR-Cas9 for genome engineering', *Cell*. Cell Press, pp. 1262–1278. doi: 10.1016/j.cell.2014.05.010.

Hsu, W. H. *et al.* (2019) 'Checkpoint Kinase 1 Inhibition Enhances Cisplatin Cytotoxicity and Overcomes Cisplatin Resistance in SCLC by Promoting Mitotic Cell Death', *Journal of Thoracic Oncology*. Elsevier Inc, 14(6), pp. 1032–1045. doi: 10.1016/j.jtho.2019.01.028.

Huang, C. Y. *et al.* (2017) 'A review on the effects of current chemotherapy drugs and natural agents in treating non-small cell lung cancer', *BioMedicine (France)*, 7(4), pp. 12–23. doi: 10.1051/bmdcn/2017070423.

Huang, F. *et al.* (2012) 'Inhibition of homologous recombination in human cells by targeting RAD51 recombinase', *Journal of Medicinal Chemistry*, 55(7), pp. 3011–3020. doi: 10.1021/jm201173g.

Huang, F. and Mazin, A. V (2011) 'Identification of specific inhibitors of human RAD51 recombinase using high-throughput screening', *ACS Chem Biol.*, 6(6), pp. 628–635. doi: 10.1021/cb100428c.Identification.

Hustedt, N. and Durocher, D. (2017) *The control of DNA repair by the cell cycle*, *Nature Cell Biology*. doi: 10.1038/ncb3452.

Jackson, S. P. and Bartek, J. (2009) 'The DNA-damage response in human biology and disease', *Nature*, 461(7267), pp. 1071–1078. doi: 10.1038/nature08467.

Jahani, M. *et al.* (2017) 'L-arginine alters the effect of 5-fluorouracil on breast cancer cells in favor of apoptosis', *Biomedicine and Pharmacotherapy*. Elsevier Masson SAS, 88, pp. 114–123. doi: 10.1016/j.biopha.2017.01.047.

Jeon, B. N. *et al.* (2014) 'Two ZNF509 (ZBTB49) isoforms induce cell-cycle arrest by activating transcription of p21/CDKN1A and RB upon exposure to genotoxic stress', *Nucleic Acids Research*, 42(18), pp. 11447–11461. doi: 10.1093/nar/gku857.

Jeong, W. J. *et al.* (2019) 'WDR76 is a RAS binding protein that functions as a tumor suppressor via RAS degradation', *Nature Communications*. Springer US, 10(1). doi: 10.1038/s41467-018-08230-6.

Jeyshka, R.-G. *et al.* (2020) 'Downstream Effectors of ILK in CIsplatin-Resistant Ovarian Cancer', *Cancers*, 12(880). doi: 10.1016/S0140-6736(09)61338-6.

Johannessen, C. M. and Boehm, J. S. (2017) 'Progress towards precision functional genomics in cancer', *Current Opinion in Systems Biology*. Elsevier Ltd, pp. 74–83. doi: 10.1016/j.coisb.2017.02.002.

Jordan, M. A. *et al.* (2005) 'The primary antimitotic mechanism of action of the synthetic halichondrin E7389 is suppression of microtubule growth', *Molecular Cancer Therapeutics*. Mol Cancer Ther, 4(7), pp. 1086–1095. doi: 10.1158/1535-7163.MCT-04-0345.

Jordan, M. A. and Wilson, L. (2004) 'Microtubules as a target for anticancer drugs', *Nature Reviews Cancer*. Nature Publishing Group, pp. 253–265. doi: 10.1038/nrc1317.

Jordan, M. and Kamath, K. (2007) 'How Do Microtubule-Targeted Drugs Work? An Overview', *Current Cancer Drug Targets*. Bentham Science Publishers Ltd., 7(8), pp. 730–742. doi: 10.2174/156800907783220417.

Jung, J. *et al.* (2016) 'TP53 mutations emerge with HDM2 inhibitor SAR405838 treatment in de-differentiated liposarcoma', *Nature Communications*, 7, pp. 1–7. doi: 10.1038/ncomms12609.

Kahata, K., Dadras, M. S. and Moustakas, A. (2018) 'TGF-β family signaling in epithelial differentiation and epithelial-mesenchymal transition', *Cold Spring Harbor Perspectives in Biology*. Cold Spring Harbor Laboratory Press, 10(1). doi: 10.1101/cshperspect.a022194.

Kalimutho, M. *et al.* (2019) 'Patterns of Genomic Instability in Breast Cancer', *Trends in Pharmacological Sciences*. Elsevier Ltd, pp. 198–211. doi: 10.1016/j.tips.2019.01.005.

Kanehisa, M. (2000) 'KEGG: Kyoto Encyclopedia of Genes and Genomes', *Nucleic Acids Research*. Oxford University Press (OUP), 28(1), pp. 27–30. doi: 10.1093/nar/28.1.27.

Kanehisa, M. *et al.* (2019) 'New approach for understanding genome variations in KEGG', *Nucleic Acids Research*. Oxford University Press, 47(D1), pp. D590–D595. doi: 10.1093/nar/gky962.

Kang, S. *et al.* (2006) 'Association between excision repair cross-complementation group 1 polymorphism and clinical outcome of platinum-based chemotherapy in patients with epithelial ovarian cancer', *Experimental and Molecular Medicine*. Korean Society of Med. Biochemistry and Mol. Biology, 38(3), pp. 320–324. doi: 10.1038/emm.2006.38.

Karanam, K. *et al.* (2012) 'Quantitative Live Cell Imaging Reveals a Gradual Shift between DNA Repair Mechanisms and a Maximal Use of HR in Mid S Phase', *Molecular Cell*. Cell Press, 47(2), pp. 320–329. doi: 10.1016/j.molcel.2012.05.052.

Karczewski, K. J. *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*. Cold Spring Harbor Laboratory, p. 531210. doi: 10.1101/531210.

Karthikeyan, S. and Hoti, S. (2015) 'Development of Fourth Generation ABC Inhibitors from Natural

Products: A Novel Approach to Overcome Cancer Multidrug Resistance', *Anti-Cancer Agents in Medicinal Chemistry*. Bentham Science Publishers Ltd., 15(5), pp. 605–615. doi: 10.2174/1871520615666150113103439.

Kastan, M. B. and Bartek, J. (2004) 'Cell-cycle checkpoints and cancer', *Nature*, 432(7015), pp. 316–323. doi: 10.1038/nature03097.

Katano, K. *et al.* (2004) 'Confocal microscopic analysis of the interaction between cisplatin and copper transporter ATP7B in human ovarian carcinoma cells', *Clinical Cancer Research*, 10(13), pp. 4578–4588. doi: 10.1158/1078-0432.CCR-03-0689.

Kausar, T. *et al.* (2015) 'Sensitization of Pancreatic Cancers to Gemcitabine Chemoradiation by WEE1 Kinase Inhibition Depends on Homologous Recombination Repair', *Neoplasia*. Elsevier Inc., 17(10), pp. 757–766. doi: 10.1016/j.neo.2015.09.006.

Kavallaris, M. (2010) 'Microtubules and resistance to tubulin-binding agents', *Nature Reviews Cancer*, pp. 194–204. doi: 10.1038/nrc2803.

Kawakami, K. *et al.* (2015) 'Integrin β4 and vinculin contained in exosomes are potential markers for progression of prostate cancer associated with taxane-resistance', *International Journal of Oncology*. Spandidos Publications, 47(1), pp. 384–390. doi: 10.3892/ijo.2015.3011.

Keenan, T. E. and Tolaney, S. M. (2020) 'Role of immunotherapy in triple-negative breast cancer', *JNCCN Journal of the National Comprehensive Cancer Network*. Harborside Press, pp. 479–489. doi: 10.6004/jnccn.2020.7554.

Keppner, S. *et al.* (2009) 'Identification and Validation of a Potent Type II Inhibitor of Inactive Polo-like Kinase 1', *ChemMedChem*. John Wiley & Sons, Ltd, 4(11), pp. 1806–1809. doi: 10.1002/cmdc.200900338.

Khrapko, K. *et al.* (1997) 'Mitochondrial mutational spectra in human cells and tissues', *Proceedings of the National Academy of Sciences of the United States of America*. Proc Natl Acad Sci U S A, 94(25), pp. 13798–13803. doi: 10.1073/pnas.94.25.13798.

Kim, C. *et al.* (2018) 'Chemoresistance Evolution in Triple-Negative Breast Cancer Delineated by Single-Cell Sequencing', *Cell*. Cell Press, 173(4), pp. 879-893.e13. doi: 10.1016/j.cell.2018.03.041.

Kim, H. K. *et al.* (2011) 'A Gene Expression Signature of Acquired Chemoresistance to Cisplatin and Fluorouracil Combination Chemotherapy in Gastric Cancer Patients', *PLoS ONE*. Edited by A. Navarro. Public Library of Science, 6(2), p. e16694. doi: 10.1371/journal.pone.0016694.

Kim, S. Y. *et al.* (2006) 'Doxorubicin-induced reactive oxygen species generation and intracellular Ca2+ increase are reciprocally modulated in rat cardiomyocytes', *Experimental and Molecular Medicine*. Korean Society of Med. Biochemistry and Mol. Biology, 38(5), pp. 535–545. doi: 10.1038/emm.2006.63.

King, C. *et al.* (2014) 'Characterization and preclinical development of LY2603618: A selective and potent Chk1 inhibitor', *Investigational New Drugs*, 32(2), pp. 213–226. doi: 10.1007/s10637-013-

0036-7.

King, C. *et al.* (2015) 'LY2606368 causes replication catastrophe and anti-tumor effects through CHK1-dependent mechanisms.', *Molecular Cancer Therapeutics*, 14(September), pp. 2004–2014. doi: 10.1158/1535-7163.MCT-14-1037.

Koboldt, D. C. *et al.* (2012) 'Comprehensive molecular portraits of human breast tumours', *Nature*, 490(7418), pp. 61–70. doi: 10.1038/nature11412.

Koch, T. (no date) 'Carboxylesterase-activated Doxazolidine-prodrug for Hepatocellular Carcinoma'.

Kodani, A. *et al.* (2015) 'Centriolar satellites assemble centrosomal microcephaly proteins to recruit CDK2 and promote centriole duplication', *eLife*, 4(AUGUST2015), pp. 1–27. doi: 10.7554/eLife.07519.

Koh, S.-B. *et al.* (2018) 'Translational Science Mechanistic Distinctions between CHK1 and WEE1 Inhibition Guide the Scheduling of Triple Therapy with Gemcitabine', *Cancer Res*, 78(11), pp. 3054–66. doi: 10.1158/0008-5472.CAN-17-3932.

Kojouharov, B. M. *et al.* (2014) 'Toll-like receptor-5 agonist entolimod broadens the therapeutic window of 5-fluorouracil by reducing its toxicity to normal tissues in mice', *Oncotarget*. Impact Journals LLC, 5(3), pp. 802–814. doi: 10.18632/oncotarget.1773.

Kolinjivadi, A. M. *et al.* (2017) 'Smarcal1-Mediated Fork Reversal Triggers Mre11-Dependent Degradation of Nascent DNA in the Absence of Brca2 and Stable Rad51 Nucleofilaments', *Molecular Cell*, 67(5), pp. 867-881.e7. doi: 10.1016/j.molcel.2017.07.001.

Konieczkowski, D. J., Johannessen, C. M. and Garraway, L. A. (2018) 'Cancer Cell Perspective A Convergence-Based Framework for Cancer Drug Resistance', *Cancer Cell*, 33, pp. 801–815. doi: 10.1016/j.ccell.2018.03.025.

KÖRBER, M. I. *et al.* (2016) 'NFκB-Associated Pathways in Progression of Chemoresistance to 5-Fluorouracil in an In Vitro Model of Colonic Carcinoma', *Anticancer Research*. International Institute of Anticancer Research, 36(4), pp. 1631–1639.

Kozyraki, R. *et al.* (1999) 'The intrinsic factor-vitamin B12 receptor, cubilin, is a high-affinity apolipoprotein A-I receptor facilitating endocytosis of high-density lipoprotein', *Nature Medicine*. Nat Med, 5(6), pp. 656–661. doi: 10.1038/9504.

Kozyraki, R. *et al.* (2001) 'Megalin-dependent cubilin-mediated endocytosis is a major pathway for the apical uptake of transferrin in polarized epithelia', *Proceedings of the National Academy of Sciences of the United States of America*. Proc Natl Acad Sci U S A, 98(22), pp. 12491–12496. doi: 10.1073/pnas.211291398.

Kroemer, G. *et al.* (2009) 'Classification of cell death: Recommendations of the Nomenclature Committee on Cell Death 2009', *Cell Death and Differentiation*. Cell Death Differ, pp. 3–11. doi: 10.1038/cdd.2008.150.

Kumagai, A. *et al.* (2006) 'TopBP1 activates the ATR-ATRIP complex', *Cell*. Elsevier, 124(5), pp. 943–955. doi: 10.1016/j.cell.2005.12.041.

Kumar, S. and Kim, J. (2015) 'PLK-1 Targeted Inhibitors and Their Potential against Tumorigenesis', *BioMed Research International*. Hindawi Publishing Corporation, 2015. doi: 10.1155/2015/705745.

Lachmann, A. *et al.* (2010) 'ChEA: Transcription factor regulation inferred from integrating genome-wide ChIP-X experiments', *Bioinformatics*, 26(19), pp. 2438–2444. doi: 10.1093/bioinformatics/btq466.

Lafranchi, L. *et al.* (2014) ' APC / C C dh1 controls Ct IP stability during the cell cycle and in response to DNA damage ', *The EMBO Journal*, 33(23), pp. 2860–2879. doi: 10.15252/embj.201489017.

Landrum, M. J. *et al.* (2014) 'ClinVar: Public archive of relationships among sequence variation and human phenotype', *Nucleic Acids Research*, 42(D1), pp. 980–985. doi: 10.1093/nar/gkt1113.

Larsen, A. K. and Skladanowski, A. (1998) 'Cellular resistance to topoisomerase-targeted drugs: From drug uptake to cell death', *Biochimica et Biophysica Acta - Gene Structure and Expression*. Biochim Biophys Acta, pp. 257–274. doi: 10.1016/S0167-4781(98)00140-7.

Laulier, C. *et al.* (2011) 'Bcl-2 inhibits nuclear homologous recombination by localizing BRCA1 to the endomembranes', *Cancer Research*. American Association for Cancer Research Inc., 71(10), pp. 3590–3602. doi: 10.1158/0008-5472.CAN-10-3119.

Lawrence, M. S. *et al.* (2014) 'Discovery and saturation analysis of cancer genes across 21 tumour types', *Nature*. Nature Publishing Group, 505(7484), pp. 495–501. doi: 10.1038/nature12912.

Lee, A. and Djamgoz, M. B. A. (2018) 'Triple negative breast cancer: Emerging therapeutic modalities and novel combination therapies', *Cancer Treatment Reviews*, 62, pp. 110–122. doi: 10.1016/j.ctrv.2017.11.003.

Lee, J. *et al.* (2018) 'Mutalisk: A web-based somatic MUTation AnaLyIS toolKit for genomic, transcriptional and epigenomic signatures', *Nucleic Acids Research*. Oxford University Press, 46(W1), pp. W102–W108. doi: 10.1093/nar/gky406.

Lee, J. H. and Paull, T. T. (2004) 'Direct Activation of the ATM Protein Kinase by the Mre11/Rad50/Nbs1 Complex', *Science*, 304(5667), pp. 93–96. doi: 10.1126/science.1091496.

Lee, J. H. and Paull, T. T. (2005) 'ATM activation by DNA double-strand breaks through the Mre11-Rad50-Nbs1 complex', *Science*. Science, 308(5721), pp. 551–554. doi: 10.1126/science.1108297.

Lee, K. S. *et al.* (2015) 'Recent Advances and New Strategies in Targeting Plk1 for Anticancer Therapy', *Trends in Pharmacological Sciences*. Elsevier Ltd, pp. 858–877. doi: 10.1016/j.tips.2015.08.013.

Lehmann, B. D. *et al.* (2011) 'Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies', *Journal of Clinical Investigation*. J Clin Invest, 121(7), pp. 2750–2767. doi: 10.1172/JCI45014.

Li, D. Q. *et al.* (2013) 'Metastasis-associated protein 1 is an integral component of the circadian molecular machinery', *Nature Communications*. Nature Publishing Group, 4. doi: 10.1038/ncomms3545.

Li, G.-M. (2008) 'Mechanisms and functions of DNA mismatch repair', *Cell Research*, 18, pp. 85–98. doi: 10.1038/cr.2007.115.

Li, H. *et al.* (2009) 'The Sequence Alignment/Map format and SAMtools', *Bioinformatics*, 25(16), pp. 2078–2079. doi: 10.1093/bioinformatics/btp352.

Li, H. (2011) 'A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data', *Bioinformatics*. Oxford University Press, 27(21), pp. 2987–2993. doi: 10.1093/bioinformatics/btr509.

Li, L. *et al.* (2017) 'Long noncoding RNA SFTA1P promoted apoptosis and increased cisplatin chemosensitivity via regulating the hnRNP-U-GADD45A axis in lung squamous cell carcinoma', *Oncotarget*, 8(57), pp. 97476–97489. doi: 10.18632/oncotarget.22138.

Li, Q. *et al.* (2018) 'NF-κB in pancreatic cancer: Its key role in chemoresistance', *Cancer Letters*. Elsevier Ireland Ltd, pp. 127–134. doi: 10.1016/j.canlet.2018.02.011.

Li, S. *et al.* (2016) 'Protein tyrosine phosphatase PTPN3 promotes drug resistance and stem cell-like characteristics in ovarian cancer', *Scientific Reports*. Nature Publishing Group, 6(1), pp. 1–10. doi: 10.1038/srep36873.

Li, Y. *et al.* (2016) 'Hyperoside induces apoptosis and inhibits growth in pancreatic cancer via Bcl-2 family and NF-κB signaling pathway both in vitro and in vivo', *Tumor Biology*. Springer Netherlands, 37(6), pp. 7345–7355. doi: 10.1007/s13277-015-4552-2.

Liang, X. *et al.* (2019) 'A comprehensive review of topoisomerase inhibitors as anticancer agents in the past decade', *European Journal of Medicinal Chemistry*. Elsevier Masson SAS, pp. 129–168. doi: 10.1016/j.ejmech.2019.03.034.

Liang, X. J. *et al.* (2005) 'Trafficking and localization of platinum complexes in cisplatin-resistant cell lines monitored by fluorescence-labeled platinum', *Journal of Cellular Physiology*. J Cell Physiol, 202(3), pp. 635–641. doi: 10.1002/jcp.20253.

Liao, H. *et al.* (2018) 'Mechanisms for stalled replication fork stabilization: new targets for synthetic lethality strategies in cancer treatments', *EMBO reports*, 19(9). doi: 10.15252/embr.201846263.

Lim, Z. F. and Ma, P. C. (2019) 'Emerging insights of tumor heterogeneity and drug resistance mechanisms in lung cancer targeted therapy', *Journal of Hematology and Oncology*, 12(1). doi: 10.1186/s13045-019-0818-2.

Lin, P. C. *et al.* (2019) 'Germline susceptibility variants impact clinical outcome and therapeutic strategies for stage III colorectal cancer', *Scientific Reports*, 9(1). doi: 10.1038/s41598-019-40571-0.

Litovchick, L. *et al.* (2007) 'Evolutionarily Conserved Multisubunit RBL2/p130 and E2F4 Protein Complex Represses Human Cell Cycle-Dependent Genes in Quiescence', *Molecular Cell*, 26(4), pp. 539–551. doi: 10.1016/j.molcel.2007.04.015.

Liu, W. *et al.* (2011) 'Identification of RNF213 as a susceptibility gene for moyamoya disease and its possible role in vascular development', *PLoS ONE*, 6(7). doi: 10.1371/journal.pone.0022542.

Longley, D. B., Harkin, D. P. and Johnston, P. G. (2003a) '5-Fluorouracil: Mechanisms of action and clinical strategies', *Nature Reviews Cancer*, 3(5), pp. 330–338. doi: 10.1038/nrc1074.

Longley, D. B., Harkin, D. P. and Johnston, P. G. (2003b) '5-Fluorouracil: Mechanisms of action and clinical strategies', *Nature Reviews Cancer*. Nat Rev Cancer, pp. 330–338. doi: 10.1038/nrc1074.

Longley, D. B. and Johnston, P. G. (2005) 'Molecular mechanisms of drug resistance', *Journal of Pathology*. John Wiley & Sons, Ltd, pp. 275–292. doi: 10.1002/path.1706.

Loo, T. W. *et al.* (2012) 'The ATPase activity of the P-glycoprotein drug pump is highly activated when the N-terminal and central regions of the nucleotide-binding domains are linked closely together', *Journal of Biological Chemistry*, 287(32), pp. 26806–26816. doi: 10.1074/jbc.M112.376202.

Lord, C. J. and Ashworth, A. (2012) 'The DNA damage response and cancer therapy', *Nature*. Nature, pp. 287–294. doi: 10.1038/nature10760.

Lord, C. J. and Ashworth, A. (2016) 'BRCAness revisited', *Nature Reviews Cancer*. Nature Publishing Group, 16(2), pp. 110–120. doi: 10.1038/nrc.2015.21.

Lord, C. J. and Ashworth, A. (2017) 'PARP inhibitors: Synthetic lethality in the clinic', *Science*, 355(6330), pp. 1152–1158. doi: 10.1126/science.aam7344.

Lu, J. F., Pokharel, D. and Bebawy, M. (2015) 'MRP1 and its role in anticancer drug resistance', *Drug Metabolism Reviews*. Taylor and Francis Ltd, pp. 406–419. doi: 10.3109/03602532.2015.1105253.

Lund, M. J. *et al.* (2009) 'Race and triple negative threats to breast cancer survival: A population-based study in Atlanta, GA', *Breast Cancer Research and Treatment*. Springer, 113(2), pp. 357–370. doi: 10.1007/s10549-008-9926-3.

Lundberg, A. S. and Weinberg, R. A. (1999) 'Control of the cell cycle and apoptosis', *European Journal of Cancer*. Elsevier Ltd, 35(14), pp. 1886–1894. doi: 10.1016/S0959-8049(99)00292-0.

Lunt, S. Y. and Vander Heiden, M. G. (2011) 'Aerobic Glycolysis: Meeting the Metabolic Requirements of Cell Proliferation', *Annual Review of Cell and Developmental Biology*, 27(1), pp. 441–464. doi: 10.1146/annurev-cellbio-092910-154237.

Luo, P. *et al.* (2018) 'A Large-scale, multicenter serum metabolite biomarker identification study for the early detection of hepatocellular carcinoma', *Hepatology*. John Wiley and Sons Inc., 67(2), pp. 662–675. doi: 10.1002/hep.29561.

Lv, Y. *et al.* (2019) 'Landscape of cancer diagnostic biomarkers from specifically expressed genes',

Briefings in Bioinformatics, 00(July), pp. 1–10. doi: 10.1093/bib/bbz131.

Lyu, Y. L. *et al.* (2006) 'Role of Topoisomerase IIβ in the Expression of Developmentally RegulatedGenes', *Molecular and Cellular Biology*. American Society for Microbiology, 26(21), pp. 7929–7941. doi: 10.1128/mcb.00617-06.

Ma, L. *et al.* (2015) 'Synthesis and biological evaluation of 5-nitropyrimidine-2,4-dione analogues as inhibitors of nitric oxide and iNOS activity', *Chemical Biology and Drug Design*, 85(3), pp. 296–299. doi: 10.1111/cbdd.12386.

Ma, L. *et al.* (2016) 'Fat4 suppression induces Yap translocation accounting for the promoted proliferation and migration of gastric cancer cells', *Cancer Biology & Therapy*, 17. doi: 10.1080/15384047.2015.1108488.

Ma, X. and Yu, H. (2006) *Global Burden of Cancer*, *YALE JOURNAL OF BIOLOGY AND MEDICINE*.

Mancilla, T. R., Iskra, B. and Aune, G. J. (2019) 'Doxorubicin-induced cardiomyopathy in children', *Comprehensive Physiology*. Wiley-Blackwell Publishing Ltd, 9(3), pp. 905–931. doi: 10.1002/cphy.c180017.

Manfredi, M. G. *et al.* (2011) 'Characterization of alisertib (MLN8237), an investigational small-molecule inhibitor of Aurora A kinase using novel in vivo pharmacodynamic assays', *Clinical Cancer Research*, 17(24), pp. 7614–7624. doi: 10.1158/1078-0432.CCR-11-1536.

Mannini, L. *et al.* (2015) 'SMC1B is present in mammalian somatic cells and interacts with mitotic cohesin proteins', *Scientific Reports*. Nature Publishing Group, 5(1), pp. 1–11. doi: 10.1038/srep18472.

Mansoori, B. *et al.* (2017a) 'The Different Mechanisms of Cancer Drug Resistance: A Brief Review', *Adv Pharm Bull*, 7(3), pp. 339–348. doi: 10.15171/apb.2017.041.

Mansoori, B. *et al.* (2017b) 'The Different Mechanisms of Cancer Drug Resistance: A Brief Review', *Tabriz University of Medical Sciences*, 7(3), pp. 339–348. doi: 10.15171/apb.2017.041.

Mardis, E. R. and Wilson, R. K. (2009) 'Cancer genome sequencing: A review', *Human Molecular Genetics*. Oxford University Press, 18(R2), p. R163. doi: 10.1093/hmg/ddp396.

Marteil, G. *et al.* (2018) 'Over-elongation of centrioles in cancer promotes centriole amplification and chromosome missegregation', *Nature Communications*, 9(1). doi: 10.1038/s41467-018-03641-x.

Martin, L. P., Hamilton, T. C. and Schilder, R. J. (2008) 'Platinum resistance: The role of DNA repair pathways', *Clinical Cancer Research*. American Association for Cancer Research, pp. 1291–1295. doi: 10.1158/1078-0432.CCR-07-2238.

Martin, N. and Smith, G. (2007) 'Targeted cancer therapies based on the inhibition of DNA strand break repair', *Oncogene*, 26, pp. 7816–7824. doi: 10.1038/sj.onc.1210879.

Marullo, R. *et al.* (2013a) 'Cisplatin induces a mitochondrial-ros response that contributes to

cytotoxicity depending on mitochondrial redox status and bioenergetic functions', *PLoS ONE*. Public Library of Science, 8(11). doi: 10.1371/journal.pone.0081162.

Marullo, R. *et al.* (2013b) 'Cisplatin induces a mitochondrial-ros response that contributes to cytotoxicity depending on mitochondrial redox status and bioenergetic functions', *PLoS ONE*, 8(11), pp. 1–15. doi: 10.1371/journal.pone.0081162.

Massihnia, D. *et al.* (2016) *Triple negative breast cancer: Shedding light onto the role of pi3k/akt/mtor pathway*, *Oncotarget*. doi: 10.18632/oncotarget.10858.

Matutino, A., Amaro, C. and Verma, S. (2018) 'CDK4/6 inhibitors in breast cancer: beyond hormone receptor-positive HER2-negative disease', *Therapeutic Advances in Medical Oncology*, 10. doi: 10.1177/1758835918818346.

McCabe, N., Kennedy, R. D. and Prise, K. M. (2016) *The role of PTEN as a cancer biomarker*, *Oncoscience*. doi: 10.18632/oncoscience.296.

McCann, K. E., Hurvitz, S. A. and McAndrew, N. (2019) 'Advances in Targeted Therapies for Triple-Negative Breast Cancer', *Drugs*. Springer International Publishing, 79(11), pp. 1217–1230. doi: 10.1007/s40265-019-01155-4.

McDermott, M., Eustance, A. J. and Stordal, B. (2014) 'In vitro Development of Chemotherapy and Targeted Therapy Drug-Resistant Cancer Cell Lines: A Practical Guide with Case Studies', *Frontiers in Oncology*, 4(March). doi: 10.3389/fonc.2014.00040.

McKenna, A. *et al.* (2010) 'The genome analysis toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data', *Genome Research*. Genome Res, 20(9), pp. 1297–1303. doi: 10.1101/gr.107524.110.

McKiernan, C. J. and Friedlander, M. (1999) 'The retinal rod Na+/Ca2+,K+ exchanger contains a noncleaved signal sequence required for translocation of the N terminus', *Journal of Biological Chemistry*. J Biol Chem, 274(53), pp. 38177–38182. doi: 10.1074/jbc.274.53.38177.

McLaren, W. *et al.* (2016) 'The Ensembl Variant Effect Predictor', *Genome Biology*. Genome Biology, 17(1), pp. 1–14. doi: 10.1186/s13059-016-0974-4.

Mehanna, J. *et al.* (2019) 'Triple-negative breast cancer: Current perspective on the evolving therapeutic landscape', *International Journal of Women's Health*, 11, pp. 431–437. doi: 10.2147/IJWH.S178349.

Mehta, J. *et al.* (2015) 'A Molecular Analysis Provides Novel Insights into Androgen Receptor Signalling in Breast Cancer', *PLOS ONE*. Edited by N. Kyprianou. Public Library of Science, 10(3), p. e0120622. doi: 10.1371/journal.pone.0120622.

Mei, L. *et al.* (2019) 'Ataxia telangiectasia and Rad3-related inhibitors and cancer therapy: Where we stand', *Journal of Hematology and Oncology*, 12(1). doi: 10.1186/s13045-019-0733-6.

Melinda, L. T. *et al.* (2016) 'Homologous recombination deficiency (hrd) score predicts response to

platinum-containing neoadjuvant chemotherapy in patients with triple-negative breast cancer', *Clinical Cancer Research*. American Association for Cancer Research Inc., 22(15), pp. 3764–3773. doi: 10.1158/1078-0432.CCR-15-2477.

Metz, R. *et al.* (2007) 'Novel tryptophan catabolic enzyme IDO2 is the preferred biochemical target of the antitumor indoleamine 2,3-dioxygenase inhibitory compound D-1-methyl-tryptophan', *Cancer Research*, 67(15), pp. 7082–7087. doi: 10.1158/0008-5472.CAN-07-1872.

Metzger, R. *et al.* (1998) 'ERCC1 mRNA levels complement thymidylate synthase mRNA levels in predicting response and survival for gastric cancer patients receiving combination cisplatin and fluorouracil chemotherapy', *Journal of Clinical Oncology*. American Society of Clinical Oncology, 16(1), pp. 309–316. doi: 10.1200/JCO.1998.16.1.309.

Meyer, F. *et al.* (2020) 'Prevention of DNA Replication Stress by CHK1 Leads to Chemoresistance Despite a DNA Repair Defect in Homologous Recombination in Breast Cancer', *Cells*, 9(1), p. 238. doi: 10.3390/cells9010238.

Meyers, M. *et al.* (2001) 'Role of the hMLH1 DNA mismatch repair protein in fluoropyrimidine-mediated cell death and cell cycle responses', *Cancer Research*, 61(13), pp. 5193–5201.

Meyers, Mark *et al.* (2001) *Role of the hMLH1 DNA Mismatch Repair Protein in Fluoropyrimidine-mediated Cell Death and Cell Cycle Responses 1*, *CANCER RESEARCH*. Available at: http://www.cancer.org. (Accessed: 22 June 2020).

Meyers, M. *et al.* (2005) 'DNA mismatch repair-dependent response to fluoropyrimidine-generated damage', *Journal of Biological Chemistry*, 280(7), pp. 5516–5526. doi: 10.1074/jbc.M412105200.

Meyerson, M., Gabriel, S. and Getz, G. (2010) 'Advances in understanding cancer genomes through second-generation sequencing', *Nature Reviews Genetics*. Nat Rev Genet, pp. 685–696. doi: 10.1038/nrg2841.

Michaelis, M. *et al.* (2009) 'Reversal of P-glycoprotein-Mediated Multidrug Resistance by the Murine Double Minute 2 Antagonist Nutlin-3', *Cancer Res*, 69(2), pp. 416–437. doi: 10.1158/0008-5472.CAN-08-1856.

Michaelis, M. *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 and Disease*, 2(12). doi: 10.1038/cddis.2011.129.

Michaelis, M. *et al.* (2012) 'Human neuroblastoma cells with acquired resistance to the p53 activator RITA retain functional p53 and sensitivity to other p53 activating agents', *Cell Death and Disease*, 3(4), pp. 1–8. doi: 10.1038/cddis.2012.35.

Michaelis, M. *et al.* (2014) 'Aurora kinases as targets in drug-resistant neuroblastoma cells', *PLoS ONE*, 9(9), pp. 1–9. doi: 10.1371/journal.pone.0108758.

Michaelis, M. *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*,

12(5), pp. 1–17. doi: 10.3390/cancers12051080.

Michaelis, M., Wass, M. N. and Cinatl, J. (2019) 'Drug-adapted cancer cell lines as preclinical models of acquired resistance', *Cancer Drug Resistance*. OAE Publishing Inc., 2(3), pp. 447–456. doi: 10.20517/cdr.2019.005.

Michels, J. *et al.* (2014) 'Predictive biomarkers for cancer therapy with PARP inhibitors', *Oncogene*. Nature Publishing Group, pp. 3894–3907. doi: 10.1038/onc.2013.352.

Mills, K. *et al.* (2017) 'A Small Molecule RAD51 Inhibitor Preferentially Affects Cells Expressing High Cytidine Deaminase Activity', *Blood*. American Society of Hematology, 130(Supplement 1), pp. 4627–4627. doi: 10.1182/BLOOD.V130.SUPPL_1.4627.4627.

Von Minckwitz, G. *et al.* (2014) 'Neoadjuvant carboplatin in patients with triple-negative and HER2-positive early breast cancer (GeparSixto; GBG 66): A randomised phase 2 trial', *The Lancet Oncology*. Lancet Publishing Group, 15(7), pp. 747–756. doi: 10.1016/S1470-2045(14)70160-3.

Mini, E. *et al.* (2006) 'Cellular pharmacology of gemcitabine', *Annals of Oncology*, 17(SUPPL. 5). doi: 10.1093/annonc/mdj941.

Mitchison, T. J. (2012) 'The proliferation rate paradox in antimitotic chemotherapy', *Molecular Biology of the Cell*. Mol Biol Cell, pp. 1–6. doi: 10.1091/mbc.E10-04-0335.

Miura, K. *et al.* (2010) '5-FU metabolism in cancer and orally-administrable 5-FU drugs', *Cancers*. Multidisciplinary Digital Publishing Institute (MDPI), pp. 1717–1730. doi: 10.3390/cancers2031717.

Mladenov, E. *et al.* (2019) 'DNA-PKcs and ATM epistatically suppress DNA end resection and hyperactivation of ATR-dependent G2-checkpoint in S-phase irradiated cells', *Scientific Reports*. Nature Publishing Group, 9(1), pp. 1–16. doi: 10.1038/s41598-019-51071-6.

Mohiuddin, I. S. and Kang, M. H. (2019) 'DNA-PK as an Emerging Therapeutic Target in Cancer', *Frontiers in Oncology*, 9(July), pp. 1–8. doi: 10.3389/fonc.2019.00635.

Moitra, K. (2015) 'Overcoming Multidrug Resistance in Cancer Stem Cells', *BioMed Research International*, 2015. doi: 10.1155/2015/635745.

Mojardín, L. *et al.* (2013) 'New Insights into the RNA-Based Mechanism of Action of the Anticancer Drug 59-Fluorouracil in Eukaryotic Cells', *PLoS ONE*, 8(11), p. 78172. doi: 10.1371/journal.pone.0078172.

Montanari, F. and Ecker, G. F. (2015) 'Prediction of drug-ABC-transporter interaction - Recent advances and future challenges', *Advanced Drug Delivery Reviews*. Elsevier, pp. 17–26. doi: 10.1016/j.addr.2015.03.001.

Montano, R. *et al.* (2012) 'Preclinical development of the novel Chk1 inhibitor SCH900776 in combination with DNA-damaging agents and antimetabolites.', *Molecular cancer therapeutics*, 11(2), pp. 427–38. doi: 10.1158/1535-7163.MCT-11-0406.

Montemorano, L., Michelle, D. S. and Bixel, L. K. (2019) 'Role of olaparib as maintenance treatment

for ovarian cancer: The evidence to date', *OncoTargets and Therapy*. Dove Medical Press Ltd., pp. 11497–11506. doi: 10.2147/OTT.S195552.

Mosmann, T. (1983) 'Rapid colorimetric assay for cellular growth and survival: Application to proliferation and cytotoxicity assays', *Journal of Immunological Methods*, 65(1–2), pp. 55–63. doi: 10.1016/0022-1759(83)90303-4.

Mouliere, F. *et al.* (2018) 'Enhanced detection of circulating tumor DNA by fragment size analysis', *Science Translational Medicine*. American Association for the Advancement of Science, 10(466). doi: 10.1126/scitranslmed.aat4921.

Mouliere, F. and Rosenfeld, N. (2015) 'Circulating tumor-derived DNA is shorter than somatic DNA in plasma', *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 112(11), pp. 3178–3179. doi: 10.1073/pnas.1501321112.

Moureau, S. *et al.* (2016) 'Therapeutic potential of novel PLK1 inhibitor CYC140 in esophageal cancer and acute leukemia', *European Journal of Cancer*. Elsevier BV, 69, p. S117. doi: 10.1016/s0959-8049(16)32948-3.

Mukhopadhyay, P. *et al.* (2011) 'Mucins in the pathogenesis of breast cancer: Implications in diagnosis, prognosis and therapy', *Biochimica et Biophysica Acta - Reviews on Cancer*, 1815(2), pp. 224–240. doi: 10.1016/j.bbcan.2011.01.001.

Muley, H. *et al.* (2020) 'Drug uptake-based chemoresistance in breast cancer treatment', *Biochemical Pharmacology*. Elsevier Inc., p. 113959. doi: 10.1016/j.bcp.2020.113959.

Myant, K. B. *et al.* (2017) 'HUWE1 is a critical colonic tumour suppressor gene that prevents MYC signalling, DNA damage accumulation and tumour initiation', *EMBO Molecular Medicine*, 9(2), pp. 181–197. doi: 10.15252/emmm.201606684.

Nair, J. *et al.* (2020) 'Resistance to the CHK1 inhibitor prexasertib involves functionally distinct CHK1 activities in BRCA wild-type ovarian cancer', *Oncogene*. Springer US, 39(33), pp. 5520–5535. doi: 10.1038/s41388-020-1383-4.

Nakamura, Y. *et al.* (2015) *Expression and intracellular localization of TBC1D9, a Rab GTPase-accelerating protein, in mouse testes*, *Experimental Animals*. doi: 10.1538/expanim.15-0016.

Nanayakkara, A. K. *et al.* (2018) 'Targeted inhibitors of P-glycoprotein increase chemotherapeutic-induced mortality of multidrug resistant tumor cells', *Scientific Reports*. Nature Publishing Group, 8(1), pp. 1–18. doi: 10.1038/s41598-018-19325-x.

Napoli, C. *et al.* (2013) 'Effects of nitric oxide on cell proliferation: Novel insights', *Journal of the American College of Cardiology*, 62(2), pp. 89–95. doi: 10.1016/j.jacc.2013.03.070.

Natarajan, K. *et al.* (2012) 'Role of breast cancer resistance protein (BCRP/ABCG2) in cancer drug resistance', *Biochemical Pharmacology*. Elsevier, pp. 1084–1103. doi: 10.1016/j.bcp.2012.01.002.

Negrini, S., Gorgoulis, V. G. and Halazonetis, T. D. (2010) 'Genomic instability an evolving hallmark

of cancer', *Nature Reviews Molecular Cell Biology*. Nat Rev Mol Cell Biol, pp. 220–228. doi: 10.1038/nrm2858.

Ng, P. C. and Henikoff, S. (2003) 'SIFT: Predicting amino acid changes that affect protein function', *Nucleic Acids Research*, 31(13), pp. 3812–3814. doi: 10.1093/nar/gkg509.

Ng, S. W. K. *et al.* (2016) 'A 17-gene stemness score for rapid determination of risk in acute leukaemia', *Nature*. Nature Publishing Group, 540(7633), pp. 433–437. doi: 10.1038/nature20598.

Nguyen, D. J. M. *et al.* (2020) *Targeting the kynurenine pathway for the treatment of cisplatin-resistant lung cancer*, *Molecular Cancer Research*. doi: 10.1158/1541-7786.MCR-19-0239.

Nichols, A. C. *et al.* (2014) 'Exploiting high-throughput cell line drug screening studies to identify candidate therapeutic agents in head and neck cancer', *BMC Pharmacology and Toxicology*, 15(1), pp. 1–10. doi: 10.1186/2050-6511-15-66.

Di Nicolantonio, F. *et al.* (2005) 'Cancer cell adaptation to chemotherapy', *BMC Cancer*. BMC Cancer, 5. doi: 10.1186/1471-2407-5-78.

Nijman, S. M. B. (2011) 'Synthetic lethality: General principles, utility and detection using genetic screens in human cells', *FEBS Letters*. Elsevier B.V., pp. 1–6. doi: 10.1016/j.febslet.2010.11.024.

Nik-Zainal, S. *et al.* (2012) 'Mutational processes molding the genomes of 21 breast cancers', *Cell*, 149(5), pp. 979–993. doi: 10.1016/j.cell.2012.04.024.

Nitiss, J. L. (2009) 'Targeting DNA topoisomerase II in cancer chemotherapy', *Nature Reviews Cancer*. Nat Rev Cancer, pp. 338–350. doi: 10.1038/nrc2607.

Niu, W. *et al.* (2020) 'BRD7 suppresses invasion and metastasis in breast cancer by negatively regulating YB1-induced epithelial-mesenchymal transition', *Journal of Experimental and Clinical Cancer Research*. BioMed Central Ltd., 39(1), p. 30. doi: 10.1186/s13046-019-1493-4.

Niu, Y., Xu, J. and Sun, T. (2019) 'Cyclin-dependent kinases 4/6 inhibitors in breast cancer: Current status, resistance, and combination strategies', *Journal of Cancer*. Ivyspring International Publisher, 10(22), pp. 5504–5517. doi: 10.7150/jca.32628.

Norquist, B. *et al.* (2011) 'Secondary somatic mutations restoring BRCA1/2 predict chemotherapy resistance in hereditary ovarian carcinomas', *Journal of Clinical Oncology*. J Clin Oncol, 29(22), pp. 3008–3015. doi: 10.1200/JCO.2010.34.2980.

Nowsheen, S. and Yang, E. S. (2012a) 'The intersection between DNA damage response and cell death pathways', *Experimental Oncology*. NIH Public Access, pp. 243–254. Available at: /pmc/articles/PMC3754840/?report=abstract (Accessed: 9 October 2020).

Nowsheen, S. and Yang, E. S. (2012b) 'The intersection between DNA damage response and cell death pathways', *Experimental Oncology*. NIH Public Access, pp. 243–254. Available at: /pmc/articles/PMC3754840/?report=abstract (Accessed: 23 June 2020).

Nykjaer, A. *et al.* (2001) 'Cubilin dysfunction causes abnormal metabolism of the steroid hormone

25(OH) vitamin D3', *Proceedings of the National Academy of Sciences of the United States of America*. Proc Natl Acad Sci U S A, 98(24), pp. 13895–13900. doi: 10.1073/pnas.241516998.

O'Neil, N. J., Bailey, M. L. and Hieter, P. (2017) 'Synthetic lethality and cancer', *Nature Reviews Genetics*, 18(10), pp. 613–623. doi: 10.1038/nrg.2017.47.

O'Shaughnessy, J., Kaklamani, V. and Kalinsky, K. (2019) 'Perspectives on the mechanism of action and clinical application of eribulin for metastatic breast cancer', *Future Oncology*. Future Medicine Ltd., pp. 1641–1653. doi: 10.2217/fon-2018-0936.

Oba, T., Izumi, H. and Ito, K. (2016) 'ABCB1 and ABCC11 confer resistance to eribulin in breast cancer cell lines', *Oncotarget*, 7(43). doi: 10.18632/oncotarget.11727.

Odintsova, T. I. *et al.* (2003) 'Characterization and analysis of posttranslational modifications of the human large cytoplasmic ribosomal subunit proteins by mass spectrometry and Edman sequencing', *Journal of Protein Chemistry*. J Protein Chem, 22(3), pp. 249–258. doi: 10.1023/A:1025068419698.

Ohga, T. *et al.* (1996) 'Role of the human Y box-binding protein YB-1 in cellular sensitivity to the DNA-damaging agents cisplatin, mitomycin C, and ultraviolet light', *Cancer Research*, 56(18), pp. 4224–4228.

Okabe, M. *et al.* (2005a) 'Characterization of the organic cation transporter SLC22A16: A doxorubicin importer', *Biochemical and Biophysical Research Communications*. Biochem Biophys Res Commun, 333(3), pp. 754–762. doi: 10.1016/j.bbrc.2005.05.174.

Okabe, M. *et al.* (2005b) *Characterization of the organic cation transporter SLC22A16: A doxorubicin importer*, *Biochemical and Biophysical Research Communications*. doi: 10.1016/j.bbrc.2005.05.174.

Okuda, T. *et al.* (2000) 'Molecular cloning, expression, and chromosomal mapping of human chondroitin 4-sulfotransferase, whose expression pattern in human tissues is different from that of chondroitin 6-sulfotransferase', *Journal of Biochemistry*. Japanese Biochemical Society, 128(5), pp. 763–770. doi: 10.1093/oxfordjournals.jbchem.a022813.

Olaussen, K. A. *et al.* (2006) 'DNA repair by ERCC1 in non-small-cell lung cancer and cisplatin-based adjuvant chemotherapy', *New England Journal of Medicine*. N Engl J Med, 355(10), pp. 983–991. doi: 10.1056/NEJMoa060570.

Oltean, A. *et al.* (2018) 'Quantifying ciliary dynamics during assembly reveals stepwise waveform maturation in airway cells', *American Journal of Respiratory Cell and Molecular Biology*. American Thoracic Society, 59(4), pp. 511–522. doi: 10.1165/rcmb.2017-0436OC.

Oronsky, B. T. *et al.* (2012) 'The Scarlet Letter of Alkylation: A Mini Review of Selective Alkylating Agents', *Translational Oncology*, 5, pp. 226–229. doi: 10.1593/tlo.12187.

Osorio, F. G. *et al.* (2016) 'Loss of the proteostasis factor AIRAPL causes myeloid transformation by deregulating IGF-1 signaling', *Nature Medicine*, 22(1), pp. 91–96. doi: 10.1038/nm.4013.

Overbeck, T. R. *et al.* (2017) 'ABCA3 phenotype in non-small cell lung cancer indicates poor outcome', *Oncology (Switzerland)*, 93(4), pp. 270–278. doi: 10.1159/000477619.

Oz, S., Ivashko-Pachima, Y. and Gozes, I. (2012) 'The ADNP Derived Peptide, NAP Modulates the Tubulin Pool: Implication for Neurotrophic and Neuroprotective Activities', *PLoS ONE*. Edited by A. I. Bush. Public Library of Science, 7(12), p. e51458. doi: 10.1371/journal.pone.0051458.

Pagani, F., Raponi, M. and Baralle, F. E. (2005) 'Synonymous mutations in CFTR exon 12 affect splicing and are not neutral in evolution', *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 102(18), pp. 6368–6372. doi: 10.1073/pnas.0502288102.

Papouli, E., Cejka, P. and Jiricny, J. (2004) 'Dependence of the cytotoxicity of DNA-damaging agents on the mismatch repair status of human cells', *Cancer Research*, 64(10), pp. 3391–3394. doi: 10.1158/0008-5472.CAN-04-0513.

Park, S. W. *et al.* (2003) 'Analysis of P-glycoprotein-mediated membrane transport in human peripheral blood lymphocytes using the UIC2 shift assay', *Cytometry*. Wiley-Liss Inc., 53A(2), pp. 67–78. doi: 10.1002/cyto.a.10039.

Parsons, J. L. *et al.* (2009) 'Ubiquitin ligase ARF-BP1/Mule modulates base excision repair', *EMBO Journal*. Nature Publishing Group, 28(20), pp. 3207–3215. doi: 10.1038/emboj.2009.243.

Pascoe, J. M. and Roberts, J. J. (1974) 'Interactions between mammalian cell DNA and inorganic platinum compounds-I. DNA interstrand cross-linking and cytotoxic properties of platinum(II) compounds', *Biochemical Pharmacology*, 23(9), pp. 1345–1357. doi: 10.1016/0006-2952(74)90354-2.

Patil, M., Pabla, N. and Dong, Z. (2013) 'Checkpoint kinase 1 in DNA damage response and cell cycle regulation', *Cellular and Molecular Life Sciences*. NIH Public Access, pp. 4009–4021. doi: 10.1007/s00018-013-1307-3.

Pattabiraman, D. R. *et al.* (2016) 'Activation of PKA leads to mesenchymal-to-epithelial transition and loss of tumor-initiating ability', *Science*. American Association for the Advancement of Science, 351(6277), p. aad3680. doi: 10.1126/science.aad3680.

Pechmann, S. and Frydman, J. (2013) 'Evolutionary conservation of codon optimality reveals hidden signatures of cotranslational folding', *Nature Structural and Molecular Biology*. Nat Struct Mol Biol, 20(2), pp. 237–243. doi: 10.1038/nsmb.2466.

Perri, F., Pisconti, S. and Vittoria Scarpati, G. Della (2016) 'P53 mutations and cancer: A tight linkage', *Annals of Translational Medicine*. AME Publishing Company, 4(24), p. 44. doi: 10.21037/atm.2016.12.40.

Peshkin, B. N., Alabek, M. L. and Isaacs, C. (2010) 'BRCA1/2 mutations and triple negative breast cancers', *Breast Disease*. Breast Dis, 32(1–2), pp. 25–33. doi: 10.3233/BD-2010-0306.

Pfister, S. X. *et al.* (2015) 'Inhibiting WEE1 Selectively Kills Histone H3K36me3-Deficient Cancers by

dNTP Starvation', *Cancer Cell*. Cell Press, 28(5), pp. 557–568. doi: 10.1016/j.ccell.2015.09.015.

Phi, L. T. H. *et al.* (2018) 'Cancer stem cells (CSCs) in drug resistance and their therapeutic implications in cancer treatment', *Stem Cells International*, 2018. doi: 10.1155/2018/5416923.

Pichler, A. *et al.* (2002) 'The nucleoporin RanBP2 has SUMO1 E3 ligase activity', *Cell*, 108(1), pp. 109–120. doi: 10.1016/S0092-8674(01)00633-X.

Pike, K. G. *et al.* (2018) 'The Identification of Potent, Selective, and Orally Available Inhibitors of Ataxia Telangiectasia Mutated (ATM) Kinase: The Discovery of AZD0156 (8-{6-[3-(Dimethylamino)propoxy]pyridin-3-yl}-3-methyl-1-(tetrahydro-2 H-pyran-4-yl)-1,3-dihydro-2 H-imidazo[4', *Journal of Medicinal Chemistry*, 61(9), pp. 3823–3841. doi: 10.1021/acs.jmedchem.7b01896.

Pluchino, K. M. *et al.* (2012) 'Collateral sensitivity as a strategy against cancer multidrug resistance', *Drug Resistance Updates*, 15(1–2), pp. 98–105. doi: 10.1016/j.drup.2012.03.002.

Plummer, E. R. *et al.* (2019) 'A first-in-human phase I/II trial of SRA737 (a Chk1 Inhibitor) in subjects with advanced cancer.', *Journal of Clinical Oncology*. American Society of Clinical Oncology (ASCO), 37(15_suppl), pp. 3094–3094. doi: 10.1200/jco.2019.37.15_suppl.3094.

Pommier, Y. *et al.* (2010) 'DNA topoisomerases and their poisoning by anticancer and antibacterial drugs', *Chemistry and Biology*. Chem Biol, pp. 421–433. doi: 10.1016/j.chembiol.2010.04.012.

Poole, L. A. and Cortez, D. (2017) 'Functions of SMARCAL1, ZRANB3, and HLTF in maintaining genome stability', *Critical Reviews in Biochemistry and Molecular Biology*. Taylor and Francis Ltd, pp. 696–714. doi: 10.1080/10409238.2017.1380597.

Post, G. C. *et al.* (2005) 'Doxazolidine, a proposed active metabolite of doxorubicin that cross-links DNA', *Journal of Medicinal Chemistry*. American Chemical Society , 48(24), pp. 7648–7657. doi: 10.1021/jm050678v.

Povey, A. C. *et al.* (2002) *DNA Alkylation and Repair in the Large Bowel: Animal and Human Studies 1,2*, *International Research Conference on Food*. Available at: https://academic.oup.com/jn/article-abstract/132/11/3518S/4687186 (Accessed: 12 June 2020).

Presnyak, V. *et al.* (2015) 'Codon optimality is a major determinant of mRNA stability', *Cell*. Cell Press, 160(6), pp. 1111–1124. doi: 10.1016/j.cell.2015.02.029.

Puccetti, M. V. *et al.* (2019) 'SMARCAL1 and ZrAnB3 protect replication forks from MYC-induced DNA replication stress', *Cancer Research*. American Association for Cancer Research Inc., 79(7), pp. 1612–1623. doi: 10.1158/0008-5472.CAN-18-2705.

Raab, M. *et al.* (2014) 'Quantitative chemical proteomics reveals a Plk1 inhibitor-compromised cell death pathway in human cells', *Cell research*. Nature Publishing Group, pp. 1141–1145. doi: 10.1038/cr.2014.86.

Rajeshkumar, N. V. *et al.* (2011) 'MK-1775, a potent Wee1 inhibitor, synergizes with gemcitabine to

achieve tumor regressions, selectively in p53-deficient pancreatic cancer xenografts', *Clinical Cancer Research*. Clin Cancer Res, 17(9), pp. 2799–2806. doi: 10.1158/1078-0432.CCR-10-2580.

Raudvere, U. *et al.* (2019) 'g:Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update)', *Nucleic Acids Research*, 47(W1), pp. W191–W198. doi: 10.1093/nar/gkz369.

Reis-Filho, J. S. and Tutt, A. N. J. (2008) 'Triple negative tumours: A critical review', *Histopathology*, 52(1), pp. 108–118. doi: 10.1111/j.1365-2559.2007.02889.x.

Reva, B., Antipin, Y. and Sander, C. (2011) 'Predicting the functional impact of protein mutations: Application to cancer genomics', *Nucleic Acids Research*, 39(17), pp. 37–43. doi: 10.1093/nar/gkr407.

Reverter, D. and Lima, C. D. (2005) 'Insights into E3 ligase activity revealed by a SUMO-RanGAP1-Ubc9-Nup358 complex', *Nature*, 435(7042), pp. 687–692. doi: 10.1038/jid.2014.371.

Rheinbay, E. *et al.* (2020) 'Analyses of non-coding somatic drivers in 2,658 cancer whole genomes', *Nature*, 578(7793), pp. 102–111. doi: 10.1038/s41586-020-1965-x.

Riches, L. C. *et al.* (2020) 'Pharmacology of the ATM inhibitor AZD0156: Potentiation of irradiation and olaparib responses preclinically', *Molecular Cancer Therapeutics*. American Association for Cancer Research Inc., 19(1), pp. 13–25. doi: 10.1158/1535-7163.MCT-18-1394.

Ringel, J. and Löhr, M. (2003) 'The MUC gene family: Their role in diagnosis and early detection of pancreatic cancer', *Molecular Cancer*, 2, pp. 7–11. doi: 10.1186/1476-4598-2-9.

Rivlin, N. *et al.* (2011) 'Mutations in the p53 tumor suppressor gene: Important milestones at the various steps of tumorigenesis', *Genes and Cancer*. Genes Cancer, pp. 466–474. doi: 10.1177/1947601911408889.

Robasky, K., Lewis, N. E. and Church, G. M. (2014) 'The role of replicates for error mitigation in next-generation sequencing', *Nature Reviews Genetics*. Nature Publishing Group, pp. 56–62. doi: 10.1038/nrg3655.

Robert, T. *et al.* (2016) 'The Topo VIB-Like protein family is required for meiotic DNA double-strand break formation', *Science*. American Association for the Advancement of Science, 351(6276), pp. 943–949. doi: 10.1126/science.aad5309.

Robson, M. E. *et al.* (2019) 'OlympiAD final overall survival and tolerability results: Olaparib versus chemotherapy treatment of physician's choice in patients with a germline BRCA mutation and HER2-negative metastatic breast cancer', *Annals of Oncology*, 30(4), pp. 558–566. doi: 10.1093/annonc/mdz012.

Rolfo, C. *et al.* (2014) 'Liquid biopsies in lung cancer: The new ambrosia of researchers', *Biochimica et Biophysica Acta - Reviews on Cancer*. Elsevier B.V., 1846(2), pp. 539–546. doi: 10.1016/j.bbcan.2014.10.001.

Román, M. *et al.* (2018) 'KRAS oncogene in non-small cell lung cancer: Clinical perspectives on the treatment of an old target', *Molecular Cancer*. BioMed Central Ltd., pp. 1–14. doi: 10.1186/s12943-018-0789-x.

Romond, E. H. *et al.* (2005) 'Trastuzumab plus adjuvant chemotherapy for operable HER2-positive breast cancer', *New England Journal of Medicine*. N Engl J Med, 353(16), pp. 1673–1684. doi: 10.1056/NEJMoa052122.

Ronchi, A. *et al.* (2020) *Current and potential immunohistochemical biomarkers for prognosis and therapeutic stratification of breast carcinoma*, *Seminars in Cancer Biology*. doi: 10.1016/j.semcancer.2020.03.002.

Roos, W. P., Thomas, A. D. and Kaina, B. (2016) 'DNA damage and the balance between survival and death in cancer biology', *Nature Reviews Cancer*, 16(1), pp. 20–33. doi: 10.1038/nrc.2015.2.

Russell, P. and Nurse, P. (1987) 'Negative regulation of mitosis by wee1+, a gene encoding a protein kinase homolog', *Cell*. Elsevier, 49(4), pp. 559–567. doi: 10.1016/0092-8674(87)90458-2.

Rutman, R. J., Cantarow, A. and Paschkis, K. E. (1954) 'Studies in 2-Acetylaminofluorene Carcinogenesis', *Cancer Research*, 14(2).

Ryu, D. W. *et al.* (2011) 'Clinical significance of morphologic characteristics in triple negative breast cancer', *Journal of the Korean Surgical Society*, 80(5), pp. 301–306. doi: 10.4174/jkss.2011.80.5.301.

S., A. *et al.* (2013) 'DNA Repair and Resistance to Cancer Therapy', in *New Research Directions in DNA Repair*. InTech. doi: 10.5772/53952.

Saadi, I. *et al.* (2011) 'Deficiency of the cytoskeletal protein SPECC1L leads to oblique facial clefting', *American Journal of Human Genetics*. The American Society of Human Genetics, 89(1), pp. 44–55. doi: 10.1016/j.ajhg.2011.05.023.

Sadasivam, S. and DeCaprio, J. A. (2013) 'The DREAM complex: Master coordinator of cell cycle-dependent gene expression', *Nature Reviews Cancer*, 13(8), pp. 585–595. doi: 10.1038/nrc3556.

Safaei, R. *et al.* (2005) 'Abnormal lysosomal trafficking and enhanced exosomal export of cisplatin in drug-resistant human ovarian carcinoma cells', *Molecular Cancer Therapeutics*, 4(10), pp. 1595–1604. doi: 10.1158/1535-7163.MCT-05-0102.

Sakurikar, N. *et al.* (2016) 'A subset of cancer cell lines is acutely sensitive to the Chk1 inhibitor MK-8776 as monotherapy due to CDK2 activation in S phase', *Oncotarget*, 7(2), pp. 1380–1394. doi: 10.18632/oncotarget.6364.

Saleh-Gohari, N. *et al.* (2005) 'Spontaneous Homologous Recombination Is Induced by Collapsed Replication Forks That Are Caused by Endogenous DNA Single-Strand Breaks', *Molecular and Cellular Biology*. American Society for Microbiology, 25(16), pp. 7158–7169. doi: 10.1128/mcb.25.16.7158-7169.2005.

Sanborn, J. Z. *et al.* (2015) 'Phylogenetic analyses of melanoma reveal complex patterns of

metastatic dissemination', *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 112(35), pp. 10995–11000. doi: 10.1073/pnas.1508074112.

Sawada, M. *et al.* (2003) 'Ku70 suppresses the apoptotic translocation of bax to mitochondria', *Nature Cell Biology*. Nat Cell Biol, 5(4), pp. 320–329. doi: 10.1038/ncb950.

Schlacher, K., Wu, H. and Jasin, M. (2012) 'A Distinct Replication Fork Protection Pathway Connects Fanconi Anemia Tumor Suppressors to RAD51-BRCA1/2', *Cancer Cell*. NIH Public Access, 22(1), pp. 106–116. doi: 10.1016/j.ccr.2012.05.015.

Schmid, P. *et al.* (2018) 'Atezolizumab and Nab-Paclitaxel in Advanced Triple-Negative Breast Cancer', *New England Journal of Medicine*. Massachussetts Medical Society, 379(22), pp. 2108–2121. doi: 10.1056/NEJMoa1809615.

Schmid, P. *et al.* (2020) 'Pembrolizumab for Early Triple-Negative Breast Cancer', *New England Journal of Medicine*. Massachussetts Medical Society, 382(9), pp. 810–821. doi: 10.1056/NEJMoa1910549.

Schmit, F. *et al.* (2007) 'LINC, a human complex that is related to pRB-containing complexes in invertebrates regulates the expression of G2/M genes', *Cell Cycle*, 6(15), pp. 1903–1913. doi: 10.4161/cc.6.15.4512.

Schmitz, C. *et al.* (2003) 'Regulation of vertebrate cellular Mg2+ homeostasis by TRPM7', *Cell*. Cell Press, 114(2), pp. 191–200. doi: 10.1016/S0092-8674(03)00556-7.

Scholz, B. *et al.* (2016) 'Endothelial RSPO3 Controls Vascular Stability and Pruning through Non-canonical WNT/Ca2+/NFAT Signaling', *Developmental Cell*. Elsevier Ltd, 36(1), pp. 79–93. doi: 10.1016/j.devcel.2015.12.015.

Schwartz, P. M. *et al.* (1985) 'Role of uridine phosphorylase in the anabolism of 5-fluorouracil', *Biochemical Pharmacology*. Elsevier Inc., 34(19), pp. 3585–3589. doi: 10.1016/0006-2952(85)90737-3.

Schwerk, C. *et al.* (2003) 'ASAP, a Novel Protein Complex Involved in RNA Processing and Apoptosis', *Molecular and Cellular Biology*. American Society for Microbiology, 23(8), pp. 2981–2990. doi: 10.1128/mcb.23.8.2981-2990.2003.

Scully, R. *et al.* (2019) 'DNA double-strand break repair-pathway choice in somatic mammalian cells', *Nature Reviews Molecular Cell Biology*. Springer US, 20(11), pp. 698–714. doi: 10.1038/s41580-019-0152-0.

Sethi, M. K. *et al.* (2010) 'Identification of glycosyltransferase 8 family members as xylosyltransferases acting on O-glucosylated Notch epidermal growth factor repeats', *Journal of Biological Chemistry*, 285(3), pp. 1582–1586. doi: 10.1074/jbc.C109.065409.

Seyfried, T. N. (2015) 'Cancer as a mitochondrial metabolic disease', *Frontiers in Cell and Developmental Biology*. Frontiers Media S.A., 3(JUL), p. 43. doi: 10.3389/fcell.2015.00043.

Sharma, S. (2009) 'Tumor markers in clinical practice: General principles and guidelines', *Indian Journal of Medical and Paediatric Oncology*. Medknow, 30(1), p. 1. doi: 10.4103/0971-5851.56328.

Shibue, T. and Weinberg, R. A. (2017a) 'EMT, CSCs, and drug resistance: the mechanistic link and clinical implications', *Nature Publishing Group*. doi: 10.1038/nrclinonc.2017.44.

Shibue, T. and Weinberg, R. A. (2017b) 'EMT , CSCs , and drug resistance : the mechanistic link and clinical implications', *Nature Reviews Clinical Oncology*. Nature Publishing Group, 14, pp. 611–629. doi: 10.1038/nrclinonc.2017.44.

Shintre, C. A. *et al.* (2013) 'Structures of ABCB10, a human ATP-binding cassette transporter in apo-And nucleotide-bound states', *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 110(24), pp. 9710–9715. doi: 10.1073/pnas.1217042110.

Shyr, C. *et al.* (2014) *FLAGS, frequently mutated genes in public exomes*, *BMC Medical Genomics*. doi: 10.1186/s12920-014-0064-y.

Siddik, Z. H. (2005) 'Mechanisms of Action of Cancer Chemotherapeutic Agents: DNA-Interactive Alkylating Agents and Antitumour Platinum-Based Drugs', in *The Cancer Handbook*. Chichester, UK: John Wiley & Sons, Ltd. doi: 10.1002/0470025077.chap84b.

Sigismund, S., Avanzato, D. and Lanzetti, L. (2018) 'Emerging functions of the EGFR in cancer', *Molecular Oncology*, 12(1), pp. 3–20. doi: 10.1002/1878-0261.12155.

Sijens, P. E., Baldwin, N. J. and Ng, T. C. (1991) 'Multinuclear MR investigation of the metabolic response of the murine RIF-1 tumor to 5-fluorouracil chemotherapy', *Magnetic Resonance in Medicine*, 19(2), pp. 373–385. doi: 10.1002/mrm.1910190228.

da Silva, J. L. *et al.* (2020) 'Triple negative breast cancer: A thorough review of biomarkers', *Critical Reviews in Oncology/Hematology*. Elsevier, 145(December 2019), p. 102855. doi: 10.1016/j.critrevonc.2019.102855.

Skehan, P. *et al.* (1990) 'New colorimetric cytotoxicity assay for anticancer-drug screening', *Journal of the National Cancer Institute*, 82(13), pp. 1107–1112. doi: 10.1093/jnci/82.13.1107.

Slamon, D. J. *et al.* (2001) 'Use of chemotherapy plus a monoclonal antibody against her2 for metastatic breast cancer that overexpresses HER2', *New England Journal of Medicine*. N Engl J Med, 344(11), pp. 783–792. doi: 10.1056/NEJM200103153441101.

Smith, J. A. *et al.* (2010) 'Eribulin binds at microtubule ends to a single site on tubulin to suppress dynamic instability', *Biochemistry*. Biochemistry, 49(6), pp. 1331–1337. doi: 10.1021/bi901810u.

Smith, P. K. *et al.* (1985) 'Measurement of protein using bicinchoninic acid', *Analytical Biochemistry*, 150(1), pp. 76–85. doi: 10.1016/0003-2697(85)90442-7.

Snyder, J. P. *et al.* (2001) 'The binding conformation of Taxol in β-tubulin: A model based on electron crystallographic density', *Proceedings of the National Academy of Sciences of the United States of*

America. National Academy of Sciences, 98(9), pp. 5312–5316. doi: 10.1073/pnas.051309398.

Solit, D. B. *et al.* (2002) 'Akt Forms an Intracellular Complex with Heat Shock Protein 90 (Hsp90) and Cdc37 and Is Destabilized by Inhibitors of Hsp90 Function', *Journal of Biological Chemistry*, 277(42), pp. 39858–39866. doi: 10.1074/jbc.m206322200.

Song, G. *et al.* (2008) 'Osteopontin promotes ovarian cancer progression and cell survival and increases HIF-1α expression through the PI3-K/Akt pathway', *Cancer Science*. Cancer Sci, 99(10), pp. 1901–1907. doi: 10.1111/j.1349-7006.2008.00911.x.

Sørensen, C. S. *et al.* (2003) 'Chk1 regulates the S phase checkpoint by coupling the physiological turnover and ionizing radiation-induced accelerated proteolysis of Cdc25A', *Cancer Cell*. Cell Press, 3(3), pp. 247–258. doi: 10.1016/S1535-6108(03)00048-5.

Spencer, C. J., Yakymchuk, C. and Ghaznavi, M. (2017) 'Visualising data distributions with kernel density estimation and reduced chi-squared statistic', *Geoscience Frontiers*. Elsevier Ltd, 8(6), pp. 1247–1252. doi: 10.1016/j.gsf.2017.05.002.

Spiliotaki, M. *et al.* (2011) 'Targeting the insulin-like growth factor I receptor inhibits proliferation and VEGF production of non-small cell lung cancer cells and enhances paclitaxel-mediated anti-tumor effect', *Lung Cancer*. Elsevier Ireland Ltd, 73(2), pp. 158–165. doi: 10.1016/j.lungcan.2010.11.010.

Squire, C. J. *et al.* (2005) 'Structure and inhibition of the human cell cycle checkpoint kinase, Wee1A kinase: An atypical tyrosine kinase with a key role in CDK1 regulation', *Structure*. Cell Press, 13(4), pp. 541–550. doi: 10.1016/j.str.2004.12.017.

Srivastava, R. C. *et al.* (1996) 'Evidence for the involvement of nitric oxide in cisplatin-induced toxicity in rats', *BioMetals*, 9(2), pp. 139–142. doi: 10.1007/BF00144618.

Steegmaier, M. *et al.* (2007) 'BI 2536, a Potent and Selective Inhibitor of Polo-like Kinase 1, Inhibits Tumor Growth In Vivo', *Current Biology*, 17(4), pp. 316–322. doi: 10.1016/j.cub.2006.12.037.

Di Stefano, V. *et al.* (2011) 'Knockdown of cyclin-dependent kinase inhibitors induces cardiomyocyte re-entry in the cell cycle', *Journal of Biological Chemistry*, 286(10), pp. 8644–8654. doi: 10.1074/jbc.M110.184549.

Stemke-Hale, K. *et al.* (2008) 'An integrative genomic and proteomic analysis of PIK3CA, PTEN, and AKT mutations in breast cancer', *Cancer Research*. Cancer Res, 68(15), pp. 6084–6091. doi: 10.1158/0008-5472.CAN-07-6854.

Stergachis, A. B. *et al.* (2013) 'Exonic transcription factor binding directs codon choice and affects protein evolution', *Science*. Science, 342(6164), pp. 1367–1372. doi: 10.1126/science.1243490.

Stewart, Z. O. E. A. and Leach, S. D. (1999) 'p21 Waf1 / Cip1 Inhibition of Cyclin E / Cdk2 Activity Prevents Endoreduplication after Mitotic Spindle Disruption', *Molecular and cellular biology*, 19(1), pp. 205–215.

Stillwell, E. E., Zhou, J. and Joshi, H. C. (2004) 'Human ninein is a centrosomal autoantigen recognized by CREST patient sera and plays a regulatory role in microtubule nucleation', *Cell Cycle*, 3(7), pp. 921–928. doi: 10.4161/cc.3.7.947.

Stratton, M. R., Campbell, P. J. and Futreal, P. A. (2009) 'The cancer genome', *Nature*. Nature, pp. 719–724. doi: 10.1038/nature07943.

Strimbu, K. and Tavel, J. A. (2010) 'What are biomarkers?', *Current Opinion in HIV and AIDS*, 5(6), pp. 463–466. doi: 10.1097/COH.0b013e32833ed177.

Sugano, T. *et al.* (2015) 'Inhibition of ABCB1 overcomes cancer stem cell-like properties and acquired resistance to MET inhibitors in non-small cell lung cancer', *Molecular Cancer Therapeutics*. American Association for Cancer Research Inc., 14(11), pp. 2433–2440. doi: 10.1158/1535-7163.MCT-15-0050.

Sun, H. *et al.* (2014) 'Aurora-A controls cancer cell radio- and chemoresistance via ATM/Chk2-mediated DNA repair networks', *Biochimica et Biophysica Acta - Molecular Cell Research*. Elsevier B.V., 1843(5), pp. 934–944. doi: 10.1016/j.bbamcr.2014.01.019.

Sun, N. K. *et al.* (2014) 'Transcriptomic profiling of taxol-resistant ovarian cancer cells identifies FKBP5 and the androgen receptor as critical markers of chemotherapeutic response', *Oncotarget*, 5(23), pp. 11939–11956. doi: 10.18632/oncotarget.2654.

Suraweera, A. *et al.* (2009) 'Functional role for senataxin, defective in ataxia oculomotor apraxia type 2, in transcriptional regulation', *Human Molecular Genetics*, 18(18), pp. 3384–3396. doi: 10.1093/hmg/ddp278.

Suwinski, P. *et al.* (2019) 'Advancing personalized medicine through the application of whole exome sequencing and big data analytics', *Frontiers in Genetics*. Frontiers Media S.A., p. 49. doi: 10.3389/fgene.2019.00049.

Swanton, C. (2012) 'Intratumor heterogeneity: Evolution through space and time', *Cancer Research*. Cancer Res, pp. 4875–4882. doi: 10.1158/0008-5472.CAN-12-2217.

T, S. *et al.* (1992) 'Expression of MDR1 and Glutathione S Transferase-Pi Genes and Chemosensitivities in Human Gastrointestinal Cancer', *Cancer*. Cancer, 69(4). doi: 10.1002/1097-0142(19920215)69:4<941::AID-CNCR2820690418>3.0.CO;2-H.

Taglialatela, A. *et al.* (2017) 'Restoration of Replication Fork Stability in BRCA1- and BRCA2-Deficient Cells by Inactivation of SNF2-Family Fork Remodelers', *Molecular Cell*. Elsevier Inc., 68(2), pp. 414-430.e8. doi: 10.1016/j.molcel.2017.09.036.

Takahashi, K. *et al.* (2001) 'DJ-1 positively regulates the androgen receptor by impairing the binding of PIASx alpha to the receptor.', *The Journal of biological chemistry*, 276(40), pp. 37556–63. doi: 10.1074/jbc.M101730200.

Takebe, N. *et al.* (2001) 'Retroviral transduction of human dihydropyrimidine dehydrogenase cDNA confers resistance to 5-fluorouracil in murine hematopoietic progenitor cells and human CD34+-

enriched peripheral blood progenitor cells', *Cancer Gene Therapy*. Cancer Gene Ther, 8(12), pp. 966–973. doi: 10.1038/sj.cgt.7700393.

Tanenbaum, M. E. *et al.* (2011) 'A complex of Kif18b and MCAK promotes microtubule depolymerization and is negatively regulated by aurora kinases', *Current Biology*, 21(16), pp. 1356–1365. doi: 10.1016/j.cub.2011.07.017.

Tang, J., Erikson, R. L. and Liu, X. (2006) 'Checkpoint kinase 1 (Chk1) is required for mitotic progression through negative regulation of polo-like 1 (Plk1)', *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 103(32), pp. 11964–11969. doi: 10.1073/pnas.0604987103.

Tang, L. *et al.* (2016) 'Plasma EGFR T790M ctDNA status is associated with clinical outcome in advanced NSCLC patients with acquired EGFR-TKI resistance', *Scientific Reports*, 6(1), pp. 1–9. doi: 10.1038/srep20913.

Taymaz-Nikerel, H. *et al.* (2018) 'Doxorubicin induces an extensive transcriptional and metabolic rewiring in yeast cells', *Scientific Reports*. Nature Publishing Group, 8(1), pp. 1–14. doi: 10.1038/s41598-018-31939-9.

Thauvin-Robinet, C. *et al.* (2014) 'The oral-facial-digital syndrome gene C2CD3 encodes a positive regulator of centriole elongation', *Nature Genetics*, 46(8), pp. 905–911. doi: 10.1038/ng.3031.

Thomas, H. D. *et al.* (2007) 'Preclinical selection of a novel poly(ADP-ribose) polymerase inhibitor for clinical trial', *Molecular Cancer Therapeutics*, 6(3), pp. 945–956. doi: 10.1158/1535-7163.MCT-06-0552.

Thompson, R. and Eastman, A. (2013) 'The cancer therapeutic potential of Chk1 inhibitors: How mechanistic studies impact on clinical trial design', *British Journal of Clinical Pharmacology*, 76(3), pp. 358–369. doi: 10.1111/bcp.12139.

Thorn, C. F. *et al.* (2011) 'Doxorubicin pathways: Pharmacodynamics and adverse effects', *Pharmacogenetics and Genomics*. NIH Public Access, 21(7), pp. 440–446. doi: 10.1097/FPC.0b013e32833ffb56.

Tochio, N. *et al.* (2010) 'Solution Structure of Histone Chaperone ANP32B: Interaction with Core Histones H3-H4 through Its Acidic Concave Domain', *Journal of Molecular Biology*. Elsevier Ltd, 401(1), pp. 97–114. doi: 10.1016/j.jmb.2010.06.005.

Tomasetti, C. and Vogelstein, B. (2015) 'Variation in cancer risk among tissues can be explained by the number of stem cell divisions', *Science*. American Association for the Advancement of Science, 347(6217), pp. 78–81. doi: 10.1126/science.1260825.

Tomoko, I. *et al.* (1999) 'Transcription factor Y-box binding protein 1 binds preferentially to cisplatin-modified DNA and interacts with proliferating cell nuclear antigen', *Cancer Research*, 59(2), pp. 342–346.

Topatana, W. *et al.* (2020) 'Advances in synthetic lethality for cancer therapy: Cellular mechanism

and clinical translation', *Journal of Hematology and Oncology*. BioMed Central Ltd, p. 118. doi: 10.1186/s13045-020-00956-5.

Toschi, L. *et al.* (2005) 'Role of gemcitabine in cancer therapy', *Future Oncology*. Future Medicine Ltd., 1(1), pp. 7–17. doi: 10.1517/14796694.1.1.7.

Tran, Q. *et al.* (2016) 'Targeting Cancer Metabolism-Revisiting the Warburg Effects', *Toxicol. Res*, 32(3), pp. 177–193. doi: 10.5487/TR.2016.32.3.177.

Trott, J. F. *et al.* (2003) 'Alternative splicing to exon 11 of human prolactin receptor gene results in multiple isoforms including a secreted prolactin-binding protein', *Journal of Molecular Endocrinology*. J Mol Endocrinol, pp. 31–47. doi: 10.1677/jme.0.0300031.

Trushina, E. *et al.* (2003) 'Microtubule destabilization and nuclear entry are sequential steps leading to toxicity in Huntington's disease', *Proceedings of the National Academy of Sciences of the United States of America*, 100(21), pp. 12171–12176. doi: 10.1073/pnas.2034961100.

Tsimberidou, A. M. *et al.* (2009) 'Ultimate fate of oncology drugs approved by the US food and drug administration without a randomized trial', *Journal of Clinical Oncology*. J Clin Oncol, pp. 6243–6250. doi: 10.1200/JCO.2009.23.6018.

Tsuruoka, S. *et al.* (2002) 'Functional analysis of ABCA8, a new drug transporter', *Biochemical and Biophysical Research Communications*. Biochem Biophys Res Commun, 298(1), pp. 41–45. doi: 10.1016/S0006-291X(02)02389-6.

Türk, Dóra & Szakács, G. (2009) 'Relevance of multidrug resistance in the age of targeted therapy', *Current opinion in drug discovery & development.*, 12(246–52).

Turner, N. C. *et al.* (2008) 'A synthetic lethal siRNA screen identifying genes mediating sensitivity to a PARP inhibitor', *EMBO Journal*. EMBO J, 27(9), pp. 1368–1377. doi: 10.1038/emboj.2008.61.

Turner, N., Tutt, A. and Ashworth, A. (2004) 'Hallmarks of "BRCAness" in sporadic cancers', *Nature Reviews Cancer*. Nature Publishing Group, pp. 814–819. doi: 10.1038/nrc1457.

Uhlen, M. *et al.* (2017) 'A pathology atlas of the human cancer transcriptome', *Science*, 357(660). doi: 10.1126/science.aan2507.

Vadlapatla, R. *et al.* (2013) 'Mechanisms of Drug Resistance in Cancer Chemotherapy: Coordinated Role and Regulation of Efflux Transporters and Metabolizing Enzymes', *Current Pharmaceutical Design*, 19(40), pp. 7126–7140. doi: 10.2174/13816128113199990493.

Vaidyanathan, A. *et al.* (2016) 'ABCB1 (MDR1) induction defines a common resistance mechanism in paclitaxel- and olaparib-resistant ovarian cancer cells', *British Journal of Cancer*. Nature Publishing Group, 115(4), pp. 431–441. doi: 10.1038/bjc.2016.203.

Veitch, Z. W. *et al.* (2019) 'Safety and tolerability of CFI-400945, a first-in-class, selective PLK4 inhibitor in advanced solid tumours: a phase 1 dose-escalation trial', *British Journal of Cancer*. Nature Publishing Group, 121(4), pp. 318–324. doi: 10.1038/s41416-019-0517-3.

Vendetti, F. P. *et al.* (2015) 'The orally active and bioavailable ATR kinase inhibitor AZD6738 potentiates the anti-tumor effects of cisplatin to resolve ATM-deficient non-small cell lung cancer in vivo.', *Oncotarget*, 6(42), pp. 44289–305. doi: 10.18632/oncotarget.6247.

Veskimäe, K. *et al.* (2018) 'Expression Analysis of Platinum Sensitive and Resistant Epithelial Ovarian Cancer Patient Samples Reveals New Candidates for Targeted Therapies', *Translational Oncology*. The Authors, 11(5), pp. 1160–1170. doi: 10.1016/j.tranon.2018.07.010.

Vidal, L. *et al.* (2016) 'Chlorambucil for the treatment of patients with chronic lymphocytic leukemia (CLL) – a systematic review and meta-analysis of randomized trials', *Leukemia and Lymphoma*. Taylor and Francis Ltd, 57(9), pp. 2047–2057. doi: 10.3109/10428194.2016.1154956.

Wagle, N. *et al.* (2011) 'Dissecting therapeutic resistance to RAF inhibition in melanoma by tumor genomic profiling', *Journal of Clinical Oncology*. J Clin Oncol, 29(22), pp. 3085–3096. doi: 10.1200/JCO.2010.33.2312.

Wagner, J. M. and Karnitz, L. M. (2009) 'Cisplatin-induced DNA damage activates replication checkpoint signaling components that differentially affect tumor cell survival', *Molecular Pharmacology*, 76(1), pp. 208–214. doi: 10.1124/mol.109.055178.

Wahba, H. A. and El-Hadaad, H. A. (2015) 'Current approaches in treatment of triple-negative breast cancer', *Cancer Biology and Medicine*. Cancer Biology and Medicine, pp. 106–116. doi: 10.7497/j.issn.2095-3941.2015.0030.

Wahlberg, E. *et al.* (2012) 'Family-wide chemical profiling and structural analysis of PARP and tankyrase inhibitors', *Nature Biotechnology*, 30(3), pp. 283–288. doi: 10.1038/nbt.2121.

Waller, A. P. *et al.* (2013) 'GLUT12 functions as a basal and insulin-independent glucose transporter in the heart', *Biochimica et Biophysica Acta - Molecular Basis of Disease*. Biochim Biophys Acta, 1832(1), pp. 121–127. doi: 10.1016/j.bbadis.2012.09.013.

Walton, M. I. *et al.* (2015) 'The clinical development candidate CCT245737 is an orally active CHK1 inhibitor with preclinical activity in RAS mutant NSCLC and Eμ-MYC driven B-cell lymphoma', *Oncotarget*, 7(3). doi: 10.18632/oncotarget.4919.

Wang, C. *et al.* (2019) 'Genome-wide CRISPR screens reveal synthetic lethality of RNASEH2 deficiency and ATR inhibition', *Oncogene*. Springer US, 38(14), pp. 2451–2463. doi: 10.1038/s41388-018-0606-4.

Wang, Hongyan *et al.* (2004) 'ATR affecting cell radiosensitivity is dependent on homologous recombination repair but independent of nonhomologous end joining', *Cancer Research*. American Association for Cancer Research, 64(19), pp. 7139–7143. doi: 10.1158/0008-5472.CAN-04-1289.

Wang, J. *et al.* (2017) 'Novel strategies to prevent the development of multidrug resistance (MDR) in cancer', *Oncotarget*, 8(48), pp. 84559–84571. doi: 10.18632/oncotarget.19187.

Wang, M., Wang, Y. and Zhong, J. (2015) 'Side population cells and drug resistance in breast cancer', *Molecular Medicine Reports*. Spandidos Publications, 11(6), pp. 4297–4302. doi:

10.3892/mmr.2015.3291.

Wang, Q. *et al.* (2017) 'Novel metrics to measure coverage in whole exome sequencing datasets reveal local and global non-uniformity', *Scientific Reports*. Nature Publishing Group, 7(1), pp. 1–11. doi: 10.1038/s41598-017-01005-x.

Wang, X., Zhang, H. and Chen, X. (2019) 'Drug resistance and combating drug resistance in cancer', *Cancer Drug Resistance*. OAE Publishing Inc. doi: 10.20517/cdr.2019.10.

Warren, N. J. H. and Eastman, A. (2020) 'in combination with DNA damage', *Oncogene*, 39(7), pp. 1389–1401. doi: 10.1038/s41388-019-1079-9.Comparison.

Weaver, B. A. (2014) 'How Taxol/paclitaxel kills cancer cells', *Molecular Biology of the Cell*, 25(18), pp. 2677–2681. doi: 10.1091/mbc.E14-04-0916.

Weinstein, J. N. (2013) 'Cancer Genome Atlas Pan-cancer analysis project', *Nat Genet*, 45(10), pp. 113–1120. doi: 10.3779/j.issn.1009-3419.2015.04.02.

Wendorff, Timothy J.; Schmidt, Bryan H.; Heslop, Pauline; Austin, Caroline A.; Berger, J. M. (2012) 'The Structure of DNA-Bound Human Topoisomerase II Alpha: Conformational Mechanisms for Coordinating Inter-Subunit Interactions with DNA Cleavage', *J Mol Biol*, 424(3–4), pp. 109–124. doi: 10.1038/jid.2014.371.

West, K. A., Castillo, S. S. and Dennis, P. A. (2002) 'Activation of the PI3K/Akt pathway and chemotherapeutic resistance', *Drug Resistance Updates*, 5(6), pp. 234–248. doi: 10.1016/S1368-7646(02)00120-6.

West, S., Gromak, N. and Proudfoot, N. J. (2004) 'Human 5' → 3' exonuclease Xm2 promotes transcription termination at co-transcriptional cleavage sites', *Nature*, 432(7016), pp. 522–525. doi: 10.1038/nature03035.

Wiegmans, A. P. *et al.* (2014) 'Rad51 supports triple negative breast cancer metastasis', *Oncotarget*. Impact Journals LLC, 5(10), pp. 3261–3272. doi: 10.18632/oncotarget.1923.

Wishart, D. S. (2006) 'DrugBank: a comprehensive resource for in silico drug discovery and exploration', *Nucleic Acids Research*, 34(90001), pp. D668–D672. doi: 10.1093/nar/gkj067.

De Witt Hamer, P. C. *et al.* (2011) 'WEE1 kinase targeting combined with DNA-damaging cancer therapy catalyzes mitotic catastrophe', *Clinical Cancer Research*. Clin Cancer Res, pp. 4200–4207. doi: 10.1158/1078-0432.CCR-10-2537.

Wohlhueter, R. M., McIvor, R. S. and Plagemann, P. G. W. (1980) 'Facilitated transport of uracil and 5-fluorouracil, and permeation of orotic acid into cultured mammalian cells', *Journal of Cellular Physiology*. J Cell Physiol, 104(3), pp. 309–319. doi: 10.1002/jcp.1041040305.

Wong, R. W., Blobel, G. and Coutavas, E. (2006) 'Rae1 interaction with NuMA is required for bipolar spindle formation', *Proceedings of the National Academy of Sciences of the United States of America*, 103(52), pp. 19783–19787. doi: 10.1073/pnas.0609582104.

Woo, J. S. *et al.* (2006) 'Structural and functional insights into the B30.2/SPRY domain', *EMBO Journal*, 25(6), pp. 1353–1363. doi: 10.1038/sj.emboj.7600994.

Wu, C.-P. (2013) 'Human ABCB1 (P-glycoprotein) and ABCG2 Mediate Resistance to BI 2536, a Potent and Selective Inhibitor of Polo-like Kinase 1', *Biochem Pharmacol*, 86(7), pp. 904–913. doi: 10.1038/jid.2014.371.

Wynne, P. *et al.* (2007) 'Enhanced repair of DNA interstrand crosslinking in ovarian cancer cells from patients following treatment with platinum-based chemotherapy', *British Journal of Cancer*, 97(7), pp. 927–933. doi: 10.1038/sj.bjc.6603973.

Xia, L. *et al.* (2018) 'Role of the NFκB-signaling pathway in cancer', *OncoTargets and Therapy*. Dove Medical Press Ltd., 11, pp. 2063–2073. doi: 10.2147/OTT.S161109.

Yang, F. *et al.* (2014) 'Doxorubicin, DNA torsion, and chromatin dynamics', *Biochimica et Biophysica Acta - Reviews on Cancer*. NIH Public Access, pp. 84–89. doi: 10.1016/j.bbcan.2013.12.002.

Yang, F., Kemp, C. J. and Henikoff, S. (2013) 'Doxorubicin enhances nucleosome turnover around promoters', *Current Biology*. Cell Press, 23(9), pp. 782–787. doi: 10.1016/j.cub.2013.03.043.

Yang, L. *et al.* (2008) 'Unbiased screening for transcriptional targets of ZKSCAN3 identifies integrin β4 and vascular endothelial growth factor as downstream targets', *Journal of Biological Chemistry*. J Biol Chem, 283(50), pp. 35295–35304. doi: 10.1074/jbc.M806965200.

Yang, S. *et al.* (2016) 'ANP32B deficiency impairs proliferation and suppresses tumor progression by regulating AKT phosphorylation', *Cell Death and Disease*, 7, pp. 1–9. doi: 10.1038/cddis.2016.8.

Yang, W. *et al.* (2013) 'Genomics of Drug Sensitivity in Cancer (GDSC): A resource for therapeutic biomarker discovery in cancer cells', *Nucleic Acids Research*, 41(D1), pp. 955–961. doi: 10.1093/nar/gks1111.

Yu, H. (2007) 'Chk1: A Double Agent in Cell Cycle Checkpoints', *Developmental Cell*, pp. 167–168. doi: 10.1016/j.devcel.2007.01.005.

Yuan, M.-L. *et al.* (2018) 'Inhibition of WEE1 Suppresses the Tumor Growth in Laryngeal Squamous Cell Carcinoma', *Frontiers in Pharmacology*, 9(September), pp. 1–10. doi: 10.3389/fphar.2018.01041.

Zandi, R. *et al.* (2007) 'Mechanisms for oncogenic activation of the epidermal growth factor receptor', *Cellular Signalling*. Cell Signal, pp. 2013–2023. doi: 10.1016/j.cellsig.2007.06.023.

Zaretsky, J. M. *et al.* (2016) 'Mutations associated with acquired resistance to PD-1 blockade in melanoma', *New England Journal of Medicine*. Massachussetts Medical Society, 375(9), pp. 819–829. doi: 10.1056/NEJMoa1604958.

Zhang, W. H. *et al.* (2008) 'DNA damage-induced S phase arrest in human breast cancer depends on Chk1, but G2 arrest can occur independently of Chk1, Chk2 or MAPKAPK2', *Cell Cycle*, 7(11), pp. 1668–1677. doi: 10.4161/cc.7.11.5982.

Zhou, F. *et al.* (2017) 'Recent advance in the pharmacogenomics of human Solute Carrier Transporters (SLCs) in drug disposition', *Advanced Drug Delivery Reviews*. Elsevier B.V., pp. 21–36. doi: 10.1016/j.addr.2016.06.004.

Zhou, J. *et al.* (2020) 'The Drug-Resistance Mechanisms of Five Platinum-Based Antitumor Agents', *Frontiers in Pharmacology*. Frontiers Media S.A., p. 343. doi: 10.3389/fphar.2020.00343.

Zhou, Y. *et al.* (2020) 'FGF/FGFR signaling pathway involved resistance in various cancer types', *Journal of Cancer*, 11(8), pp. 2000–2007. doi: 10.7150/jca.40531.

Zhu, L. and Chen, L. (2019) 'Progress in research on paclitaxel and tumor immunotherapy', *Cellular and Molecular Biology Letters*, 24(1). doi: 10.1186/s11658-019-0164-y.

Zhu, X. R. *et al.* (1999) 'Characterization of human Kv4.2 mediating a rapidly-inactivating transient voltage-sensitive K+ current', *Receptors and Channels*, 6(5), pp. 387–400.

De Zio, D., Cianfanelli, V. and Cecconi, F. (2013) 'New insights into the link between DNA damage and apoptosis', *Antioxidants and Redox Signaling*. Mary Ann Liebert, Inc., pp. 559–571. doi: 10.1089/ars.2012.4938.

Zou, L. and Elledge, S. J. (2003) 'Sensing DNA damage through ATRIP recognition of RPA-ssDNA complexes', *Science*. American Association for the Advancement of Science, 300(5625), pp. 1542–1548. doi: 10.1126/science.1083430.