University of Sussex

A University of Sussex PhD thesis

Available online via Sussex Research Online:

http://sro.sussex.ac.uk/

This thesis is protected by copyright which belongs to the author.

This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the Author

The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the Author

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given

Please visit Sussex Research Online for more information and further details

TDP2 as a biomarker of sensitivity to TOP2 targeting agents and as a novel therapeutic target

A thesis submitted to the University of Sussex for the degree of

Doctor of Philosophy

By

Ioanna Ntai

May 2017

Acknowledgements

UNIVERSITY OF SUSSEX

IOANNA NTAI

DOCTOR OF PHILOSOPHY IN BIOCHEMISTRY

TDP2 AS A BIOMARKER OF SENSITIVITY TO TOP2 TARGETING AGENTS AND AS A NOVEL THERAPEUTIC TARGET

SUMMARY

TDP2, a DNA phosphodiesterase that removes trapped topoisomerase 2 (TOP2) from 5'-DNA termini, is required for efficient repair of TOP2-induced DNA double-strand breaks (DSBs). Cellular depletion of TDP2 was shown to result in a substantially increased sensitivity to TOP2-induced DSBs and TOP2 poisons, such as etoposide, in various types of human cancer cell lines. In addition, over-expression of TDP2 has been shown to increase resistance to etoposide. Recent data suggest that expression levels of TDP2 vary greatly in different cancer cell lines. However, there are no reported studies addressing the possible role of TDP2 as a clinical predictor of anti-cancer therapy outcome, or wider studies correlating TDP2 over-expression with resistance to TOP2 poisons. TDP2 has the potential to be a good target for pharmacological inhibition potentially increasing tumour sensitivity to TOP2 poisons, particularly those that develop resistance during the course of treatment. There is already a lot of interest in the development of TDP2 inhibitors and the *in vitro* results are very promising with many possible small molecule inhibitors showing selectiveness in their target. In my thesis I aim to further establish the range of TDP2 and TOP2 mRNA and protein levels in a panel of lung and breast cancer cell lines. In addition, I will explore the possibility of a correlation between TDP2 protein levels or TOP2/TDP2 protein ratios and sensitivity to the TOP2 poison etoposide. This likely complex relationship will be further defined by possible mutation effects based on available literature and studies for the cancer cell lines accessible for this project. Furthermore, as etoposide has been tested in clinical trials not only alone but also in combination with other treatments, I aim to include other accessible drugs, either currently in use for cancer treatment or at a promising clinical trial stage, such as estradiol and PARP1 inhibitors.

There is a recent model, which suggests that induction of transcriptional programs by stimulating breast cancer cells with estrogens, or prostate cancer cells with androgens, can involve the formation of TOP2B mediated DSBs and the recruitment of DSB repair proteins. TOP2B is believed to be recruited with the estrogen/androgen receptor to regulatory sites on target genes. It is hypothesized that the formation of these DSBs by TOP2B could also be exploited therapeutically. In my thesis I aim to utilise a combination treatment that will first induce transcription with estradiol in breast cancer cells, which will cause transient TOP2B-mediated DSBs, and then transform those breaks into abortive breaks with etoposide. Cells will then be overwhelmed with DSBs and apoptosis will be promoted. This will also be explored in TDP2-depleted MCF7 breast cancer cells. Such a strategy could possibly find particular use in hormone dependent cancers, where other types of treatment have failed.

Contents

List of Figures

List of Tables

Abbreviations

APS = Ammonium persulfate

ATCC = American type culture collection

ATP = Adenosine triphosphate

BCP = 1-bromo-3-chloro-propane

BER = Base excision repair

bp = Base pair

BT474 = BT-474

Ca = Calcium

cDNA = Complementary DNA

CO_2 = Carbon dioxide

C-terminus = Carboxyl terminus

DAPI = 4',6-diamidino-2-phenylindole

ddH_2O = Double distilled water

dH_2O = Distilled water

DMEM = Dulbecco's Modified Eagle's Medium

DMSO = Dimethyl sulfoxide

DNA = Deoxynucleic acid

DNAse = Deoxyribonuclease

DSB = Double-strand break

DSBR = Double-strand break repair

DTT = Dithiothreitol

E.coli = *Escherichia coli*

E2 = Estradiol

EAPII = ETS1-associated protein II

EdU = 5-Ethynyl-2´-deoxyuridine

ERα = Estrogen receptor α

EtOH = Ethanol

FBS = Foetal bovine serum

g = Gram

GREB1 = Growth regulation by estrogen in breast cancer 1

h = hour

H2AX = Histone H2A variant X

H_2O = water

HCl = Hydrogen chloride

HR = Homologous recombination

kDa = Kilodalton

LB = Luria-Bertani medium

LN_2 = Liquid nitrogen

M = Molar

M phase = Mitosis

MCF7 = Michigan Cancer Foundation-7

Mg = Magnesium

min = Minute

ml = Millilitre

mM = Millimolar

MM361 = MDA-MB-361

MM453 = MDA-MB-453

MMR = Mismatch repair

Mn = Manganese

mRNA = Messenger RNA

mt = Mutant

MTT = Microculture Tetrazolium Assay

NER = Nucleotide excision repair

ng = Nanogram

NHEJ = Non-homologous end joining

nM = Nanomolar

N-terminus = Amino-terminus

p53 = Tumour protein 53

PAGE = polyacrylamide gel electrophoresis

PARP1 = Poly (ADP-Ribose) Polymerase 1

PBS = Phosphate buffered saline

PFA = Paraformaldehyde

qPCR = Quantitative polymerase chain reaction

RNA = Ribonucleic acid

RNAi = RNA interference

RNAse = Ribonuclease

RPLP0 = Ribosomal protein, large, P0

RPMI = Roswell Park Memorial Institute medium

RT-qPCR = Real-time quantitative polymerase chain reaction

S phase = Synthesis phase

SDS = sodium dodecyl sulphate

sec = Second

SSB = Single-strand break

SSBR = Single-strand break repair

TAM = Tamoxifen

TBP = TATA-binding protein

TBST = Tris-buffered saline and Tween 20

TDP1 = Tyrosyl-DNA phosphodiesterase 1

TDP2 = Tyrosyl-DNA phosphodiesterase 2

TEMED = Tetramethylethylenediamine

TFF1 = Trefoil factor 1

TNF = tumor necrosis factor

TOP1 = DNA topoisomerase 1

TOP2 = DNA topoisomerase 2

TOP2A = DNA topoisomerase 2-alpha

TOP2B = DNA topoisomerase 2-beta

TRAF = TNF receptor-associated factors

Tris = 2-Amino-2-hydroxymethyl-propane-1,3-diol

TTRAP = TRAF and TNF receptor-associated protein

V = Volt

v/v = Volume/volume

VP16 = Etoposide

w/v = Weight/volume

wt = Wild type

Zn = Zinc

ZR751 = ZR-75-1

μg = Microgram

μM = Micromolar

1. CHAPTER ONE - Introduction

1.1 DNA Damage

DNA is the repository of genetic information. Genomic integrity and stability are fundamental to the survival of every cell. The human genome is fragile and under a lot of pressure to maintain its integrity during exposure to both internal and external (environmental) sources of damage (Lindahl, 1993, Jackson and Bartek, 2009). Not repairing this damage can lead to harmful mutations and even disease. As the consequences of unrepaired DNA lesions are generally detrimental, the cells have a number of DNA repair responses and pathways to prevent accumulation of lethal DNA damage and ensure genome integrity and stability. Figure 1.1 shows in short the types of DNA damage and the repair process responsible for repairing them.
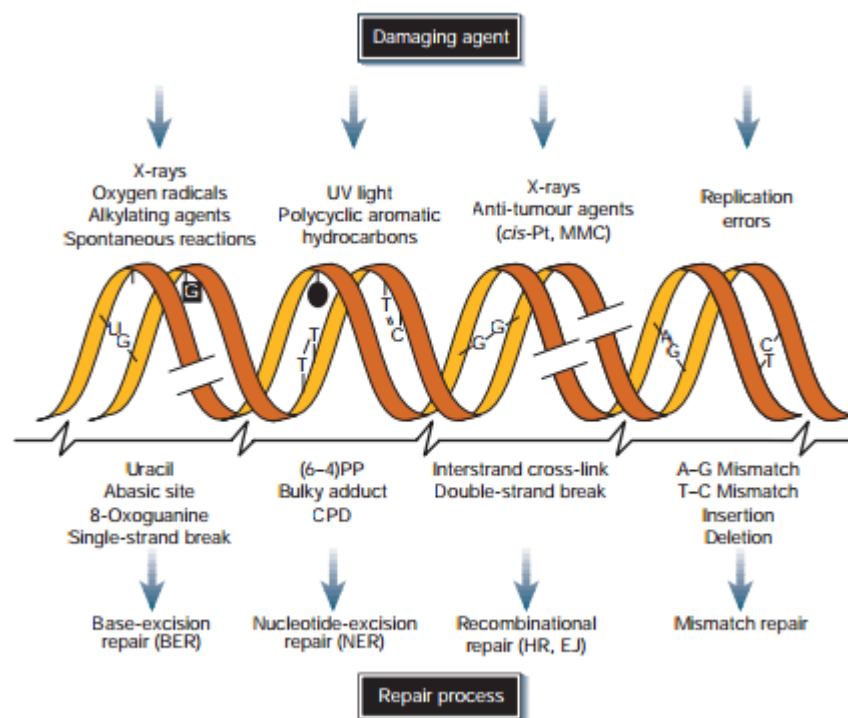
Figure 1.1: From top to bottom: Sources of DNA damage, examples of DNA lesions induced by these sources and most relevant DNA repair mechanisms for removing these lesions. (Adapted from (Hoeijmakers, 2001).

The external sources of damage fall under three main categories: ultraviolet radiation, ionising radiation and exogenous genotoxic agents. Ultraviolet radiation (commonly known as UV) is an electromagnetic radiation present in sunlight with a wavelength from 10nm to 400nm. It can potentially cause DNA lesions, such as cyclobutane pyrimidine dimers and 6-4 photoproducts that distort the DNA structure (Hoeijmakers, 2001). Ionising radiation, which can be in the form of gamma rays, X-rays or the higher ultraviolet part of the electromagnetic spectrum, can induce both single- and double-strand breaks to the DNA double helix, as well as base damage by producing hydroxyl (OH) radicals (Ward, 1975, Ward et al., 1987). In addition to radiation, DNA damage can result from exogenous genotoxic agents such as the polycyclic aromatic hydrocarbons in tobacco smoke (Clansy, 2008, Hoeijmakers, 2001).

Despite the multitude of external sources of DNA damage, there are many more internal sources. By-products of the cell's own metabolism can lead to oxidative DNA lesions (Beckman and Ames, 1997, Cadet et al., 2003, Azzam et al., 2012, Maynard et al., 2009). For example, endogenous reactive oxygen species (ROS) can be generated during the immune response, lipid peroxidation in peroxisomes and during the electron transfer chain (Cooke et al., 2003). Furthermore, certain chemical bonds in the DNA can disintegrate spontaneously under physiological conditions; for example nucleotide residues can hydrolyse leading to abasic sites and cytosine, adenine, guanine or 5-methylcytosine can de-aminate and result in uracil, hypoxanthine, xanthine and thymidine, respectively (Lindahl, 1993). Even the very process of DNA replication is prone to errors and damage, such as A-G and T-C mismatch, as well as insertions and deletions during a process called strand slippage (Pray, 2008).

Unrepaired DNA can lead to multiple problems of different severity for the cell. DNA replication or transcription can be halted, because polymerases cannot bypass DNA lesions. This can lead to cell-cycle arrest and/or cell death. There are links between bulky DNA adducts in smokers and increased lung and bladder cancer risk (Veglia et al., 2003). Accumulation of

point mutations contribute greatly to oncogenesis and can result to the loss of tumour-suppressor genes or the untimely activation of oncogenes (Hoeijmakers, 2001, Kastan and Bartek, 2004, Stratton et al., 2009). p53 point mutation is a well studied example (Goh et al., 1995). Another example is oxidative 8-hydroxyguanine that results from reactive oxygen species, a particularly mutagenic lesion that binds with cytosine or adenine with equal affinity, that can lead to G to T and C to A substitutions in the genome (Cheng et al., 1992).

There is a diverse spectrum of DNA damage categories. These include interstrand crosslink, intrastrand crosslink, DNA-protein crosslink, base mismatch, base modification, single-strand break (SSB) and double-strand break (DSB). Figure 1.2 depicts that spectrum and highlights the corresponding DNA repair pathway that cells use to cope with the damage.
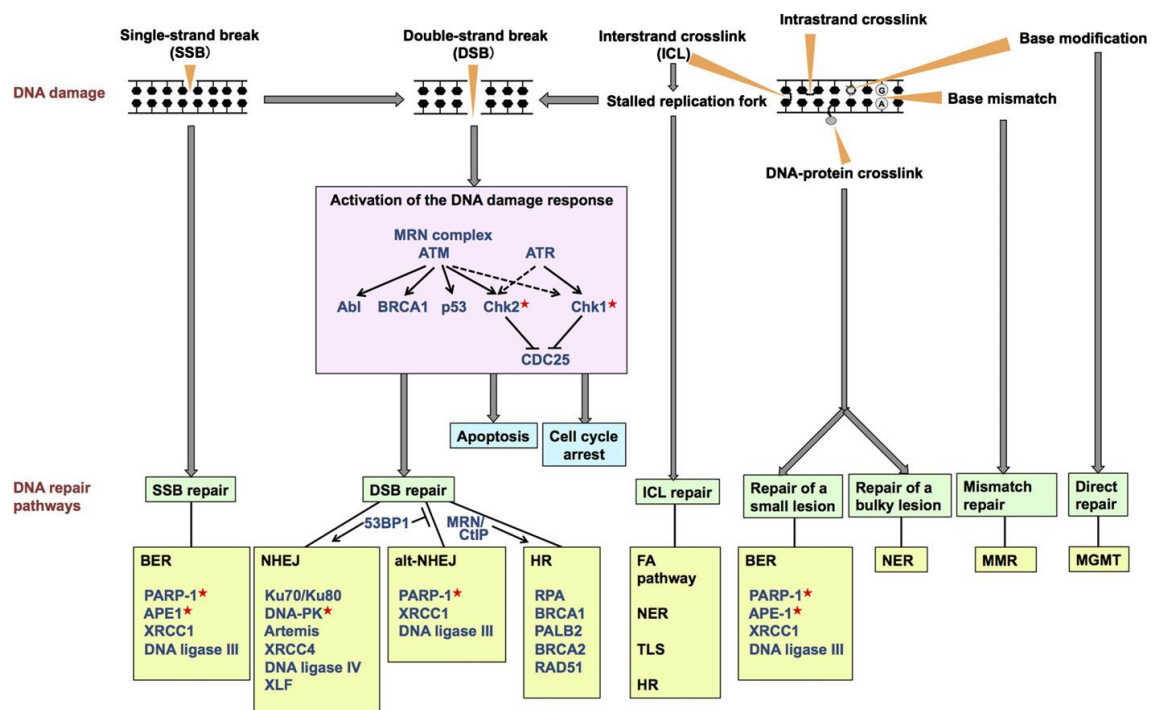
Figure 1.2: Overview of the wide spectrum of types of DNA damage, the major DNA damage response mechanisms and their key proteins (Hosoya and Miyagawa, 2014).

DNA SSBs are discontinuities in one of the DNA strands of the double helix and are the most common types of DNA lesions in the cell. In fact, they are so common that it is estimated that as many as 10,000 SSBs can arise in a single cell per day (Beckman and Ames, 1997, Caldecott,

2001, Lindahl and Nyberg, 1972, Lindahl, 1993). They usually co-occur with loss of a single nucleotide and damaged 5'- and/or 3'-termini at the site of the break (Caldecott, 2008). If not repaired rapidly, they pose a serious threat to genetic integrity. Even though they are not regarded as genotoxic themselves, they can be converted to DSBs by collision with the machinery of a replication or transcription fork (Caldecott, 2008, Hoeijmakers, 2001, Kouzminova and Kuzminov, 2006, Kuzminov, 2001). The cell has a mechanism to respond to such breaks, but an acute increase of SSB levels can saturate this pathway, leading to cell death (Caldecott, 2001). Under certain physiological conditions, high levels of SSBs can cause excessive activation of the SSB sensor protein poly (ADP-ribose) polymerase 1 (PARP1) (Caldecott, 2008, Heeres and Hergenrother, 2007, Moroni, 2008). This continuous activation can cause a deficiency in cellular NAD+ and ATP and/or release of apoptosis-inducing factor (AIF) from mitochondria (Moroni, 2008, Heeres and Hergenrother, 2007, Caldecott, 2008). A common source of SSBs is oxidative attack by endogenous reactive oxygen species (ROS). They can occur directly, by disintegration of the oxidized sugar or indirectly, during the DNA base-excision repair of oxidized bases at damaged, altered or abasic sites (Caldecott, 2008, Demple and DeMott, 2002, Hegde et al., 2008, Pogozelski and Tullius, 1998). Another internal source of SSBs is the abortive activity of DNA topoisomerase 1 (TOP1) (Caldecott, 2008).

DSBs are discontinuities in both DNA strands of the double helix and can arise from both internal and external DNA damage sources. They do not occur as often as SSBs, as it is estimated that only 10 spontaneous DSBs arise per cell per day, as seen in early passage primary human fibroblasts (Martin et al., 1985, Lieber et al., 2003, Lieber and Karanjawala, 2004). They are considered very serious and a single one can be responsible for the induction of apoptosis (Rich et al., 2000). In addition, they can lead to chromosomal rearrangements, such as translocations, deletions and mis-segregation of chromosomes during mitosis (Hoeijmakers, 2001). Ionising radiation (IR) is a major source of DSBs. The IR particles create localised clusters of ROS (Azzam et al., 2012). Similar to SSBs, they can be generated directly if

the attack of free radicals on the sugar-phosphate backbone of the DNA molecule results in the dissolution of two oxidised sugars in close proximity from one another on anti-parallel strands of the DNA (Ward et al., 1987, Azzam et al., 2012). Endogenous DSBs can be a result of various processes. As previously mentioned, they can be caused by the collision of unrepaired SSBs with the replication machinery, but also due to the mechanical stress on chromosomes during chromosomal segregation in mitosis and more importantly for the subject of this thesis, by the abortive activity of DNA topoisomerase 2 (TOP2), which transiently nicks both strands of the DNA during processes such as replication and transcription as a response to topological stress (Adachi et al., 2003, Kouzminova and Kuzminov, 2006, Kuzminov, 2001, Hoeijmakers, 2001).

1.2 DNA Repair Mechanisms

As seen in Figure 1.2, there is a variety of DNA repair mechanisms. The more prevalent pathways for the repair of interstrand crosslink, intrastrand crosslink, DNA-protein crosslink, base mismatch and base modification are base excision repair (BER), mismatch repair (MMR) and nucleotide excision repair (NER). These can be used in cases of damage to one of the single strands, where the other strand can be used as a template to guide the correction of the damaged one. Single-strand break repair (SSBR) is important in its own merit and the same is the case for the double-strand break repair (DSBR) response, which consists of homologous recombination (HR) and non-homologous end joining (NHEJ).

1.2.1 Base Excision Repair

Base excision repair (BER) is a DNA damage response mechanism that can repair DNA throughout the cell cycle. It is primarily responsible for removing small base lesions that do not distort the double helix, such as SSBs (Dianov and Hübscher, 2013, Hosoya and Miyagawa, 2014). This is important, because these damaged bases could otherwise lead to mutations by mis-pairing or DNA breaks during replication. BER works by deploying glycosylases, which recognize and remove specific damaged or inappropriate bases, forming apurinic–apyrimidinic

(AP) sites. These AP sites are cleaved by AP endonuclease I (APE1), which results in a SSB that can be processed by either short-patch (where a single nucleotide is replaced) or long-patch (where 2-10 new nucleotides are synthesized) BER (Liu et al., 2007).

1.2.2 Mismatch Repair

Mismatch repair (MMR) is a system that recognises and repairs inaccurate insertion, deletion and mis-incorporation of bases and has a ubiquitous presence (Hsieh and Yamane, 2008). These problems can occur during DNA replication, recombination and during the repair of other forms of DNA damage. In the example of DNA replication, newly synthesised daughter DNA strands will commonly include errors. MMR is strand-specific and can distinguish the daughter strand from the parental one (Hsieh and Yamane, 2008). MMR is initiated when mismatches are recognized by the highly conserved MutS protein. MutS and a second conserved protein, MutL, activate the endonucleolytic cleavage of the newly synthesised strand close to the region of the damage by a third protein, MutH (Hsieh and Yamane, 2008). This can result in removal of a few or up to several thousand base pairs from the newly synthesized DNA strand. DNA polymerase re-synthesises the missing part and DNA ligase seals the nick.

1.2.3 Nucleotide Excision Repair

Nucleotide excision repair (NER) is a very conserved DNA repair pathway. It is the major pathway for removal of bulky lesions generated by chemical and physical agents (Kamileri et al., 2012). For example, the common UV-generated cyclobutane pyrimidine dimers (CPDs) can only be repaired by NER in humans. NER operates in three distinct steps. Firstly, DNA damage is recognised and the six NER repair factors are assembled at the damage site in a random order, but in a co-operative manner. They form the so-called "closed complex", which is an unstable one. Subsequently, the damage-containing oligomer is removed after nicks have been

created on both of its sides. Finally, using the complementary strand as a template, the missing oligomer is re-synthesised and ligated (For this paragraph: (Reardon and Sancar, 2006).

1.2.4 Single-Strand Break Repair

As mentioned previously, SSBs can arise in different ways. Depending on the cause of the damage, the exact variation of the repair pathway that will be followed differs.

SSBs can arise in an indirect manner during BER by the APE1-led incision of an AP site or the lyase activity of a bi-functional DNA glycosylase, as mentioned previously. When that is the case, the SSBs are recognized by damage-specific proteins such as APE1, which then recruit Polb, XRCC1 and DNA Ligase IIIa, with the latter two forming a stable complex in order to complete the repair process (Hosoya and Miyagawa, 2014). After detection and end processing, the SSBs can be repaired by either short patch or long patch BER as discussed previously (Caldecott, 2008). PARP1 is not thought to play the same role in repair of such breaks, even though it has been shown to accumulate at sites of base damage (Durkacz et al., 1980).

When SSBs arise directly, the detection and end processing pathway is different. Direct SSBs can be the result of ROS-induced disintegration of oxidized deoxyribose, and during abortive TOP1 activity that creates a TOP1-linked SSB (Caldecott, 2008). The major proteins that play a role in repairing direct SSBs are PARP1, XRCC1, DNA Ligase IIIa, and APE1 (Hosoya and Miyagawa, 2014). PARP1 detects and binds to the SSBs and ensures rapid accumulation of downstream repair factors at the breaks (D'Amours et al., 1999, Kim et al., 2005). There is an abundance of published evidence that show that PARP1 interacts with, and is required for the rapid accumulation of the core SSBR protein X-ray cross complementing protein 1 (XRCC1) at sites of oxidative damage (Caldecott et al., 1996, El-Khamisy et al., 2003, Masson et al., 1998). PARP1 has high affinity for DNA breaks and following binding on the break, it catalyses the polymerisation of long branched chains and other target molecules (Benjamin and Gill, 1980,

Bramson et al., 1993, Weinfeld et al., 1997). It then autoribosylates and disassociates itself from the break (Zahradka and Ebisuzaki, 1982, Ferro and Olivera, 1982). It is then targeted by the enzyme poly(ADP-ribose) glycohydrolase (PARG) , which catalyses the degradation of the poly(ADP-ribose) polymers, leading to PARP1 molecule recycling (Cortes et al., 2004, Fisher et al., 2007). Inhibition, depletion or deletion of PARP1 has been shown to decrease the rate of SSBR (Fisher et al., 2007, Godon et al., 2008, Le Page et al., 2003).

1.2.5 Double-Strand Break Repair

Double-strand breaks (DSBs) are one of the most hazardous types of DNA damage, as mentioned previously. A single unrepaired DSB can lead to apoptosis, and inaccurate repair to deletions, chromosomal aberrations or even cancer (Khanna and Jackson, 2001). The accurate repair of these breaks is of vital importance to maintain genomic stability, which is why the cell has multiple pathways of dealing with these events (van Gent et al., 2001). There are two major pathways for repairing DSBs. These are homologous recombination (HR) and non-homologous end joining (NHEJ). They not only differ in their fidelity of DSB repair, but also in their requirement for a homologous template DNA (Mehta and Haber, 2014).

1.2.5.1 Homologous Recombination

Homologous recombination (HR) comprises a series of related pathways, responsible for repairing DSBs (Pâques and Haber, 1999, Hartlerode and Scully, 2009). It is an error-free pathway that is functional mainly during the late-S and G2 cell cycle phases due to the proximity of the sister chromatids, as well as during meiosis to produce new combinations of DNA sequences (Delacôte and Lopez, 2008, Sonoda et al., 2006). It is generally error-free, because it utilizes as a template for repair, the genetic information contained in the undamaged sister chromatid (Cannan and Pederson, 2016).

HR varies among different cell types, but it can be summed up into three steps: presynapsis, synapsis and postsynapsis. When a DSB occurs, the DNA ends around the break are recognised

by the heterotrimeric MRN complex, which consists of Mre11, Rad50 and Nbs1. The MRN complex resects the ends in a 5'-3' direction, which generates short 3' single-stranded overhangs with a 3'OH (Uziel et al., 2003). CtIP has also been shown to be required for the resection step (Huertas and Jackson, 2009). This resection is followed by the recruitment of additional proteins such as RPA, BRCA1, BRCA2, Rad51, Rad52, and Rad54 (Renodon-Cornière Axelle, 2013). RPA quickly binds the single-stranded overhang and Rad51 and Rad52 are recruited to the DSB; RPA and Rad52 help Rad51 to load onto the single-stranded end to form nucleoprotein filaments (van Attikum and Gasser, 2005). This is followed by strand invasion, where the filaments "invade" similar or identical DNA molecules that are intact, once they have identified a homologous DNA sequence (Forget and Kowalczykowski, 2010). This procedure involves the assistance of BRCA1, BRCA2 and the Rad51-like proteins XRCC2, XRCC3, RAD51B, RAD51C and RAD51D (Sy et al., 2009, San Filippo et al., 2006, Pellegrini et al., 2002, Zhang et al., 2009, Wong et al., 1997, Forget and Kowalczykowski, 2010). DNA polymerase η synthesises new DNA using the homologous region as a template and DNA Ligase I carries out the ligation to yield the Holliday junction, a four-way junction structure of intermediate nature (McIlwraith et al., 2005). There are three ways of resolving this structure and all of them result in the error-free repair of the DSB. Figure 1.3 depicts a schematic overview of HR.
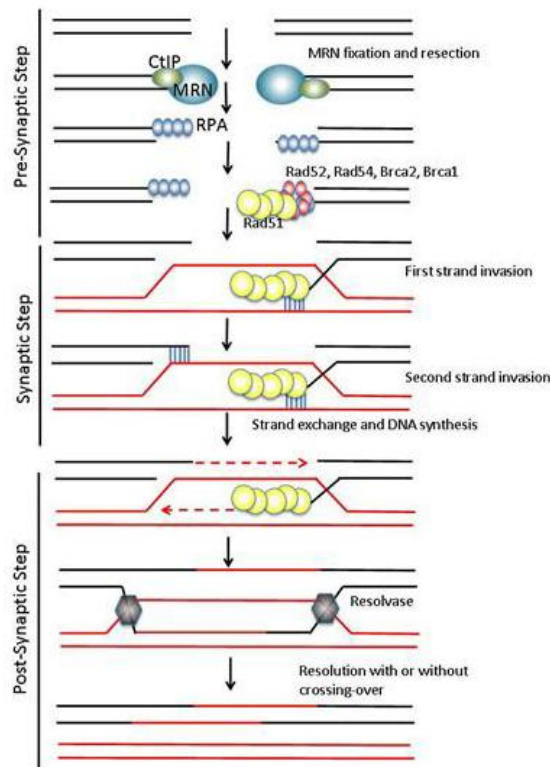
Figure 1.3: Schematic overview of how DNA double strand breaks are repaired by homologous recombination. The key factors involved are also depicted (Adapted from (Renodon-Cornière Axelle, 2013).

1.2.5.2 Non-Homologous End Joining

Non-homologous end joining (NHEJ) is an error-prone pathway which functions at all stages of the cell cycle, with an increase in activity in G1 (Hartlerode and Scully, 2009, Lieber, 2010, Mahaney et al., 2009). Unlike HR it is error-prone, because it involves elimination of DSBs by direct ligation of the broken ends. NHEJ does not require a homologous sequence for accurate repair, which is why its existence is so important during G0, G1 and early-S phase, when sister chromatids are not available to be used as templates (Delacôte and Lopez, 2008). The main protein factors involved in NHEJ are XRCC4-like factor (XLF), XRCC4/DNA Ligase IV and DNA proteinkinase (DNA–PK), with the latter being comprised of a protein kinase catalytic subunit (DNA–PKcs) and a Ku70/80 heterodimer (Davis et al., 2014, Grundy et al., 2016). Figure 1.4 depicts a schematic overview of NHEJ.
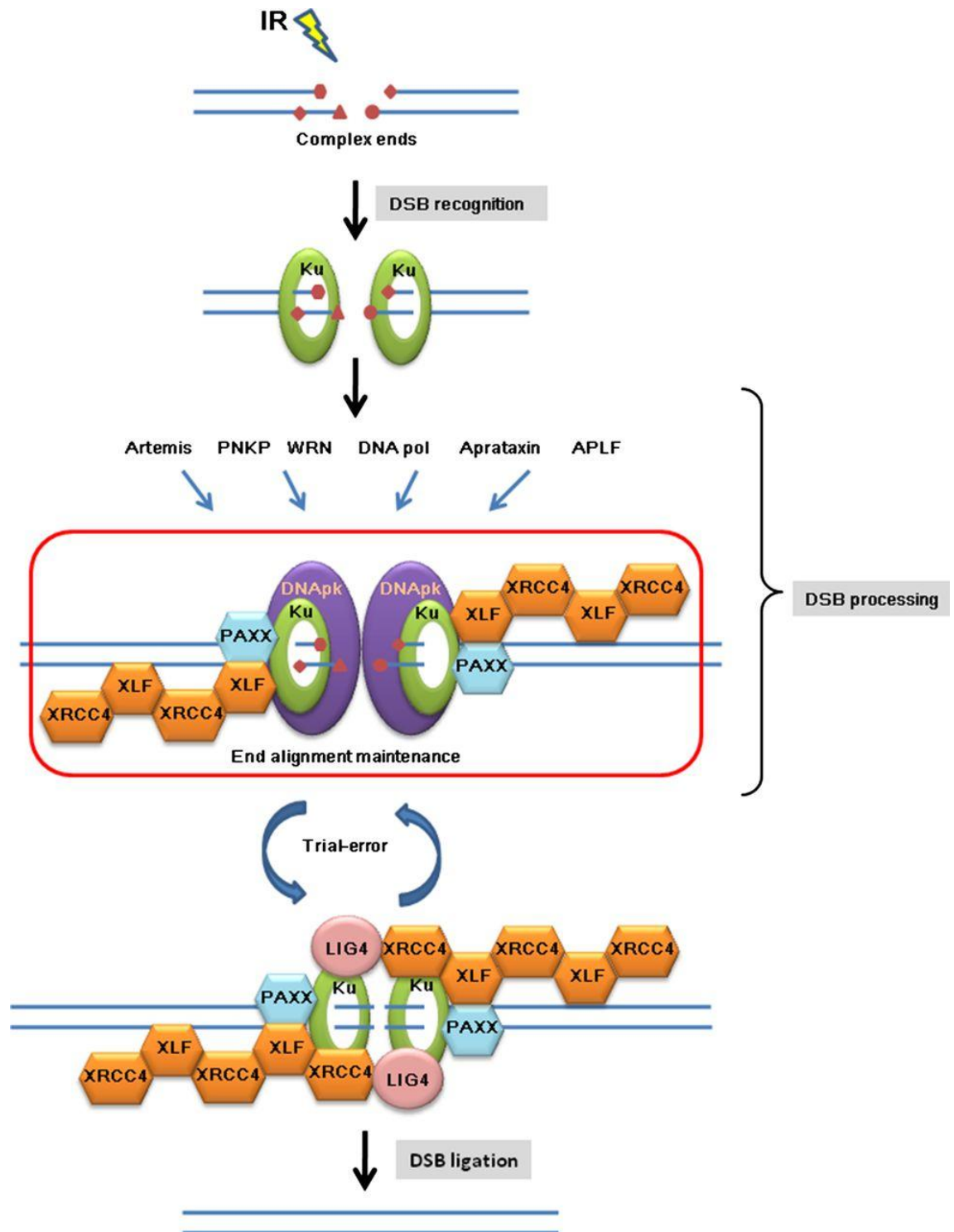
Figure 1.4: Schematic overview of the repair of double strand breaks by non-homologous end joining. Ku70/80 recognizes the break and recruits other proteins of this pathway, including DNA-PKcs, XRCC4, XLF and PAXX. These proteins form a stable complex that preserves the end alignment, and promotes end-processing and ligation likely by 'trial-and-error' (Yang et al., 2016).

NHEJ is initiated by the binding of Ku70/80 heterodimers to both ends of the DSB (Mimori and

Hardin, 1986). Once bound, the Ku70/80 heterodimer forms a ring-like structure that encircles

3-4bp of DNA and translocates inwards, making the DSB ends more accessible to other repair proteins, such as DNA dependent protein kinase (DNA-PK), which is then recruited to the site and forms a stable complex in the presence of the DNA ends (Walker et al., 2001, Yoo and Dynan, 1999, Menon and Povirk, 2016). This results in the interaction of Ku70/80 with DNA polymerases μ and λ, the XRCC4/DNA Ligase IV complex, and the XRCC4 paralogs XLF and PAXX (Chen et al., 2000, Ochi et al., 2015, Menon and Povirk, 2016). If the DNA ends are complementary and undamaged, they can be re-joined by the XRCC4/DNA Ligase IV (Reynolds et al., 2012).

In the case of non-complementary or damaged ends, ligation cannot progress without end processing to restore conventional ligatable 3'OH and 5'P. When that is the case, a group of enzymes including PNKP, APLF, Artemis and APTX are recruited for the process of the structures surrounding the DSB, before ligation (Riballo et al., 2004). Artemis, an endonuclease, is phosphorylated by DNA-PKcs at the C-terminal motif of residues 399–404 and is recruited to the DSB by the same factor (Soubeyrand et al., 2006, Ma et al., 2002). Ku70/80 does not bind directly to Artemis but there is evidence that Artemis' function, once recruited, is Ku70/80 dependent, further establishing Ku70/80's importance in NHEJ (Chang et al., 2015, Yang et al., 2016). Many publications have shown that Artemis is involved in multiple aspects of NHEJ, including performing endonucleolytic nicks of damaged overhangs, but also interacting with DNA Ligase IV in facilitating NHEJ (Malu et al., 2012, De Ioannes et al., 2012)

Once end processing and the gap filling are completed, the DNA gets ligated by DNA Ligase IV in a complex with XRCC4 (Critchlow et al., 1997, Grawunder et al., 1997, Grawunder et al., 1998). XRCC4 has no enzymatic activity itself, but acts as a scaffolding protein to facilitate the recruitment of repair factors to the DSB (Mari et al., 2006). Further interactions with Ku70/80, DNA-PK, APTX (aprataxin) and APLF (PNK-like factor) mediate the recruitment and stability of XRCC4/DNA Ligase IV (Rulten et al., 2011, Rulten et al., 2008, Mari et al., 2006, Mehrotra et al.,

2011, Macrae et al., 2008, Iles et al., 2007). XLF (XRCC4–like factor) is also recruited to the DSB by Ku70/80 and stimulates the DNA ligation by promoting the re-adenylation of DNA Ligase IV following end joining (Yang et al., 2016, Ahnesorg et al., 2006, Hammel et al., 2011, Gu et al., 2007).

1.3 DNA Topoisomerases

The DNA is a very large molecule that is usually packed very densely in the nucleus of the cell. It is negatively supercoiled in all species, both prokaryotes and eukaryotes (McClendon and Osheroff, 2007, Baranello et al., 2012, Wang, 1996, Schvartzman and Stasiak, 2004). This makes the complementary DNA strands more accessible during cell processes such as replication and transcription. But once replication or transcription has been initiated, helicases are only able to separate the two DNA strands and not unwind them. Therefore, once the different components of the replication or transcription machinery start to move along the DNA strand, it causes topological changes, such as positive supercoiling and tangles. This makes it very difficult for the two DNA strands to be separated, something that could potentially interfere with essential cell processes (For this paragraph: (McClendon and Osheroff, 2007, Wang, 1996, Schvartzman and Stasiak, 2004).

There are two types of topoisomerases, TOP1 (topoisomerase 1) and TOP2 (topoisomerase 2) with two subclasses of TOP2, TOP2A and TOP2B, which have been found to be crucial to chromosome metabolism in processes such as transcription and replication (Champoux, 2001, Wang, 2002). Both have also been found to be involved in the transient cleavage of DNA (Salerno et al., 2010). TOP1 is a ubiquitous nuclear enzyme, which is responsible for catalysing the relaxation of superhelical DNA, while generating a transient SSB in the DNA via cycles of cleavage and re-ligation (Salerno et al., 2010). TOP2 is also ubiquitous and is involved in the ATP-dependent induction of DSBs (Salerno et al., 2010).

When these breaks occur near sites of endogenous DNA damage or while topoisomerase poisons are present, the result can be abortive topoisomerase-induced DNA strand breaks, which feature covalent linkage of the enzyme to the DNA termini by a 3'- or 5'-phosphotyrosyl bond (Li and Liu, 2001, Pourquier and Pommier, 2001). These have been implicated in several human diseases, such as hereditary and neurodegenerative disease (El-Khamisy et al., 2005, Takashima et al., 2002, Yang et al., 1996), cancer and chromosomal instability (Deweese and Osheroff, 2009, Nitiss, 2009b), as well as having an impact on the clinical efficacy of anti-tumour poisons (Li and Liu, 2001, Nitiss, 2009b, Pommier, 2006). Therefore, the importance of liberating DNA termini from trapped topoisomerases and repairing the single- or double-strand breaks is apparent.

1.3.1 The role of type 2 topoisomerases in DNA double-strand break repair

For the purpose of this thesis, this part of the introduction will focus on TOP2. As previously mentioned, eukaryotic cells express two types of topoisomerase 2, TOP2A and TOP2B, which are encoded by different genes, share about 70% of sequence similarity and have molecular weights of 170kDa and 180kDa respectively (Wang, 2002, McClendon and Osheroff, 2007). Their function is to alter the topological properties of DNA, by introducing transient DNA DSBs and allow another DNA strand (also known as transport or T-segment) to pass through the gap (known as gate or G-segment), in order to relieve supercoiling and tangled DNA tension. TOP2 is also critical for recombination, proper chromosome structure and the separation of daughter chromosomes (Wang, 1996, Schvartzman and Stasiak, 2004, Wang, 2002, McClendon and Osheroff, 2007). TOP2 is formed of homodimers, which is crucial as it allows the enzymes to form the gate, which the T-segment can use to pass through the TOP2-DNA complex. Once the T-segment has passed through the gate, the DNA is re-ligated by TOP2 (Deweese and Osheroff, 2009, Pendleton et al., 2014). The central domain of TOP2 contains the active site tyrosine, which is critical for both the cleavage of the DNA and re-ligation. The N-terminal domain of TOP2 contains the site of ATP binding and hydrolysis, which is required for the

quicker translocation of the T-segment through the gate (Champoux, 2001, Baird et al., 1999).

The C-terminal domain is different between the two types of TOP2 and it is the least studied

part of the enzyme (Pendleton et al., 2014). TOP2 also requires a bivalent cation, Mg^{2+}, to

complete the chemistry of its enzymatic activity, in addition to binding the DNA (Deweese and

Osheroff, 2010). A covalent phosphotyrosine bond is generated between the 5'-terminus and

TOP2, in order to maintain genomic integrity (Sander and Hsieh, 1983); (Pendleton et al.,

2014). The resulting structure is known as the ''cleavage complex''. TOP2A and TOP2B

enzymatic activity is very similar, but they have different physiological functions.
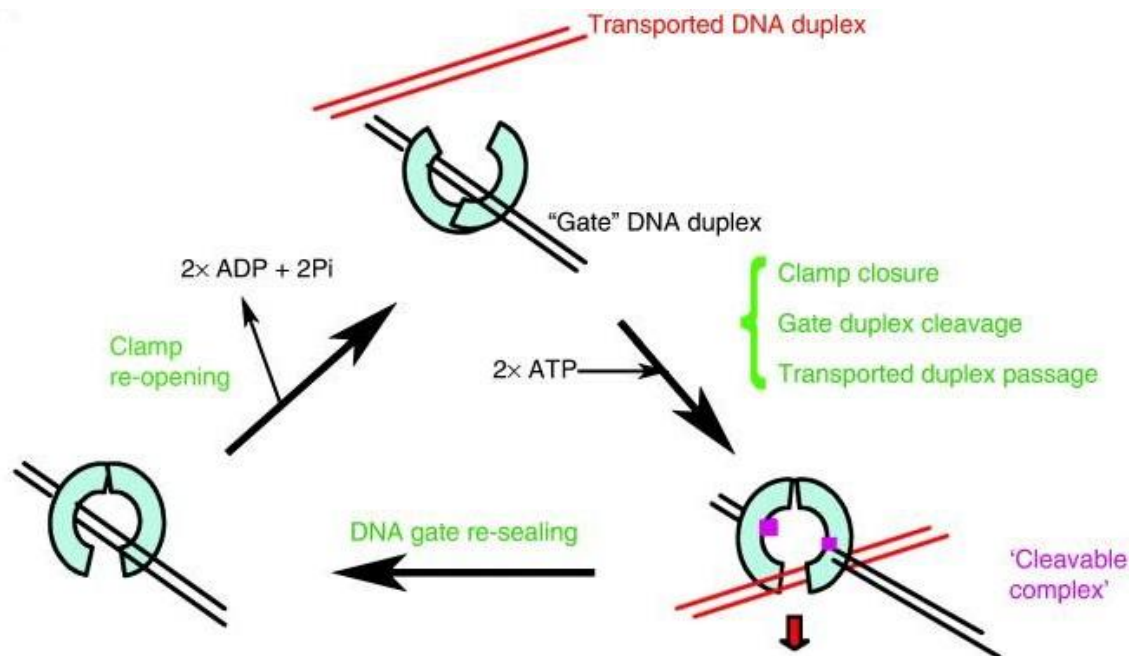
Figure 1.5: TOP2 catalytic cycle. The TOP2 homodimer (light blue semi-circles) binds the gate DNA. It then changes the spatial arrangement on ATP binding, while introducing a transient DSB in the gate DNA. TOP2 remains bound on the DSB and it transports the T-segment through it. This represents the 'cleavage complex'. The transient covalent bonds between the 5' phosphates on either side of the DSB and the tyr835 of each TOP2 subunit are depicted as pink squares. The DSB is then resealed. Hydrolysis and release of ATP is necessary for clamp re-opening (Germe et al., 2009).

TOP2A presence is crucial for actively proliferating cells, where it is found in abundance and is

upregulated (Heck and Earnshaw, 1986, Hsiang et al., 1988). Its concentration increases during

the cell cycle and peaks during G2/M (Heck et al., 1988, Kimura et al., 1994). TOP2A has been

associated with replication forks and during the duration of mitosis, it can be found tightly

bound on the chromosomes (Deweese and Osheroff, 2009, Pendleton et al., 2014). In addition, TOP2A is less frequently found in non-proliferating and differentiated cells (Pendleton et al., 2014).

In contrast, TOP2B concentration does not depend on proliferation and it can be found in high levels in most types of cells. Unlike TOP2A, during mitosis TOP2B dissociates from the chromosomes (McClendon and Osheroff, 2007, Deweese and Osheroff, 2009, Pendleton et al., 2014). Even though TOP2B is not as essential as TOP2A and it cannot compensate for the loss of TOP2A in mammalian cells (Grue et al., 1998), it is still required for proper neural development and function in mammals (Gómez-Herreros et al., 2014, Grue et al., 1998). Most importantly, there is emerging data that has linked TOP2B to the transcription of hormonally regulated genes (Ju et al., 2006, Cowell et al., 2012), something that will be discussed in more detail later.

1.4 The role of tyrosyl-DNA phosphodiesterase 2 in the restoration of 5'-phosphate termini at DNA double-strand breaks

The enzyme that cleaves 3'-phosphotyrosyl bonds had previously been identified as tyrosyl-DNA-phosphodiesterase 1 (TDP1) (Nitiss et al., 2006). On the other hand, the complementary enzyme that cleaved 5'-phosphotyrosyl bonds was not identified until a few years later, even though the effect of DNA DSBs containing such termini had been implicated in chromosome instability and cancer for some time.

In 2009, Cortes-Ledesma et al. identified an enzyme in human cells that efficiently restores 5'-phosphate termini at DNA DSBs in preparation for DNA ligation (Cortes Ledesma et al., 2009). This enzyme has been previously reported as TTRAP and EAPII and is a member of the Mg^{2+}/Mn^{2+}-dependent family of phosphodiesterases (Cortes Ledesma et al., 2009). As TTRAP (TRAF and TNF receptor-associated protein), it was identified through association with the cytoplasmic domain of CD40, TNF (tumor necrosis factor) receptor-75 and TRAFs (TNF

receptor-associated factors) (Pype et al., 2000). Through this role, it inhibits NFκB activation (Pype et al., 2000). It has also been linked to cancer cell proliferation (Xu et al., 2008). In addition, as EAPII (ETS1-associated protein II), it was identified as a transcription factor involved in tumorigenesis and metastasis (Hahne et al., 2008, Seth and Watson, 2005, Pei et al., 2005), through its involvement in the transcription activity of ETS1 (Pei et al., 2003). On the other hand, Cortes-Ledesma et al. showed that cellular depletion of this enzyme resulted in increased susceptibility and sensitivity to TOP2-induced DNA DSBs (Cortes Ledesma et al., 2009). Therefore, it was appropriately named tyrosyl-DNA-phospodiesterase 2 (TDP2) (Cortes Ledesma et al., 2009).

Sensitivity to TOP2 targeting agents, such as doxorubicin and etoposide (also known as VP16), as well as the catalytic inhibitor ICRF-193 in cells lacking TDP2 has also been shown (Cortes Ledesma et al., 2009, Zeng et al., 2011, Gómez-Herreros et al., 2014). More recent work from the same group has also revealed that TDP2 knock-out mice demonstrate lymphoid toxicity, severe intestinal damage and increased genome instability in the bone marrow, when treated with the TOP2 targeting agent etoposide (Gomez-Herreros et al., 2013). Furthermore, in a study published recently, it was shown that cells from affected individuals demonstrating intellectual disability, seizures and ataxia were very sensitive to TOP2-induced DSBs (Gómez-Herreros et al., 2014). The same study demonstrated that depleting human cells of TDP2 leads to inhibition of TOP2-dependent gene transcription, something the authors also observed in mouse post-mitotic neurons following abortive TOP2 activity (Gómez-Herreros et al., 2014). Finally, in was also demonstrated that TDP2 is crucial for maintaining regular levels of many gene transcripts in the developing mouse brain (Gómez-Herreros et al., 2014).

TDP2 is an enzyme that demonstrates specificity for phosphotyrosyl bonds (Cortes Ledesma et al., 2009, Gao et al., 2012). It is able to remove the tyrosine moiety from the 5'-end of the DNA, leaving a phosphate group ready for re-ligation (Cortes Ledesma et al., 2009).

Furthermore, it has been found to hydrolyze 5'-tyrosine-DNA adducts that mimic abortive

TOP2 cleavage complexes, while its preferred substrate was shown to be single-stranded or

double-stranded DNA with a four-base pair overhang, consistent with the known structure of

TOP2 cleavage complex (Gao et al., 2012). The specificity constant kcat/Km of TDP2 for single-

stranded 5'-tyrosyl-DNA has been measured as 4×10^5 s^{-1}m^{-1} (Gao et al., 2012). In addition to

Mg^{2+} and Mn^{2+}, TDP2 has been found to be weakly active with Ca^{2+} or Zn^{2+} (Gao et al., 2012).

Furthermore, TDP2 appears as up-regulated in p53 mutant tumour cells (Do et al., 2012).

Crystal structures of TDP2 for mouse, *Caenorhabditis elegans* and zebra fish have been

published (Schellenberg et al., 2012, Shi et al., 2012). These structures show a catalytic site

formed of four residues chelating a magnesium ion (Schellenberg et al., 2012, Shi et al., 2012).

In addition, the mouse and the zebra fish structures explain that a minimum of two single-

stranded, unpaired nucleotides are necessary to access the catalytic site (Schellenberg et al.,

2012, Shi et al., 2012, Pommier et al., 2014). In Figure 1.6 the structure of the human TDP2

catalytic domain can be seen (Hornyak et al., 2016).
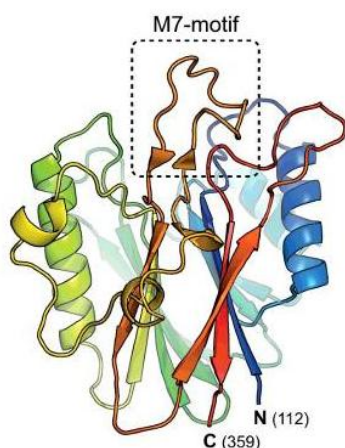
Figure 1.6: X-ray crystal structure of hTDP2^{CAT}. The molecular representation is coloured from blue to red, from the visible N- to C-terminus. Also featured is the position of the M7-motif (Hornyak et al., 2016).

As mentioned previously, DSBs can be repaired either by NHEJ or HR (Shrivastav et al., 2008).

TDP2 appears to 'function in a NHEJ-mediated and HR-independent pathway for the repair of

TOP2-induced DSBs' (Gomez-Herreros et al., 2013). It has been previously shown that TDP2

liberates DSB termini of the TOP2 cleavage complex by generating 5'phosphate and 3'-hydroxyl termini 'clean' DSBs with 4-bp overhangs that serve as direct substrate for end-joining by Ku70/80 and DNA Ligase IV, two of the most important NHEJ components, suggesting an error-free NHEJ mechanism (Fig.1.6) (Gomez-Herreros et al., 2013, Hoa et al., 2016).

Furthermore, it has been shown that deleting TDP2, Ku70 or TDP2 and Ku70 simultaneously in DT40 cells increased sensitivity to etoposide, but the $TDP2^{-/-/-}/Ku70^{-/-}$ double mutant demonstrated the same sensitivity as the $Ku70^{-/-}$ single mutant (Zeng et al., 2011). If cleavage complexes repaired by TDP2 were preferentially repaired by NHEJ, then loss of TDP2 would be followed by enhanced repair by HR (Nitiss and Nitiss, 2013). A recent study showed this in the form of increased Rad51 foci in TDP2-deficient cells treated with etoposide compared to wild type cells, as well as an increase in etoposide-induced sister chromatid exchanges, a molecular hallmark of HR (Gomez-Herreros et al., 2013).

Another interesting observation about TDP2 is that overexpression significantly reverses genome instability and mortality of Mre11 mutant TK6 human lymphoblastoid cells (Hoa et al., 2016). The same paper also suggests that Mre11 and TDP2 operate in separate pathways for the repair of etoposide-induced breaks. Moreover, it was also shown that the heterozygous mutant $MRE11^{+/-}$, but not the heterozygous $TDP2^{+/-}$ cells were more sensitive to etoposide than wild type cells (Hoa et al., 2016). Interestingly, $TDP2^{-/-}$ cells showed no significant genome instability unlike $MRE11^{-/H129N}/TDP2^{-/-}$ cells, suggesting a functional redundancy between the nuclease activity of Mre11 and TDP2 in terms of genome stability (Hoa et al., 2016).
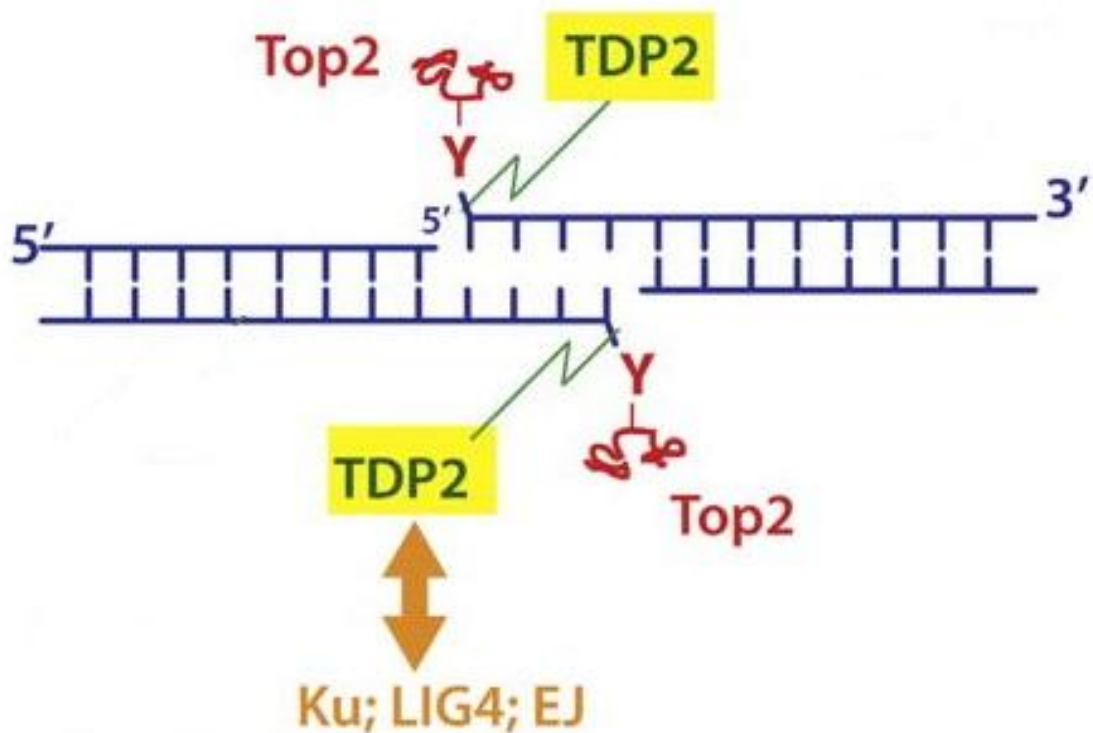
Figure 1.7: TDP2 pathway for removal of abortive TOP2 cleavage complexes. TDP2 hydrolyzes the tyrosyl-DNA links following proteolysis and/or denaturation of TOP2. Non-homologous end joining enzymes Ku and Ligase IV are essential downstream of TDP2 to complete DNA repair (Pommier et al., 2014).

1.5 Targeting DNA repair proteins in cancer therapy

Chemotherapy is one of the main strategies used to combat cancer. It is so named because it utilises one or more anti-cancer drugs or chemotherapeutic agents. Most of these agents are designed to kill cancer cells by inducing DNA damage. As discussed earlier, DNA damage is recognised and repaired by a variety of DNA damage response mechanisms. If the DNA damage response is successful, the cancer cell will survive. The increasing understanding of the DNA repair mechanisms has allowed the use of target-specific drugs that take advantage of certain aspects of cancer cells. For example, in cancer cells there are usually specific DNA repair factors that are upregulated and those can serve as biomarkers of therapeutic outcome. Inhibition of the DNA repair factors is a logical strategy to increase the effect of anti-cancer drugs, which is why it has been the subject of a lot of research groups and pharmaceutical

companies for several decades. There is a large number of inhibitors that target a variety of DNA damage response proteins at different stages of clinical trials at this moment or in the past (Hosoya and Miyagawa, 2014). These include drugs targeting PARP1 and TOP2.

1.5.1 Topoisomerase 2 inhibitors

There are several classes of drugs that target TOP2. Some of those drugs work by increasing the levels of cleavable complexes, while other drugs inhibit the catalytic site of the protein, obstructing the re-ligation of the DNA; this has led to the categorisation of TOP2 drugs as TOP2 poisons or catalytic inhibitors (Rocha et al., 2016, Pommier, 2013, Nitiss, 2009b, Nitiss, 2009a). The TOP2 anticancer agents that are used more commonly are anthracyclines (doxorubicin and daunorubicin), mitoxantrone and etoposide (Pommier et al., 2010). These drugs are considered TOP2 poisons as their effect is to increase levels of TOP2/DNA cleavage complexes (Rocha et al., 2016).

Etoposide is a semi-synthetic derivative of podophyllotoxin, a product of extraction from the mandrake root *Podophyllum peltatum*. It binds to the TOP2/DNA complex preventing the re-ligation of the DNA strands and causes an abortive break (Fleming et al., 1989). Depending on the molar ratio between etoposide and TOP2, the formation of both SSBs and DSBs is possible (Bromberg et al., 2003). This leads to accumulation of DSBs, which in turn inhibits DNA replication and/or transcription, leading to apoptosis. It acts primarily in G2 and S phases of the cell cycle. It has been used in cancer treatment for over 20 years for a variety of cancers. TOP2 poisons are used frequently for lymphomas and lung, ovarian, breasts and colorectal cancers (Pommier, 2013). It is not uncommon for patients to develop resistance to etoposide and there is data that shows a link between that resistance and TDP2 (Li et al., 2011b, Rocha et al., 2016).

There is evidence that high etoposide concentrations can trigger the process of apotosis, which involves caspases (a family of cysteine-dependent proteases) and progresses via the

cytocrome c/caspase 9 pathway (Riedl and Shi, 2004). The pathway in short involves the release of cytocrome c into the cytoplasm from the mitochondria. It then binds to the apoptotic protease activating factor 1 (apaf1) forming the apoptosome. The apoptosome binds and cleaves procaspase 9 leading to its activation. The activated caspase 9 is then able to in turn activate other regulatory caspases, such as caspase 7 and caspase 3 triggering further apoptotic processes (Fig.1.8) (Montecucco et al., 2015). In addition, etoposide has been shown to have a similar effect in non-replicating human T-lymphocytes, therefore interfering with transcription. It activates DNA damage response and induces phosphorylation of ATM and its substrates, H2AX and p53. The pro-apoptotic PUMA protein, a p53 upregulated modulator of apoptosis and the caspases are also then activated, leading to apoptosis (Korwek et al., 2012, Montecucco et al., 2015). It is also interesting to mention that PARP1 is cleaved in two fragments during apoptosis, something that has become a useful hallmark of topoisomerase inhibitor-induced apoptosis, and data shows that caspases 3 and 7 are responsible (Germain et al., 1999, Curtin, 2005).
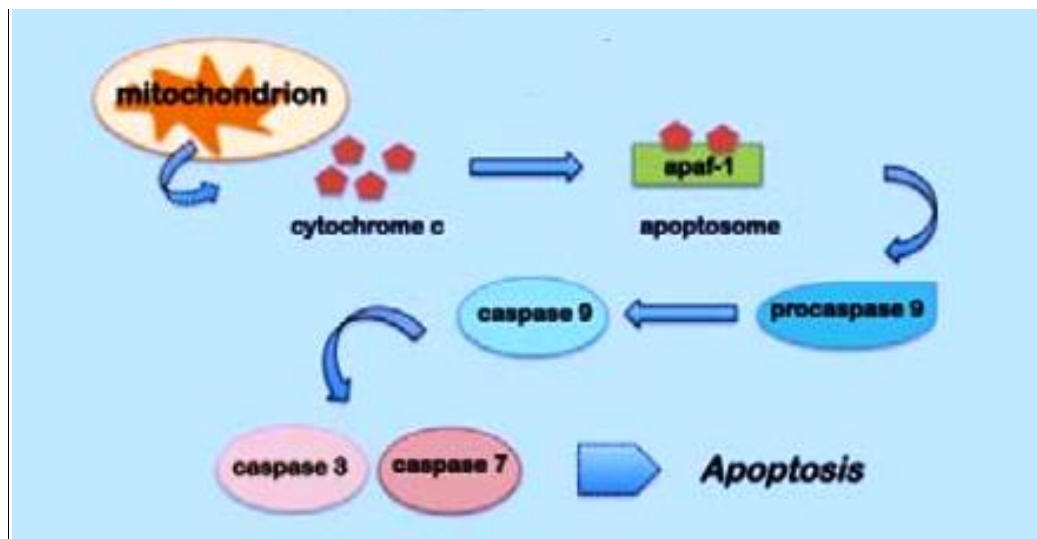
Figure 1.8: Schematic overview of the genotoxic stress-induced apoptotic pathway, showing the key factors involved. (Adapted from (Montecucco et al., 2015)

1.5.2 PARP1 inhibitors

It has already been discussed how important the maintenance of genomic integrity is and what a big part PARP1 plays in it. It is unsurprising then that for over 30 years, PARP1 inhibitors have been the most promising drugs that target DNA repair (Purnell and Whish, 1980, Rouleau et al., 2010). As mentioned Poly(ADP-ribose) polymerase 1 (PARP1) is a protein that has been implicated in the repair of SSBs, but it has also been shown to be recruited to sites of DSBs (Rouleau et al., 2010). Furthermore, if such SSBs persist unrepaired until the DNA is replicated, then DSBs can form (Rouleau et al., 2010). Drugs that work as PARP1 inhibitors can induce the formation of such DSBs by interfering with the repair of SSBs, but that is not the only proposed explanation (Haffner et al., 2011, Bareschino et al., 2011).

In 2005, two groups independently showed the ability of PARP inhibitors to selectively inhibit the growth of cells with defects in either the BRCA1 or the BRCA2 genes, suggesting a synthetic lethality use (Bryant et al., 2005, Farmer et al., 2005). There have been a number of studies investigating the use of PARP1 inhibitors as single agents, especially in BRCA-mutated breast, ovarian and prostate tumours, with the first two showing the most favourable response (Fong et al., 2009, Audeh et al., 2010). In December 2014, the FDA and the EMA approved Olaparib, a PARP1 inhibitor as monotherapy for the treatment of recurrent ovarian cancer in women with a *BRCA1* or *BRCA2* mutation (Frampton, 2015). "BRCAness" is a profile of tumours that have common traits with BRCA-mutated tumours and is also of interest for PARP1 inhibition (Weil and Chen, 2011).

There have also been studies looking into combination therapies, where a PARP1 inhibitor along with one or more other factors of the DNA damage response system, are inhibited (Hosoya and Miyagawa, 2014). There is currently a trial (NCT01642251) that is testing a combination of the PARP1 inhibitor veliparib, cisplatin and etoposide for extensive stage small-cell lung cancer, metastatic large cell neuroendocrine NSCLC and small-cell carcinoma of

unknown primary or extrapulmonary origin. This trial is at Phase II and due to conclude in January 2018, demonstrating that combination therapy targeting both PARP1 and TOP2 simultaneously has merits after progression from its first phase (Hosoya and Miyagawa, 2014).

The exact mechanism of action of PARP1 inhibitors with regard to their effects in cancer cells is not fully understood, but there is evidence that PARP1 inhibitors trap PARP1 enzymes at the locations of damaged DNA; those PARP1-DNA complexes were more cytotoxic than unrepaired SSBs caused by PARP1 inactivation, and if we were to draw a parallel with TOP2 poisons, we could say that such PARP1 inhibitors act in a PARP1-poison manner (Murai et al., 2012)

1.6 TDP2 as a biomarker to sensitivity and as a novel therapeutic target

As mentioned previously, both types of topoisomerase 2 are the targets of several anticancer drugs for the treatment of both early and advanced cancers. Cellular depletion of TDP2 has been shown to result in increased sensitivity to TOP2-induced DNA breaks and TOP2 poisons (Cortes Ledesma et al., 2009, Zeng et al., 2011). On the other hand, over-expression of TDP2 increases resistance to TOP2 poisons (Do et al., 2012).

There are recent data which suggest that expression levels of TDP2 vary greatly in different types of cancer cell lines, such as breast and lung, ranging from almost no expression to very high levels of expression (Li et al., 2011a). However, there are no reported studies addressing the possible roles of TDP2 as determinants of outcome or clinical predictors, or studies correlating over-expression with resistance to TOP2 poisons. In addition to variations in TDP2 expression across different tumour types, chemoresistance frequently occurs (Li et al., 2011b). One possible contributing factor to this is increased repair of the DNA, and TDP2 has been reported to contribute to such resistance directly (Li et al., 2011b).

Given TDP2's role in repairing TOP2-mediated DSBs, TDP2 protein levels could be influential for cancer cell resistance to TOP2 poisons, which are widely used in cancer treatment. There are

recent studies that reinforce this hypothesis. TDP2 depletion by RNAi in the lung cancer cell line A549 resulted in increased sensitivity to etoposide (Cortes Ledesma et al., 2009). Similar observations were made in chicken DT40 cells, in which deletion of TDP2 led to increased sensitivity to etoposide (Zeng et al., 2011). It has also been observed in both H460 and A549 lung cancer cell lines that knocking down TDP2 sensitises these cells to etoposide-induced cell growth inhibition (Bian et al., 2016). This suggests that types of cancer with high levels of TDP2 might be relatively resistant to TOP2 poisons, while types of cancers with low levels of TDP2 might be sensitive to such treatment. Consistent with this idea, increased levels of TDP2 have been observed in several non-small cell lung cancer cell lines, and a possible oncogenic role of TDP2 in such cancers has been suggested (Li et al., 2011a).

Two groups of cancer cell lines that could serve as a focus group for studies correlating over-expression with resistance to TOP2 poisons, or to further explore the potential of TDP2 as determinants of outcome or clinical predictors would be a panel of lung cancer and breast cancer cell lines. Non-small cell lung cancers in particular, represent 80-85% of all lung cancer cases, something that makes them clinically important (Breathnach et al., 2001, Bareschino et al., 2011). Breast cancer is the most common type of cancer in women worldwide (Siegel et al., 2015). Around 20-30% of breast cancer cases will become metastatic (O'Shaughnessy, 2005). In addition, about 6% of non pre-existing breast cancer cases begin as metastatic, also known as Stage 4 (Program, 2015). Etoposide is a commonly used drug that has been increasingly used over the past few years for patients that suffer from metastatic breast cancer and have received previous treatment (Yuan et al., 2015, Valabrega et al., 2015).

TDP2 has the potential to be a good target for pharmacological inhibition to potentially increase tumour sensitivity to TOP2 poisons, particularly those that develop resistance during the course of treatment. Recently, there has been an increase of information about studies looking into the possibility of a TDP2 inhibitor. One group reported that toxoflavin derivatives

and deazaflavins were identified as sub-micromolar and selective inhibitors of TDP2 via a high-throughput screening, however, even though they both demonstrated a structure-activity relationship, unfortunately both of them posed challenges (Raoof et al., 2013). Toxoflavins demonstrated a key redox liability, early in vitro metabolism and pharmacokinetic issues and deazaflavins low cell permeability (Raoof et al., 2013). A different group reported isoquinoline-1,3-dione compounds as selective TDP2 inhibitors, which were identified by a screening of in-house collection of synthetic compounds, followed by a structure-activity relationship study (Kankanala et al., 2016). Another publication also presented data about deazaflavin being a good candidate for TDP2 inhibition, but also went one step further and included data that showed potent synergy in combination with etoposide in human prostate cancer cells DU145 and TDP2-dependent synergy in TK6 human lymphoblast and avian DT40 cells, also showing that TDP2 knock out cells did not show significant synergy (Marchand et al., 2016). Another paper identified a selective inhibitor via a high-throughput screening and subsequent validation studies (Kont et al., 2016). NSC111041 was shown to inhibit TDP2's binding to DNA without intercalating into DNA and also demonstrated a synergistic effect with etoposide in TDP2 expressing cells but not cells with depleted TDP2. In addition, the inhibitor increased the number of etoposide-dependent γ-H2AX foci, a biomarker for DSBs (Kont et al., 2016). A different group reported another collective of TDP2 inhibitors that emerged from experimentally verifying the selective inhibitory activity of relevant molecules from the NCI diversity small molecule database (Kossmann et al., 2016). This group also reported on the binding mode of these inhibitors by molecular dynamics simulations, which showed association to the TDP2 DNA-binding cleft (Kossmann et al., 2016). Finally, another group confirmed that deazaflavin inhibits TDP2; a fluorescence-based assay was used, which is suitable for high-throughput screening, but it was taken one step further and the crystal structures of these compounds bound to a 'humanized' form of murine TDP2 were determined. Interestingly, these structures demonstrated that the compounds act as

competitive ligands for the binding site of an incoming DNA substrate, something that could potentially point the way towards generating new and effective TDP2 inhibitors (Hornyak et al., 2016).

There is a recent model, which suggests that induction of transcriptional programs, for example by stimulating breast or prostate cells with estrogens or androgens respectively, can involve the formation of TOP2B-mediated DNA DSBs and the recruitment of DSBR proteins (Haffner et al., 2011). It has been previously shown that for successful signal-dependent activation of transcription by transcription factors such as the estrogen receptor (ER), formation of transient TOP2B-mediated DSBs is required at regulatory elements (Ju et al., 2006). Ju et. al. showed that rapid increase of TOP2B and PARP1 levels were detected after treatment with estradiol (also known as E2) and were also co-purified with Ku80, Ku70 and DNA-dependent protein kinase (DNA-PK), all of which are components of the DNA damage and repair machinery (Ju et al., 2006). Furthermore, following estradiol treatment, TOP2B and the components of the repair machinery were all stably recruited to the pS2 gene promoter but not the coding region of the gene (Ju et al., 2006). The pS2 gene encodes the pS2 estrogen response element (specific to the breast cancer cell line MCF7) that encompasses structural features similar to some growth factors and whose transcription is stimulated by estrogen (Stack et al., 1988, Kim et al., 2000). Similar findings have been demonstrated for the androgen receptor (Haffner et al., 2010). TOP2B is believed to be recruited with the estrogen or androgen receptor to regulatory sites on target genes for efficient activation of those genes (Fig.1.9) (Haffner et al., 2011).
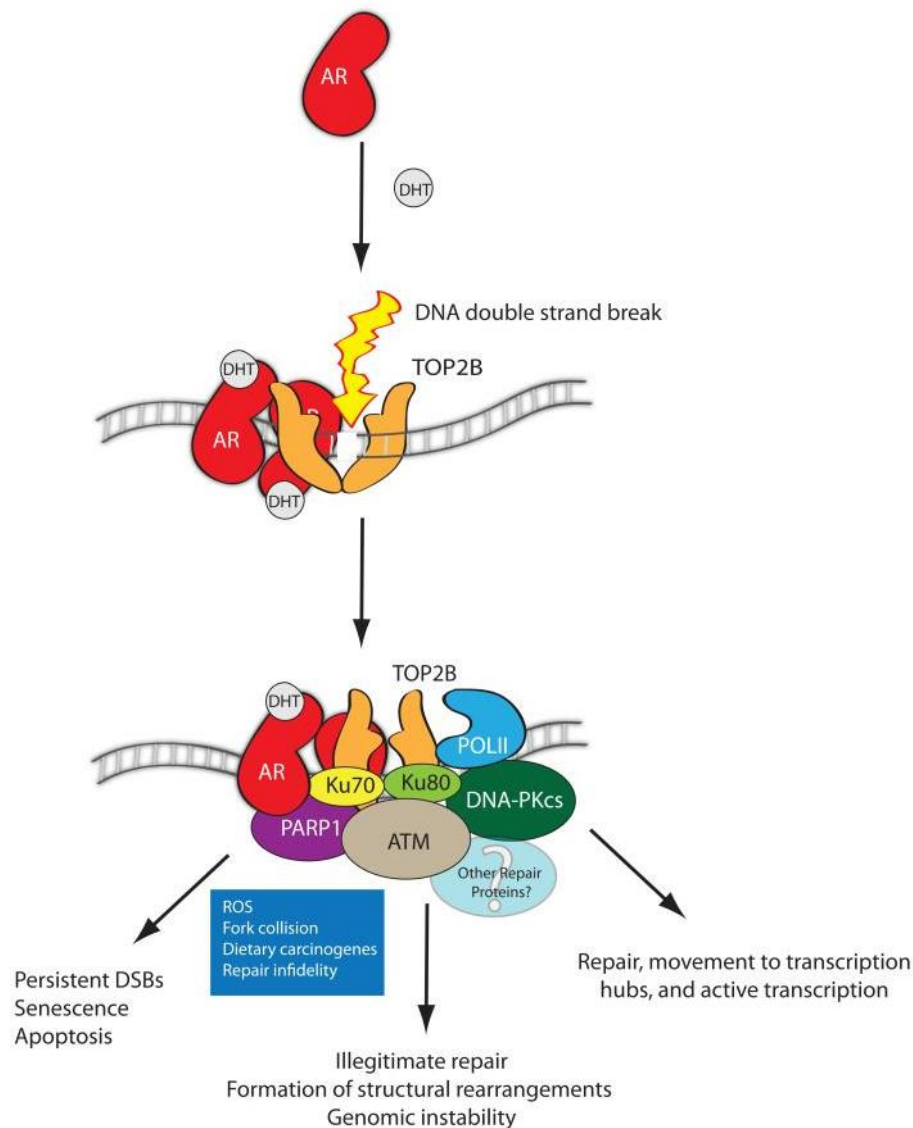
Figure 1.9: Activation of transcription by the androgen receptor AR is followed by association with TOP2B. Together they bind AR target sites. Recruitment of the DNA repair proteins Ku70, Ku80, PARP1, ATM and DNA-PK follows. TOP2B introduces a DSB, which can be rendered abortive by reactive oxygen species (ROS), dietary carcinogens, fork collisions or repair infidelity. Non-repaired abortive DSBs can lead to serious complications for the cell, including apoptosis and genomic instability (Haffner et al., 2011).

It has been shown that tumours experiencing acute or chronic hypoxia (such as solid or fast growing tumours) are sensitive to PARP1 inhibitors (Chan et al., 2010). Furthermore, there is data which suggests that certain types of breast cancer might be sensitive to PARP1 inhibitors (Weil and Chen, 2011). Whilst the generation of a transient SSB is central to the activity of TOP1, the two subunits of TOP2 introduce DNA breaks in an independent but coordinated

process and therefore can be trapped as a SSB if TOP2 activity on both strands is not coordinated, which normally would be a target for PARP1 (Zechiedrich et al., 1989, Nitiss, 2009b).

Estradiol is one of the most common estrogen hormones and is sometimes used as part of hormone cancer therapy in both women and men, especially breast cancer (Ingle, 2002). In addition, it has been involved in higher breast and ovarian cancer risk in women that have been prescribed estradiol for menopause (Beral, 2003, Beral et al., 2007).

Breast cancers are usually hormone dependent and they typically respond well to hormone therapies. Even so, a large percentage of those tumours will develop resistance to this type of treatment, while still expressing ER, implying that they still depend on the ER pathway (Risbridger et al., 2010). It is hypothesised that this could be exploited therapeutically. A combination treatment could be established that will first induce transcription (for example with estradiol), which will cause transient TOP2B-mediated DNA DSBs, and then transform those breaks into abortive breaks using a TOP2 poison. Failure to repair these DSBs due to the TOP2B-targeting drug, could lead to the cells being overwhelmed with DSBs, thus promoting apoptosis. Such a strategy could possibly find particular use in slow proliferating hormone-dependent cancers such as breast tumours, where resistance has developed or other types of treatment have failed. More specifically, this could potentially be more effective in breast cancer cell lines with low levels of TDP2 and higher levels of TOP2B, as this would mean there would not be enough TDP2 to remove the stalled TOP2B, increasing the chance of success of such a treatment. Furthermore, to increase the success of a cycle treatment as the one proposed, the addition of a PARP1 inhibitor could also be utilised to further expand on the observation that it is recruited with TOP2B after treatment with estradiol.

1.7 Lung and breast cancer cell lines as tools for investigating novel therapeutic targets

Non-small cell lung cancer cell lines are widely used in scientific research. Most of them are easy to obtain and maintain. A549 is an epithelial carcinoma cell line (American Type Culture Collection - ATCC). It grows in an adherent manner and it was established in 1972 through explant culture of lung carcinomatous tissue from a 58-year-old Caucasian male (Giard et al., 1973). NCI-H292 (H292) is an epithelia mucoepidermoid pulmonary carcinoma cell line (ATCC). It grows in an adherent manner and has been previously been identified as Hut 292 (Lechner et al., 1983). It was deposited by A. Gazdar (National Cancer Institute-Navy Medical Oncology Branch, Bethesda, MD) and it was derived from a 32-year-old Black female (Carney et al., 1985). NCI-H1650 (H1650) is an epithelial stage 3B adenocarcinoma and specifically bronchoalveolar carcinoma derived from a metastatic site (pleural effusion) that grows in an adherent manner (ATCC). It was deposited by A. Gazdar and J. Minna in May of 1987 and was established from a 27-year-old Caucasian male (Gazdar, 1996). NCI-H460 (H460) is an epithelial carcinoma derived from a male patient that grows in an adherent manner (ATCC). It was deposited by A. Gazdar and J. Minna in 1982 (Banks-Schlegel et al., 1985). NCI-H226 (H226) is an epithelial squamous cell carcinoma, specifically mesothelioma, and was isolated from a metastatic site (pleural effusion) (ATCC). It grows in an adherent manner (ATCC). It was deposited by A. Gazdar and J. Minna in March of 1980 and was established from a male patient (Gazdar, 1996). MOR is an epithelial-like adenocarcinoma cell line that grows in an adherent manner in colonies (Public Health England Culture Collections). It was deposited by Dr. P. Twentyman and Dr. K. Wright (UKCCR, PO Box 123, Lincolns Inn Fields, London WC2A 3PX) (Twentyman et al., 1986). Hop92 is an epithelial carcinoma cell line, which was contributed by M. Liu. (US National Cancer Institute). It was derived from a 62-year old male patient (US National Cancer Institute). HCC95 (prefix HCC for Hamon Cancer Center) is an

epithelial squamous cell line derived from a metastatic site (pleural effusion) of a 65-year old

male patient in 1992 (Wistuba et al., 1999).

Similar to non-small cell lung cancer, breast cancer lines are also very common in scientific

research. As previously mentioned, breast cancer is one of the most common types of cancer.

T47D is an epithelial ductal carcinoma derived from a metastatic site (pleural effusion) of

mammary gland tissue (ATCC). It was isolated from a 54-year old female and deposited by I.

Keydar (Keydar et al., 1979). MDA-MB-361 (MM361) is an epithelial adenocarcinoma cell line

derived from a metastatic site (brain) originating in mammary gland and breast tissue, from a

40-year old Caucasian female (ATCC). It was deposited by R. Cailleau (Cailleau et al., 1978).

MDA-MB-453 (MM453) is an epithelial metastatic cancinoma cell line, derived from a

metastatic site (pericardial effusion) originally from mammary gland and breast tissue (ATCC).

It was isolated from a 48-year-old Caucasian female and was deposited by R. Cailleau (Cailleau

et al., 1978). BT-474 (BT474) is an epithelial ductal carcinoma cell line from mammary gland

tissue (ATCC). It was isolated from a 60-year-old Caucasian female and was deposited by E.

Lasfargues (Lasfargues et al., 1978). ZR-75-1 (ZR751) is an epithelial ductal carcinoma derived

from a metastatic site (ascites) originally from mammary gland tissue (ATCC). It was isolated

from a 63-year-old Caucasian female and was deposited by L. Engel (Engel et al., 1978). MCF7

(Michigan Cancer Foundation-7) is an epithelial adenocarcinoma cell line derived from a

metastatic site (pleural effusion) originating from mammary gland tissue (ATCC). It was

isolated by a 69-year old Caucasian female by H. Soule and co-workers in 1973 and was

deposited by C. McGrath (Soule et al., 1973). It is also ERα positive and grows in an estrogen-

dependent manner. That means that types of cancer with the MCF7 profile usually respond

well to inhibitors of ERα activity or drugs that block estrogen synthesis. MLET5 is an MCF7

derivative, which whilst ERα positive, it grows in an estrogen-independent manner. It was

established by over-expressing ERα in MCF7 cells using adenovirus gene transduction, which

resulted in ligand-independent expression of pS2 and PR (estrogen-regulated genes), and the

cell line to be able to grow in an estrogen-independent manner (Tolhurst et al., 2011). Similarly, LCC9 is an MCF7 derivative that grows in an estrogen-independent manner, but it was acquired in a different way. Brünner et al., had previously established an MCF7 derivative, LCC1 that is able to grow *in vivo* and *in vitro* independent of estrogen, but still remain sensitive to antiestrogens, by cultural selection for resistance (Brunner et al., 1993a). The same group was also able to grow an MCF7/LCC1 derivative, LCC2, which in addition to estrogen-independent growth, it is also resistant to 4OH-TAM, again via cultural selection for resistance, but was not cross-resistant to the steroidal entiestrogens ICI 182,780 and ICI 164,384 (Brunner et al., 1993b). ICI 182,780 and ICI 164,384 are C7-acyl-substituted analogues of estradiol and have a higher affinity for ER than TAM (Wakeling et al., 1991, Thompson et al., 1989). Therefore, a few years later, Brunner et al. published a paper in which they reveal another MCF7/LCC1 derivative, LCC9, which was selected against ICI 182,780, but through that it also acquired cross-selection for TAM (Brünner et al., 1997).

T47D, Hop92, A549, H460, MCF7, H226 are part of the NCI-60 cell line panel. In the late 1980s, the US National Cancer Institute (NCI) created and developed an anticancer drug discovery and screen project containing 60 human tumour cell lines (Shoemaker, 2006). The intention of the project was to screen *in vitro* thousands of pharmacological compounds for anticancer activity, replacing earlier *in vivo* models. These cell lines included nine types of cancer: lung, breast, prostate, ovary, colon, central nervous system, kidney, leukemia and melanoma, which formed a panel referred to as NCI-60 (Abaan et al., 2013).

1.8 Aims of the thesis

This thesis aims to initially establish the mRNA and protein levels of TDP2, TOP2A and TOP2B in a panel of non-small cell lung cancer cell lines and a panel of breast cancer cell lines and explore whether these levels exhibit a large variation. Following that, the project will investigate the possibility of a correlation between the levels of TDP2, TOP2A and TOP2B and

resistance to the TOP2-targeting agent etoposide. It is hypothesised that cell lines exhibiting low levels of TDP2 and high levels of TOP2A and TOP2B will be more sensitive to etoposide, compared to cell lines that have showcased high levels of TDP2 and low levels of TOP2A and TOP2B. This is based on previous observations where depriving cells of TDP2 has lead to increased etoposide sensitivity compared to wild type cells (Cortes Ledesma et al., 2009, Zeng et al., 2011). This will serve to establish TDP2 as a biomarker for sensitivity to etoposide and a clinical predictor of treatment outcome.

In addition, the emerging model of TOP2B-mediated DSBs during transcription will be utilised to explore the possibility of TDP2 as a novel therapeutic target for inhibition. As previously mentioned, this is based on the hypothesis that inducing transcription in selected breast cancer cell lines by the estrogen estradiol, thus inducing TOP2B-mediated DSBs, and treating them with etoposide to shift the transient DSBs to abortive would overwhelm the cell's TDP2 and render it unable to cope with the high number of abortive DSBs, leading to apoptosis. Furthermore, it would be interesting to investigate whether the introduction of TOP2B-mediated DSBs by treatment with estradiol, could be severe enough in itself to affect cell survival in cancer cells with low levels of TDP2 or TDP2-depleted cell lines. If the results are promising, then perhaps the next step would be a compound that is able to inhibit TDP2 directly without the need of a knock-out.

2. CHAPTER TWO - Materials and Methods

2.1 General Chemicals and Equipment

All chemicals were obtained from Fisher Scientific or Sigma unless otherwise stated. All enzymes were obtained from Roche or New England Biolabs unless otherwise stated. All cell culture media was obtained from Gibco® by Life Technologies. All cell culture flasks, dishes and cryovials from Corning® Sigma.

All centrifugation steps, with the exception of cell culture, were carried out using a Thermo Scientific Heraeus Pico 21 centrifuge, a Thermo Scientific Heraeus Fresco 21 centrifuge or a Beckman Coulter Allegra X-22R centrifuge using S4180, F0630 and F24020H rotors.

All cell culture centrifugation steps were carried out using a Denley B5400 centrifuge or a Hettich® Rotina 38R refrigerated centrifuge.

The thermal cyclers that were used were Techne TC-5000, MWG-Biotech Primus 25 and PeqLab Advanced Primus 25.

Microscopy was carried out with a Nikon E400 or a Zeiss Axioplan 2.

The colony counter was a Stuart SC6.

2.2 Mammalian Cell Culture

The following cell lines were used:

•	Lung cancer cell lines: A549, H292, H1650, H460, H226, HCC95, Hop92, and MOR

•	Breast cancer cell lines: MDA-MB-453 (MM453), MDA-MB-361 (MM361), T47D, MCF7, ZR-75-1 (ZR751), BT474, LCC9 and MLET5.

The cell lines were maintained in the appropriate media. RPMI was used for H292, H1650, H460, H226, Hop92, MOR, T47D, ZR751 and BT474, Leibovitz's was used for MM361, DMEM

was used for A549, HCC95, MM453 and MCF7, and phenol-red-free DMEM + 1% pyruvate was used for MLET5 and LCC9. RPMI and DMEM media was supplemented with 2mM L-glutamine, 50μg/ml streptomycin, 0.5 units/ml penicillin (all from Gibco® by Life Technologies) and 10% heat-inactivated foetal bovine serum (FBS – from Pan-Biotech). Leibovitz's media was supplemented with 2mM L-glutamine, 50μg/ml streptomycin, 0.5 units/ml penicillin and 20% heat-inactivated FBS. Phenol-red-free DMEM media was supplemented with 2mM L-glutamine, 50μg/ml streptomycin, 0.5 units/ml penicillin, 1% pyruvate and 10% charcoal-stripped FBS (from Sigma). They were all grown in a $37^{o}C$ humidified incubator with 5% CO_2.

Cells were regularly tested for mycoplasma contamination. A Lonza MycoAlert Mycoplasma Detection Kit was used, according to manufacturer's instructions.

2.3 Analysis of cellular extracts by gene expression and SDS-PAGE immunoblotting

2.3.1 Whole cell extracts

Cells were washed with phosphate buffered saline (PBS) and 0.25% Trypsin powder in PBS was added for cell dissociation. Cell pellets were obtained at different time points of cell culture and were immediately frozen in LN_2 and kept at $-20^{o}C$ after that.

2.3.2 RNA Extraction

RNA was extracted for the above cell lines using an optimised protocol. The cells were homogenised with 1ml TRI reagent and incubated at room temperature for 5 minutes (min). Then, the samples were centrifuged at $4^{o}C$ for 5 min at 12000g. The supernatant was transferred to a new eppendorf tube and 100μl of 1-bromo-3-chloro-propane (BCP) was added. The samples were vortexed for 15 seconds (sec) to get a homogenous mixture. They were then incubated for 15 min at room temperature. Subsequently, they were centrifuged at $4^{o}C$ for 10 min at 12000g. The aqueous phase was rescued and precipitated with 500ml 100%

isopropanol. The samples were then vortexed for 10 sec and incubated for 10 min at room temperature, followed by a 5 min spin at 12000g at 4°C. The supernatant was discarded and 1ml of 75% ethanol was added. The samples were once again centrifuged at 4°C for 5 min at 7500g. The ethanol was removed and the pellet was left to air dry for 5 to 10 min. Finally, the RNA was re-suspended in 100µl RNAse-free water. It was quantified with a spectrophotometer (Nanodrop, Thermo Scientific).

2.3.3 Gene expression analysis

High-capacity cDNA reverse transcription kit without RNAse inhibitor by Applied Biosystems was used and the provided protocol was followed. 100ng of RNA were retrotransformed into cDNA for each cell line. Cycling conditions were as follows: 25°C for 10 min, 37°C for 120 min and 85°C for 5 min on one of the thermal cyclers mentioned in section 2.1. qPCR was run in triplicate using 1ng of cDNA for each reaction and TaqMan® gene expression assay by Life Technologies was used for each gene (Table 2.1). The volumes used for each single reaction were 5µl of 2x master mix, 0.5µl primer & probe, 3.5µl RNAse free water and 1µl sample. A two-step cycling program was used: 95°C for 10 min and 40 cycles of 95°C for 15 sec followed by 60°C for 1 min. The qPCR machine that was used was an MX3005P Stratagen, with the accompanying MxPro QPCR Software.

Table 2.1: TaqMan® gene expression assay genes		
Gene	Catalogue Number	Company
TDP2	Hs01099017_m1	Life Technologies
TOP2A	Hs00172214_m1	Life Technologies
TOP2B	Hs00172259_m1	Life Technologies
RPLP0	Hs99999902_m1	Life Technologies

2.3.4 SDS-Polyacrylamide gel electrophoresis

10% sodium dodecyl sulphate (SDS) polyacrylamide gel electrophoresis (PAGE) separating gels were used. These were prepared by mixing 4ml of distilled water, 3.3ml of 30% acrylamide, 2.5ml 1.5M Tris-HCl pH 8.8, 100μl 10% SDS, 100μl 10% APS and 4μl TEMED (Sambrook and Russell, 2006). The stacking gels were prepared by mixing 3.4ml distilled water (dH$_2$0), 830μl 30% acrylamide, 630μl 1M Tris-HCl pH 6.8, 50μl 10% SDS, 50μl 10% APS and 5μl TEMED.

Samples were prepared with the addition of 1x SDS-PAGE loading buffer (50mM Tris pH 8.0, 2% w/v SDS, 10% w/v glycerol, 0.1% w/v bromophenol blue 200mM DTT). In order to avoid degradation, SDS-PAGE loading buffer heated to 100°C was used on samples that were removed from -20°C and kept on ice shortly before the preparation for the western blot. Then, the samples were incubated at 100°C for 5 min, vortexed for 20 sec, incubated at 100°C for 5 min again, vortexed again and centrifuged at 30000rpm for 1 min. The samples were separated on SDS-PAGE using Precision Plus Protein standards, dual colour (BioRad) as markers. Electrophoresis was carried out using a power pack and a Mini Protean III Cell obtained from BioRad, in 1x running buffer (25mM Tris, 250mM glycine, 0.01% w/v SDS) at 200V for 70 min.

The first time SDS-PAGE was done for a group of replicates, 5μl, 10μl and 20μl of each cell line extract were run on the same gel and total protein was visualised by staining in Coomassie brilliant blue solution (0.25% w/v coomassie brilliant blue G-250 BioRad, 50% v/v methanol, 10% v/v acetic acid) for 30 min, destaining (10% v/v methanol, 10% v/v acetic acid) as required and dried. Based on the subsequent strength of the bands, approximate volumes that would give the same band strength among all cell lines were picked. These volumes were then doubled and doubled again, in order to obtain concentrations with the following correlation: 4C1=2C2=C3.

2.3.5 Immunoblotting of proteins

Total protein was transferred from polyacrylamide gels to nitrocellulose membrane (Hybond-C Extra Membrane from GE Healthcare) by wet transfer using Mini Protean III Trans-Blot Cell (BioRad), in 1x Towbin buffer (20mM Tris, 150mM glycine, 10% v/v methanol) at 80V for 1.5h or 10V for 999 min. Successful transfer was determined by non-permanent staining of total protein on the membrane with 0.1% w/v Ponceau S in 5% acetic acid. The stain was removed by washing in 1x TBST (10mM Tris, 140mM sodium chloride, 0.1% v/v Tween 20, pH 7.9). Membranes were blocked by incubation with 5% (w/v) milk powder in 1x TBST for 1h at room temperature, shaking. The rocker that was used was a Stuart See-saw SSL4. The membranes were then incubated in primary antibody under the appropriate conditions (Table 2.2). Actin was used as a loading control. TDP2, TOP2A and TOP2B were immunostained on different blots, to reduce the likelihood of shared technical artefacts between the different immunoblots. Primary antibodies were diluted in 5% (w/v) milk powder in 1x TBST. Following that, they were washed in 1x TBST (3x 10 min) at room temperature, shaking. Finally, they were incubated in the appropriate secondary antibody at room temperature, shaking (Table 2.3). Secondary antibodies were diluted in 1x TBST. The membranes were washed again as described for the primary antibody and incubated in enhanced chemiluminescence (ECL) western blotting detection reagent (GE Healthcare) for 1 min. To visualise the desired proteins, ImageQuant LAS 4000 system was used. To quantify the data obtained, its accompanying software ImageQuant TL was used. The software allows for background elimination for each individual band.

Table 2.2: Primary antibodies for western blotting

Primary Antibodies	Type	Organism	Company	Catalog No.
Actin	Monoclonal	Mouse	Sigma	A4700
TDP2	Polyclonal	Rabbit	Biogenes	Custom (Thomson et al., 2013)
TOP2A	Polyclonal	Goat	Santa Cruz	S-20
TOP2B	Polyclonal	Rabbit	Santa Cruz	H-286

Table 2.3: Secondary antibodies for western blotting

Secondary Antibodies	Company	Catalog No.
Rabbit anti-mouse HRP	Dako	P0260
Goat anti-rabbit HRP	Dako	P0448
Rabbit anti-goat HRP	Dako	P0449

2.3.6 Stripping for reprobing western blots

Stripping is the procedure that allows the removal of primary and secondary antibodies from a western blot membrane, in order to be able to reprobe the same western blot with a different antibody. This is particularly useful when one wants to use a loading control that has a similar size to the protein of interest, such as TDP2 and Actin. This procedure was also followed for TOP2A and TOP2B. A protocol for mild stripping by abcam was used. The blots were incubated for 10 min at room temperature, shaking, in stripping buffer (for 100ml total volume, 1.5g glycine, 1ml of 10% SDS and 1ml Tween 20 was used), which was then discarded. The blot was incubated for another 10 min at the same conditions. The buffer was discarded again. That was followed by two 10 min incubations with PBS at room temperature, shaking and

subsequently 2x 5 min incubation with TBST at room temperature, shaking. At this stage, the blots were ready for the immunoblotting protocol to resume from the blocking stage onwards.

2.4 Analysis of relationship of TDP2, TOP2A and TOP2B protein levels in lung and breast cancer cell lines and sensitivity to TOP2 poisons

2.4.1 Alamar Blue Assay

In order to explore the effect of different TDP2 levels to etoposide sensitivity in breast cancer cell lines, Alamar Blue assays (Invitrogen) were performed. This assay is used to determine whether cells are still metabolically active. The protocol was provided by Invitrogen, which is as follows: On the first day, the appropriate amount of cells was seeded (which depended on cell line growth curves) for each cell line in triplicates in a 96-well plate (flat bottom Corning® Costar® cell culture plates) with a final volume of 150µl per well, and on the next day they were treated with different concentrations of etoposide (0µM, 5µM, 10µM, 20µM, 30µM). These concentrations were chosen, because it was determined that after 30µM cell survival was very low and therefore not relevant to the objective of the assay, which was to observe any differences in cell survival among the selected cell lines. After 3h of treatment, the cells were washed twice with 100µl of PBS, and 150µl of appropriate media was added. On the third day, the media was changed, but this time the volume was 100µl, and 30µl of Alamar Blue reagent was added. The cells were then incubated at $37^{o}C$ for 4 hours protected from direct light, which was followed by measurement of fluorescence (excitation wavelength 540-570nm and fluorescence emission 580-610nm) with a BioTek Synergy Multi-Mode Reader. The protocol for the third day was done every other day, until there were five fluorescence measurements. For each triplicate, the average was calculated. Then for each biological replicate, the average was calculated. From these averaged fluorescence values, the standard error of the mean was calculated and the results were plotted with Microsoft Excel.

The same assay was utilised for three lung cancer cell lines, MOR, HCC95 and H292 and was carried out in pairs: MOR-HCC95 and MOR-H292. For these cell lines the same protocol was used, although there was an additional concentration of etoposide that was used, which was 2µM. Finally, for two lung cancer cell lines, LCC9 and ZR751, a continuous treatment was utilised, during which the etoposide concentrations were 0nM, 250nM, 500nM, 750nM and 1µM. Also, the etoposide was not washed until it was time for the Alamar Blue reagent to be added.

2.4.2 Clonogenic Cell Survival Assay

In order to explore the effect of different TDP2 levels to etoposide sensitivity in lung cancer cell lines, clonogenic cell survival assays were utilised. This kind of assay is used to evaluate the ability of a cell to proliferate after treatment, in this instance with etoposide, by forming a viable colony. The subsequent cell survival curve is therefore defined as the relationship between the doses of etoposide and the percentage of cells retaining their ability to proliferate after treatment compared to untreated cells. The following optimised protocol was followed: On the first day, the appropriate amount of cells was seeded (which was determined by trying different numbers of cells, until the appropriate number that produced enough viable colonies after two weeks was calculated) and on the next day, they were treated with different concentrations of etoposide (0µM, 5µM, 15µM, 30µM). After 3h, the cells were washed twice with PBS, appropriate media was added and the cells were left to grow for 14 days. On the 14th day, the colonies were fixed with 100% ethanol, washed with PBS and stained with 1% methylene blue in 70% ethanol.

2.5 Selection process and statistical analysis for mutational profiles of breast and lung cancer cell lines

2.5.1 Selection of Databases

In order to explore the effect of various mutations on the relationship between etoposide sensitivity and TDP2 protein levels, TOP2A/TDP2 protein level ratios or TOP2B/TDP2 protein level ratios in the previously used panels of lung and breast cancer cell lines, databases that provided information on mutations found in those cancer cell lines were required. The selection of the appropriate databases was done through review of the relevant literature and consultation with people familiar with the field (Futreal et al., 2004, Pavlopoulou et al., 2015, Abaan et al., 2013, Yang et al., 2015b). This resulted in the selection of Cancer Cell Line Encyclopedia v2.17 (CCLE), IARC TP53 Database R18, Cell Lines Project v78 and COSMIC Database v78 (Barretina et al., 2012, Bouaoun et al., 2016, Forbes et al., 2015).

These databases are expert-curated. For the COSMIC Database, the primary source is papers in scientific journals. It was verified that the samples listed were of the same type as the cell lines previously used, as multiple instances of the same sample name exist as separate entries, indicating that it was not clear for the curator whether these samples were the same. Samples which were not or there was personal doubt that they were indeed the same, were not used. It was also verified that the mutated or wild type form of the genes stated in the database matched those stated in the publications. All databases were last accessed in September 2016, to include as many up-to-date entries as possible.

2.5.2 Selection of Mutations

After being collected, the appropriate data from the databases were organised per cancer type (breast and lung) in two excel spreadsheets. Each spreadsheet listed all the genes for which information for at least one cell line of that type was found, either a mutation or a wild type entry. For the breast cancer cell lines that was a total of 3686 genes and for the lung cancer

cell lines 1150 genes. Out of the 8 breast cancer cell lines studied (LCC9, MCF7, MM361, BT474, MLET5, T47D, MM453 and ZR751), information was available for at least one gene for 6 cell lines; BT474, MCF7, MM361, MM453, T47D and ZR751. Out of the 8 lung cancer cell lines studied (MOR, HCC95, Hop92, H226, H460, H292, A549 and H1650), information for at least one gene was available for 7 cell lines; A549, H226, H292, H460, H1650, HCC95 and Hop92. Genes for which there was conflicting publications for cell line mutational status, were only included if there were at least three reports that agreed and a minimum ratio of 3 to 1.

Information for the gene status in a minimum of four cell lines, with at least one of the cell lines' mutation status being different than the rest, was the minimum requirement for the statistical analysis performed. This criterion reduced the number of genes to 7 for the breast cancer cell lines and 8 for the lung cancer cell lines. All mutant forms of each gene were grouped together so as to maximise the number of genes, as not all mutations were characterised at that level or their effects defined.

2.5.3 Statistical Analysis

Analysis of this data was performed using R statistical software (Team, 2015). The effect of each gene (as per the selection process described) on the relationship between TDP2 levels or TOP2/TDP2 ratios was independently assessed by estimating parameters of a linear regression model, where etoposide IC50 (for breast cancer cell lines) or D37 (for lung cancer cell lines) concentrations and gene status of a cell line predict the TDP2 levels or TOP2/TDP2 ratios. The significance for each parameter was assessed using an F test from an analysis of variance (ANOVA) between models with and without terms of interest.

The interaction between IC50/D37 values and gene status, i.e. giving a different slope to the mt and wt lines of best fit, was not considered, because it would be a substantially more complex model requiring more data to estimate parameters. Both the effect of IC50/D37 values and gene status were considered on their own as predictors of TDP2 levels or

TOP2/TDP2 ratios or in conjunction with the other term (e.g. the effect of IC50 alone on TDP2 levels, or the effect of IC50 on TDP2 levels once the effect of gene mutation status has been taken into account). This leads to two estimates and two p-values for each term as shown in Tables 6.2-6.5. These may be very different, because for example high IC50 values may have high TDP2 levels but actually result from a mutant form of the gene which has high IC50 value and high TDP2 levels. As is clear from this example it is important to know the effect of each variable on TDP2 levels or TOP2/TDP2 ratios independently which is why these values are reported in Tables 6.2-6.5. Additionally the effect of gene status on etoposide IC50 values can allow better interpretation so these values (again from an ANOVA) are reported in Table 6.1.

2.6 Selection process and statistical analysis of drugs for which TDP2 could potentially be used as a biomarker of sensitivity

2.6.1 Drug Selection

The first step in selecting drugs that were good candidates to be studied for a possible correlation with TDP2 protein levels was to find a database with accurate, curated information about the most drugs and their pharmacological profiling across the largest number of cancer cell lines. This resulted in the selection of CancerDR, a database that has collected information from COSMIC and CCLE databases, curated that information, organised it and presented it in a user-friendly manner (Kumar et al., 2013).

The first list made contained anti-cancer drugs, for which there was information on IC50 values for the breast and lung cancer cell lines that had been studied in Chapters 3 and 4 and had been characterised for TDP2 protein levels and TOP2/TDP2 protein ratios. For both breast and lung cancer cell lines separately, the initial list of drugs (113 for breast and 194 for lung) was reduced to drugs for which IC50 values were available for 5 or more cell lines and where there was some variation in IC50 values. This left 16 drugs for the breast cancer cell lines and 25 for the lung cancer cell lines.

2.6.2 Statistical Analysis

Analysis of this data was performed using R statistical software (Team, 2015). For each drug independently a Pearson correlation coefficient was calculated for the correlation between the IC50 value for the drug of interest and either TDP2 levels, TOP2A/TDP2 ratio or TOP2B/TDP2 ratio. The Pearson correlation coefficient measures the amount that two variables change in a similar manner (co-vary) relative to the variation of the two components. It varies between -1 and 1 with large magnitude values (either positive or negative) showing strong positive and negative correlation for positive or negative correlation coefficients respectively. The significance of the correlation was tested against a null hypothesis that the correlation coefficient was zero (no correlation) using a t-test with degrees of freedom equal to the number of cell lines minus 2. Correlation coefficients (r-values) and associated p-values are shown in Tables 7.1 and 7.2.

2.7 Analysis of relationship of impact of a combination treatment of estradiol, PARP1 inhibitor and etoposide on TDP2 repair in breast cancer cell lines

2.7.1 Alamar Blue Assay and Combination Treatment

For these experiments, the protocol for Alamar Blue was used as described previously, with a few differences described below. On the first day, MCF7, LCC9 and MLET5 cells were seeded, using 10% charcoal-stripped foetal bovine serum in phenol-red-free DMEM with streptomycin, penicillin, L-glutamine and pyruvate to allow the cells to attach. The next day, the cells were treated with 500nM PARP1 inhibitor Ku58948 (AstraZeneca) or DMSO for 30 min, 100nM estradiol or EtOH for 3h and 15μM, 30μM, 45μM etoposide or DMSO for 3h, followed by two washes with PBS and the addition of the appropriate media for each cell line as described previously. The next day, the protocol for Alamar Blue proceeded as detailed in the 'Alamar Blue Assay' section.

2.7.2 Clonogenic Cell Survival Assay and Combination Treatment

For these experiments, the protocol for clonogenic cell survival assay was used as described previously, with the only difference being the media and the treatment. On the first day, MCF7 cells were seeded, using 1% charcoal-stripped foetal bovine serum in phenol-red-free DMEM with streptomycin, penicillin, L-glutamine and pyruvate to allow the cells to attach. The next day, media was changed to 0.1% DMEM as above, to begin cell cycle synchronisation. Five days later, the cells were treated with different combination treatments.

For the first set of experiments, the treatment that was used was either only etoposide (0µM, 15µM, 30µM, 45µM and 60µM) for 3h or 500nM PARP1 inhibitor for 30 min, 100nM estradiol for 3h and 0µM, 15µM, 30µM, 45µM and 60µM etoposide for 3h.

For the second set of experiments, the treatment consisted of 500nM PARP1 inhibitor or DMSO for 30 min, 100nM estradiol or EtOH for 3h and 30µM, 45µM, 60µM etoposide or DMSO for 3h.

For the cells treated with etoposide only the control was DMSO and EtOH, for cells treated with estradiol and etoposide the control was DMSO and estradiol, for cell treated with etoposide and PARP1 inhibitor the control was DMSO/EtOH and PARP1 inhibitor and finally for cells treated with etoposide, estradiol and PARP1 inhibitor, the control was DMSO, estradiol and PARP1 inhibitor.

After the different treatments, the cells were washed twice with PBS and the appropriate media was added, as described under section 'Cell Lines'. After 14 days, cells were fixed as described under section 'Clonogenic Cell Survival Assay'.

2.8 Analysis of the impact of TOP2B-mediated DSBs during estradiol-induced transcription on cell viability and repair kinetics in TDP2-depleted breast cancer cell line MCF7

2.8.1 Establishing a stable TDP2 knock-down in MCF7 cell lines

The first step in creating a stable TDP2 knock-down in MCF7 breast cancer cell lines is to obtain the right plasmid vectors. In this instance the relevant plasmids were pSUPER empty vector for the control knock-down and pSUPER-TDP2 for the TDP2 knock-down. pcD2E was used as the selectable marker, as it will allow transfected MCF7 cells to grow in the presence of G418.

In order to obtain enough quantities of the vectors, a competent *Escherichia coli* (*E.coli*) transformation protocol was used. The *E.coli* strain used was DH5α. Initially, *E.coli* cells were slowly thawed on ice. They were then separated equally (50µl per tube) in as many pre-labelled eppendorf tubes as the number of different plasmid DNA that would be used. 1-5µl of each plasmid DNA was added in the appropriate eppendorf tube and it was mixed gently for a few seconds. That was followed by incubation on ice of 15 min. Then the cells were heat shocked for 45 sec in a water bath at 42°C and left on ice for 1 min afterwards. 500µl of LB was added and the tubes were transferred to a shaking incubator for 1h at 37°C. Finally, the cells were seeded on culture plates with ampicillin and were left to incubate at 37°C for 16h overnight.

The following day, single colonies were picked and were incubated at 37°C shaking for 16h overnight in 100ml of LB containing ampicillin, as a starter culture for Midi preps. The next day, QIAGEN's QIAfilter Midi kit was utilised for purifying the plasmid DNA for each vector. The provided protocol was followed.

The next step was transfection of MCF7 cells with the appropriate vectors. On the first day, 3x 10^{5} MCF7 cells per well were seeded in a 6-well-plate and were incubated at 37°C overnight

until they had reached 50-80% confluency. The following day, 100µl of Opti-MEM media was added in a sterile tube for each well plate that would be transfected. The vector ratio that was used for this transfection was 2:1 for pcD2E, and pSUPER or pSUPPER-TDP2 respectively. It was established through optimisation that the best transfection conditions were 2.5µl of GeneJuice Transfection Reagent for every 0.5µg of plasmid DNA used. Therefore, 7.5µl GeneJuice Transfection Reagent was added drop-wise into each tube containing the Opti-MEM media. They were then mixed by vortexing and incubated at room temperature for 5 min. Following that, 0.5µg of pSUPER or pSUPER-TDP2 and 1µg of pcD2E plasmid DNA were added per tube and were mixed gently by pipetting. The mixture was then incubated for 20 min in room temperature. It was then distributed drop-wise over the entire surface of the well containing the MCF7 cells that were ready for transfection. The cells were placed in an incubator set at 37°C for 24h. The next day, G418 was added to allow for the selection process to begin. After 10 days, the cells were re-seeded in 10mm plates in order for them to grow in G418-selectable colonies. Once the size of the colonies was visible enough to be identified, several colonies were picked and re-seeded to a number large enough for a western blot. They were screened via western blot as described previously for TDP2 and the successful TDP2 knock-down MCF7 cell line was identified, as well as a mock knock-down cell line.

2.8.2 Characterisation of the TDP2 knock-down MCF7 cell line

The first step in characterising the TDP2 knock-down MCF7 cell line was to explore its sensitivity to etoposide, compared to the mock knock-down and the wild type MCF7 cell lines. This was carried out by following the previously described protocol for clonogenic cell curvival assay. Three biological replicates were done. MCF7 cells were treated with increasing concentrations of etoposide (0µM, 15µM, 30µM, 45µM and 60µM) for 3h and then were left to grow for two weeks. On the 14th day, the colonies were fixed with 100% ethanol, washed with PBS and stained with 1% methylene blue in 70% ethanol.

An additional approach to clonogenic cell survival assays was investigating etoposide repair kinetics. In order to do so, a protocol for γH2AX immunolabelling was followed for both TDP2 and mock knock-down MCF7 cell lines. 10^4 cells per well were seeded in normal media on a 4-well plate each containing a single glass coverslip per well. One well per condition was used and the cells were left to attach overnight at 37^oC. The next day, the cells were washed with PBS and 10% charcoal-stripped foetal bovine serum, phenol-red-free DMEM media was added. The cells were left to grow for 3 days. On the fourth day, the cells were treated with DMSO or 20µM etoposide for 30 min. Then the cells were washed with PBS twice and fresh normal media was added. Cells were then collected after 0h, 3h or 6h after treatment and were fixed with 4% PFA for 5 min and then washed with PBS. That was followed by permeabilising the cells by treating them with 0.2% Triton in PBS for 2 min. The cells were then blocked with 5% BSA in PBS for 1h. Blocking was followed by the primary antibodies incubation. The antibodies used were mouse monoclonal anti-γH2AX (Millipore) and rabbit polyclonal anti-CENP-F (Sigma) (Table 2.4). They were diluted in 1% BSA in PBS at 1:1000 and 1:500 dilutions respectively. The incubation was at room temperature for 1h rocking. Following that, the cells were washed three times with PBS with 0.1% Tween and 0.02% SDS (which will be known as PBS-TS) for 5 min each time. That was followed by the secondary antibodies incubation. The antibodies were goat anti-rabbit IgG Alexa Fluor 555 and goat anti-mouse IgG Alexa Fluor 488 (both Invitrogen) (Table 2.5). The antibodies were diluted in 1% BSA in PBS at 1:500 for both. The incubation was carried out at room temperature for 1h on rocker. The cells were then washed three times with PBS-TS, for 5 min each time. 1:10000 DAPI in PBS was then added for 2 min at room temperature protected from light, rocking and then the cells were washed twice in PBS and once in dH_2O, 5 min each time. Finally, the glass coverslips were mounted on a drop of Vectashield (Vector Labs) placed on slides and the edges were sealed with clear nail polish. Microscopy was then carried out with a Nikon E400 or a Zeiss Axioplan 2. Only cells negative for CENP-F were counted.

Table 2.4: Primary antibodies for immunofluorescence

Primary Antibodies	Type	Organism	Company
anti-γH2AX	Monoclonal	Mouse	Millipore
anti-CENP-F	Polyclonal	Rabbit	Sigma

Table 2.5: Secondary antibodies for immunofluorescence

Secondary Antibodies	Company
Goat anti-rabbit IgG Alexa Fluor 555	Invitogen
Goat anti-mouse IgG Alexa Fluor 488	Invitrogen

To further characterise the TDP2 knock-down MCF7 cell line, Alamar Blue was also utilised. The protocol that was followed was the same one described previously. The TDP2 knock-down cell line was compared to the mock knock-down and the wild type MCF7 cell lines. Initially, high concentrations of etoposide were used in an acute manner, which were 0μM, 25μM, 50μM, 100μM and 150μM. Then another set of experiments was carried out with lower etoposide concentrations (0μM, 5μM, 10μM, 20μM and 40μM). Finally, a set of experiments with continuous etoposide treatment was utilised and was done the same way as previously described. The concentrations were 0nM, 250nM, 500nM, 750nM and 1μM.

2.8.3 Effects of etoposide on TDP2 knock-down MCF7 breast cancer cells after induction of transcription by estradiol treatment

In order to explore the effects of treating TDP2 knock-down MCF7 breast cancer cells with etoposide once transcription has been induced by estradiol treatment, the first thing is to establish the best conditions for induction of transcription in MCF7 breast cancer cells. The conditions that were used were 100nM estradiol for 3h, 6h, 12h and 24h (24h EtOH was the

control). 5x 10^5 TDP2 knock-down, mock knock-down or wild type MCF7 cells were seeded on 100mm plates in normal media and left to attach overnight. The next day, the media was changed to 1% charcoal-stripped foetal bovine serum in phenol-red-free DMEM and the cells were left to synchronise for 5 days. After synchronisation, the cells were treated with estradiol for the amount of time stated above. Once the treatment time was completed, cells were washed with PBS, were detached from the plate with Trypsin and frozen in LN_2 until all the samples had been collected.

The samples were then processed for RNA extraction. QIAGEN's RNeasy Plus Mini Kit was used and the provided protocol was followed with the modifications below. For the Buffer RPE, 300µl was added and the column was spun for 3 min, then 10µl of RQ1 DNAse and 10x Buffer (8µl RNAse free H_2O, 1µl RQ1 DNAse and 1µl 10x Buffer) was added and the samples were incubated at room temperature for 5 min. That was followed by the addition of 400µl Buffer RPE and a 3-min spin, and finally the addition of 300µl Buffer RPE and another 3-min spin. The protocol was then resumed with placing the RNeasy spin columns in new 2ml collection tubes. Once all the RNA samples were obtained, they were quantified by NanoDrop and its accompanying software. Two measurements were taken per sample and the average was calculated. High-capacity cDNA reverse transcription kit without RNAse inhibitor by Applied Biosystems was used to obtain cDNA from these samples, as described previously. SYBR Green qPCR was then utilised to assess induction of transcription. The genes that were used were TFF1 and GREB1, with TBP and Actin as controls. Primers can be seen in Table 2.6.

Table 2.6: SYBR Green qPCR primers
TFF1 Forward Primer Sequence GAG-AAC-AAG-GTG-ATC-TGC-GC
TFF1 Reverse Primer Sequence TGG-TAT-TAG-GAT-AGA-AGC-ACC
GREB1 Forward Primer Sequence GTG-GTA-GCC-GAG-TGG-ACA-AT
GREB1 Reverse Primer Sequence AAA-CCC-GTC-TGT-GGT-ACA-GC
Actin Forward Primer Sequence GAA-GTG-TGA-CGT-GGA-CAT-CC
Actin Reverse Primer Sequence CTC-GTC-ATA-CTC-CTG-CTT-GC
TBP Forward Primer Sequence CAC-GAA-CCA-CGG-CAC-TGA-TT
TBP Reverse Primer Sequence TTT-TCT-TGC-TGC-CAG-TCT-GGA-C

For each sample, 12.5µl QIAGEN's SYBR Green mix, 10.5µl ddH$_2$0, 1µl of primer mix (containing both forward and reverse primers for each gene, 10nM concentration) and 1µl cDNA (10ng/µl concentration) was added in a qPCR tube (Agilent Technologies). All samples were run in technical triplicates. The cycling parameters used were:

- PCR initial heat activation – 95°C for 10 min

- 3-Step cycling (x40):

  - Denaturation – 95°C for 15 sec

  - Annealing – 53°C for 15 sec

  - Extension – 60°C for 30 sec

The qPCR machine that was used was an MX3005P Stratagen, with the accompanying MxPro QPCR Software.

The next step was to visualise TOP2B-induced DSBs during induction of transcription by estradiol. The protocol for γH2AX immunolabelling was followed as described previously, for TDP2 and mock knock-down MCF7 cell lines, as well as wild type MCF7. The differences were during the treatment. The cells were treated with estradiol for 2h, 6h, 24h and 48h. The control used was a 48h-treatment with EtOH. As with the γH2AX repair kinetics described above, γH2AX foci counts were conducted blind with either a Nikon E400 or a Zeiss Axioplan 2. Only cells negative for CENP-F were counted.

In order to verify beyond any doubt that only G1 cells have been counted and not G2 or S-phase, and any difference between γH2AX foci numbers is due to the lack of TDP2, EdU labelling was utilised. The treatment part of the experiment was performed as described in the previous paragraph. 15 min before the cells were to be collected, 10μM EdU was added so it could be incorporated in the cells that are synthesising DNA to distinguish S-phase cells. Following that the cells were fixed as previously described with PFA. Invitrogen's Click-iT Imaging Kit was then used and the accompanying protocol was followed from step 3.6 to step 4.1. After that, DAPI was used as previously described. The cells were visualised with either a Nikon E400 or a Zeiss Axioplan 2. In addition to the acute EdU labelling described, a continuous EdU labelling protocol was followed. The only difference between the acute and the continuous was that during the continuous, EdU was added at the same time as the estradiol instead of 15 min before the cells were due to be collected.

The next step was to explore any differences between the TDP2 and the mock knock-down MCF7 cell lines by clonogenic cell survival assays. The protocol that was followed was the same as described previously. The differences were that 2000 cells/plate were seeded in 10% charcoal-stripped foetal bovine serum in phenol-red-free DMEM. The treatments were the following for each condition:

- 100nM estradiol or ETOH for 3h, then 15µM, 30µM, 45µM and 60µM etoposide of DMSO for 3h together with the estradiol.

- 100nM estradiol or EtOH for control for 3h.

Finally, one more experiment that utilised γH2AX foci was done. The protocol for visualising DSBs via γH2AX foci was followed, but the treatment was either estradiol, etoposide (50nM), or both for 24h only. The cells were collected and fixed with PFA, as previously described and the γH2AX foci protocol was followed. As before, the counts were conducted blind with either a Nikon E400 or a Zeiss Axioplan 2. Only cells negative for CENP-F were counted.

3. CHAPTER THREE - Characterisation of TDP2, TOP2A and TOP2B mRNA and protein levels in lung cancer cell lines

3.1 Introduction

Recently, data that have emerged suggest that expression levels of TDP2 vary greatly in different cancer cell lines, ranging from almost no expression to very high levels of expression (Li et al., 2011a). Since TDP2 repairs abortive TOP2-mediated DSBs, it is possible that levels of TDP2 correlate with cancer cell resistance to TOP2 poisons such as etoposide, which are widely used in cancer treatment. Published studies that have directly manipulated TDP2 protein levels are encouraging in this respect. For example, TDP2 depletion by RNAi in the lung cancer cell line A549 results in increased sensitivity to etoposide, an established and specific TOP2 poison, while the addition of recombinant human TDP2 increases resistance (Cortes Ledesma et al., 2009). The same is true in chicken DT40 cells, in which deletion of TDP2 similarly results in an increased sensitivity to etoposide (Zeng et al., 2011). This suggests that cancers possessing high levels of TDP2 might be relatively resistant to TOP2 poisons, while cancers with low levels of TDP2 might be hypersensitive to such treatment. Consistent with this idea, increased levels of TDP2 have been observed in several non-small cell lung cancer cell lines and a possible oncogenic role of TDP2 in such cancers has been suggested (Li et al., 2011a). Non-small cell lung cancers represent 80-85% of all lung cancer cases and are thus of major clinical concern (Breathnach et al., 2001, Bareschino et al., 2011). In addition to variations in TDP2 expression across different tumour types, chemoresistance frequently occurs. One possible contributing factor to this is increased DNA repair, and TDP2 has been reported to contribute to such resistance (Li et al., 2011b). As a first step to establishing the utility of TDP2 as biomarker of sensitivity to topoisomerase 2-targeting agents in lung cancer, I examined whether differences exist in TDP2 mRNA and protein levels across a panel of lung cancer cell lines.

3.1.1 Aims of this chapter

The aim of this chapter was to examine the possible utility of TDP2 as biomarker for therapeutic sensitivity to TOP2 poisons, by measuring its mRNA expression and protein levels in a panel of lung cancer cell lines.

3.2 Results

3.2.1 TDP2, TOP2A and TOP2B mRNA expression levels vary significantly in lung cancer cell lines.

To measure TDP2 expression levels in a panel of lung cancer cell lines, I first measured TDP2 mRNA using RT-qPCR. Eight lung cancer cell lines were employed, which were A549, H292, H1650, H460, H226, MOR, Hop92 and HCC95. All eight cell lines are non-small cell lung cancer cell lines, which is the most common type of lung cancer. These cell lines were chosen because they are all well characterised.

Cell pellets were collected for each cell line and were immediately frozen in LN_2, and stored at -20°C until use. RNA was extracted and RT-qPCR was performed in triplicate to obtain the average Ct values for each sample. Human RPLP0 (large ribosomal protein) was employed as a "housekeeping" control for the calculation of the relative expression of TDP2. The RPLP0 cDNA nucleotide sequence is consisted of highly conserved regions in the 5-prime end of its open reading frame (Liu et al., 2005, Lyng et al., 2008). The expression levels of RPLP0 remain the same across different tissue types, which will become more relevant in the next chapter, where breast cancer cell lines were used for a similar set of experiments (Liu et al., 2005, Lyng et al., 2008). In addition, when compared to the transcription levels of other common reference genes, RPLP0 was found to be more consistent and reliable (Schmid Lab; unpublished observations), which agrees with previous studies (Lyng et al., 2008).

TDP2 expression levels, relative to the RPLP0 control, are shown in Figure 3.1A, with a representative amplification profile for TDP2 and the control shown in Figure 3.1B. In the amplification curve, the highest level of expression is that for which the amplification fluorescence crosses the threshold value in the smallest number of cycles; in this case 27 cycles. From Figure 3.1A it can be seen that Hop92 is the cell line with the highest TDP2 expression level, whereas H292 and H226 are the cell lines with the lowest TDP2 expression levels. The difference in expression levels between Hop92 and H292/H226 is 4.5-fold, and is statistically significant. Of those cell lines between these extremes, A549, H1650, and MOR have ~2-fold more TDP2 mRNA than H292 and H226.

Similarly, the expression levels of TOP2A (Fig.3.2) and TOP2B (Fig.3.3) were also measured using RT-qPCR. As TDP2 is responsible for repairing abortive TOP2-induced DNA double-strand breaks, I believed it was important to determine the levels of both of these proteins, as it would be an essential part of the direction of the project in the future, especially when trying to find a possible correlation between sensitivity to a TOP2 poison and TDP2 levels (Chapter 7). Both Figures 3.2 and 3.3 clearly show a variation in TOP2A and TOP2B mRNA levels. Figure 3.2A indicates that the cell lines with the lowest levels of TOP2A mRNA are H460 and Hop92, with A549, H1650, MOR and HCC95 exhibiting the highest expression levels. The difference between H460/Hop92 and A549/H1650/MOR/HCC95 is more than 2-fold in magnitude. Figure 3.3A reveals that the highest level of TOP2B mRNA is present in H1650, with all of the other cell lines (A549, H292, H460, H226, MOR, Hop92 and HCC95) exhibiting 2-4-fold lower levels.
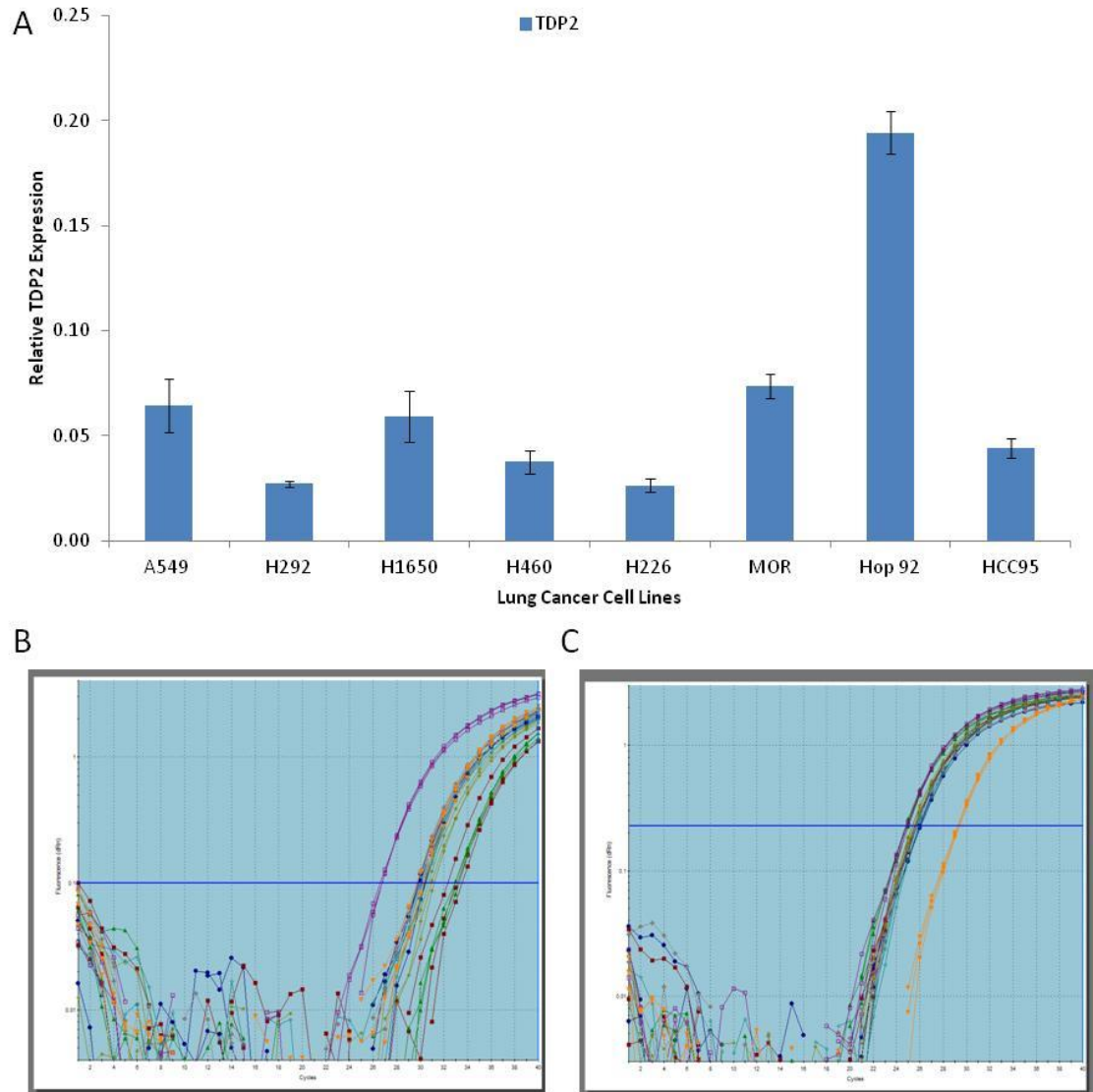
Figure 3.1: TDP2 mRNA levels in lung cancer cell lines. A) TDP2 mRNA expression levels were measured using RT-qPCR in the indicated lung cancer cell lines, and are expressed relative to the expression level of the RPLP0 housekeeping gene. Relative expression (fold-change) was calculated from the equation relative expression=$2^{-\Delta Ct}$, where ΔCt is calculated by subtracting the Ct value of RPLP0 from the Ct value of TDP2. Data are the mean (+/- SEM) from three independent experiments. B) Representative RT-qPCR amplification plot of TDP2. Amplification plots represent the accumulation of product over the duration of the RT-qPCR experiment. Rn is the fluorescence of the TDP2 primer divided by the fluorescence of the internal reference dye (ROX). dRn is Rn minus the baseline (adjusted automatically by MXPro for qPCR software),and is plotted against PCR cycle number. C) Representative RT-qPCR amplification plot of RPLP0 housekeeping gene.
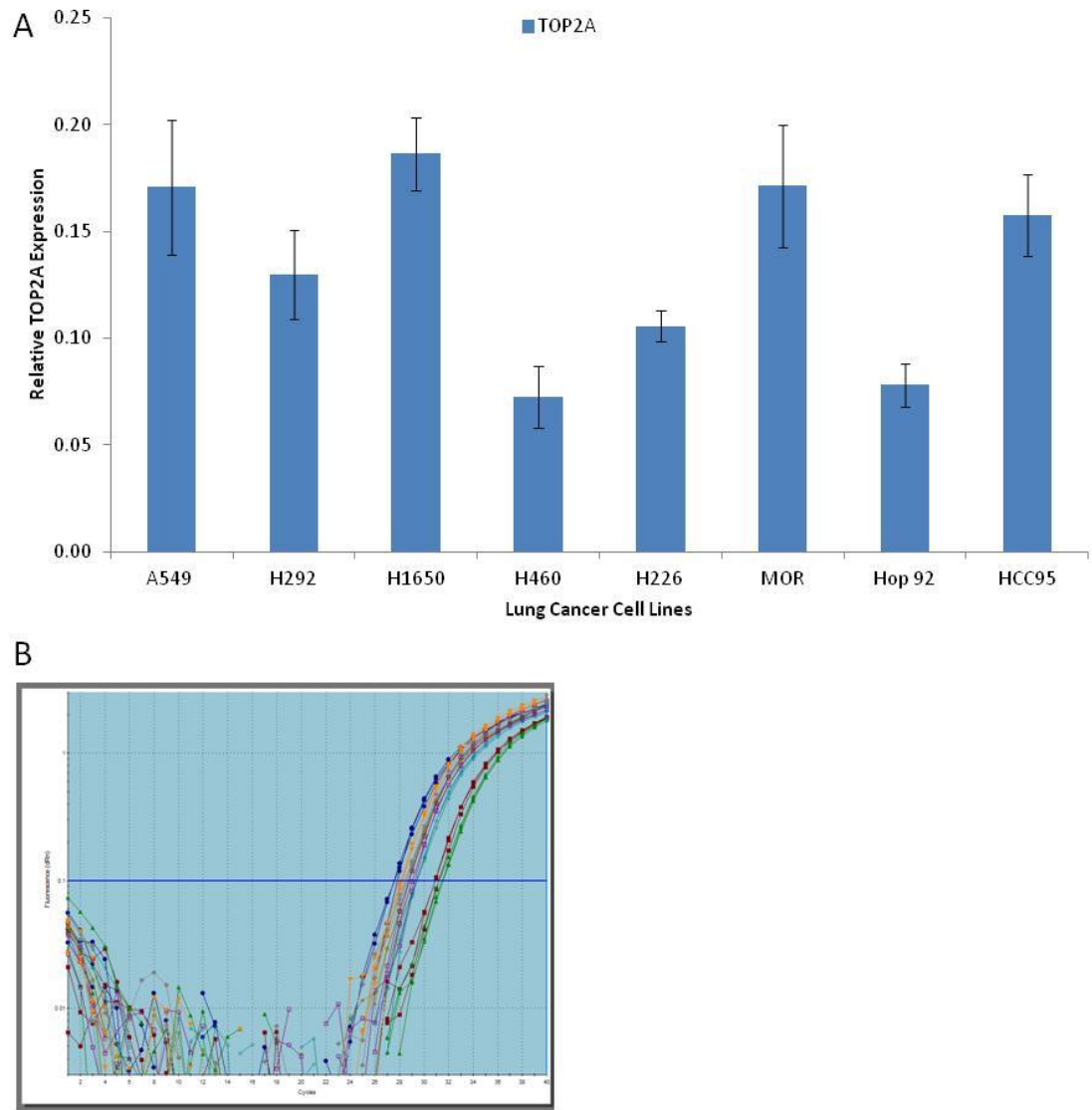
Figure 3.2: TOP2A mRNA levels in lung cancer cell lines. A) TOP2A mRNA expression levels were measured using RT-qPCR in the indicated lung cancer cell lines, and are expressed relative to the expression level of the RPLP0 housekeeping gene. Relative expression (fold-change) was calculated from the equation relative expression=$2^{-\Delta Ct}$, where ΔCt is calculated by subtracting the Ct value of RPLP0 from the Ct value of TOP2A. Data are the mean (+/- SEM) from three independent experiments. B) Representative RT-qPCR amplification plot of TOP2A. Amplification plots represent the accumulation of product over the duration of the RT-qPCR experiment. Rn is the fluorescence of the TOP2A primer divided by the fluorescence of the internal reference dye (ROX). dRn is Rn minus the baseline (adjusted automatically with MXPro for qPCR software),and is plotted against PCR cycle number.
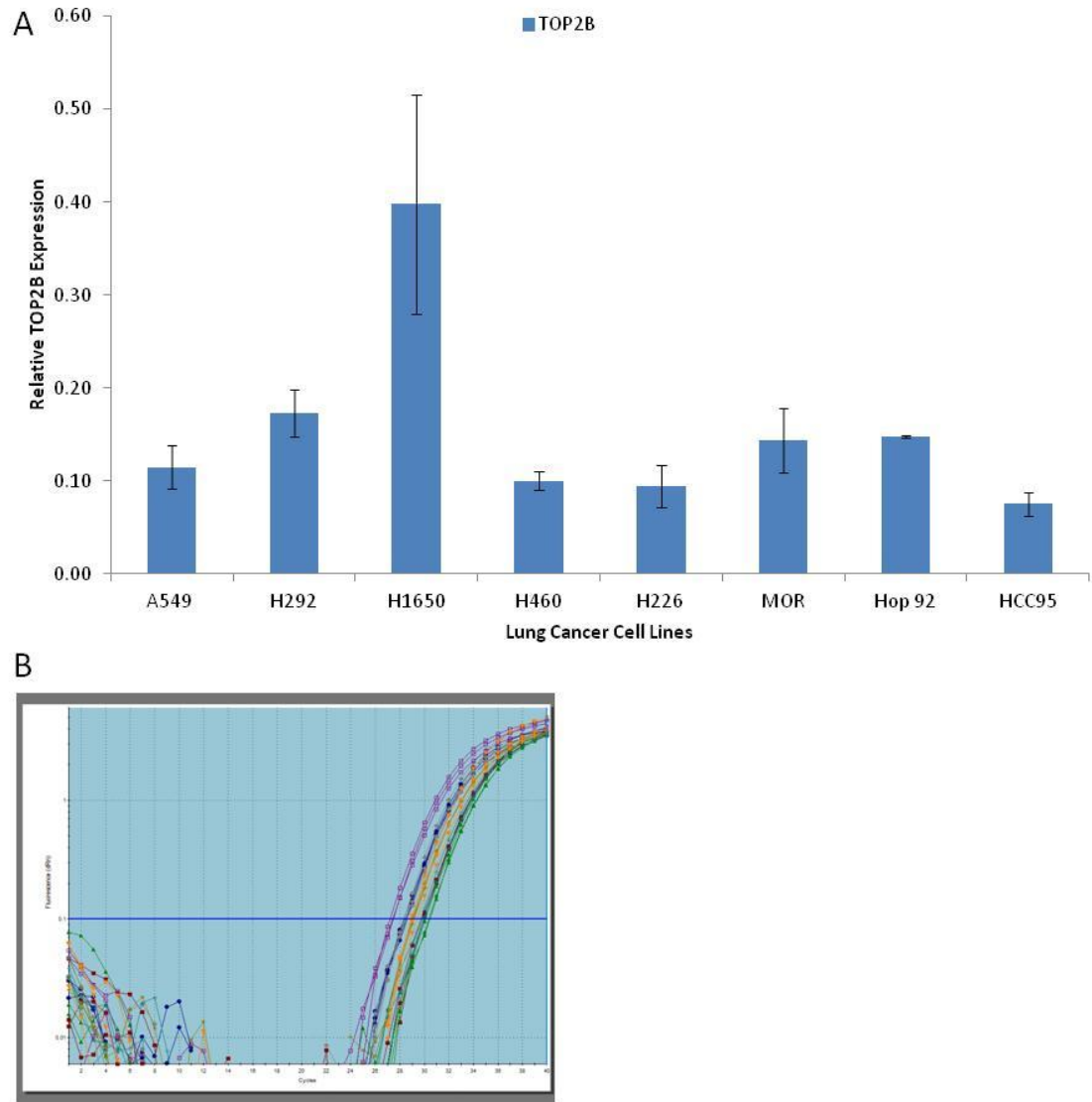
Figure 3.3: TOP2B mRNA levels in lung cancer cell lines. A) TOP2B mRNA expression levels were measured using RT-qPCR in the indicated lung cancer cell lines, and are expressed relative to the expression level of the RPLP0 housekeeping gene. Relative expression (fold-change) was calculated from the equation relative expression=$2^{-\Delta Ct}$, where ΔCt is calculated by subtracting the Ct value of RPLP0 from the Ct value of TOP2B. Data are the mean (+/- SEM) from three independent experiments. B) Representative RT-qPCR amplification plot of TOP2B. Amplification plots represent the accumulation of product over the duration of the RT-qPCR experiment. Rn is the fluorescence of the TOP2B primer divided by the fluorescence of the internal reference dye (ROX). dRn is Rn minus the baseline (adjusted automatically by MXPro for qPCR software),and is plotted against PCR cycle number.

3.2.2 TDP2, TOP2A and TOP2B protein levels vary significantly in lung cancer cell lines.

Because mRNA levels do not necessarily reflect the level of protein expression, I conducted western blotting to compare levels of TDP2, TOP2A and TOP2B protein.

Whole cell extracts prepared from each of the eight cell lines were analyzed by western blotting using antibodies against TDP2, TOP2A, TOP2B, and Actin. Actin was used as a control, to normalise the TDP2, TOP2A and TOP2B signals for differences in sample loading. Three different cell extract protein concentrations were employed for each sample, enabling me to ensure that quantification of the western blot signal was linearly related to protein quantity and was thus accurate. The first protein concentration for each sample is denoted C1, the second is denoted C2 and was twice that of C1, and the third is denoted C3 and was four-fold that of C1. Figure 3.4 reveals that MOR and HCC95 possess the highest levels of TDP2 protein, and H1650 the lowest levels, while the remaining cell lines (A549, H292, H460, H226 and Hop92) somewhere in between. The difference in TDP2 protein levels between H1650 (lowest expressor) and MOR/HCC95 (highest expressors) was 6-10-fold, depending on the amount of cell extract protein employed. Figure 3.4E and 3.4F reveal a linear relationship between TDP2 band intensity and amount of cell extract protein loaded on the gels, with correlations (R^2) close to 1 in all cases.

Figure 3.5 indicates that the cell lines A549, H226, and MOR possess the highest levels of TOP2A, and that H1650 and Hop92 possess the lowest. The difference between highest and lowest TOP2A levels was 6-12-fold. Most of the R^2 values (Fig.3.5E and 3.5F) were close to 1, with the exception of H226, suggesting that the relationship between band intensity and quantity of protein extract employed was close to linear. Similarly, H226 and MOR possessed the highest levels of TOP2B, with H1650, H460, Hop92 and HCC95 the lowest, with a variation of ~5-fold. Once again, R^2 values were very close to 1 (Fig.3.6).
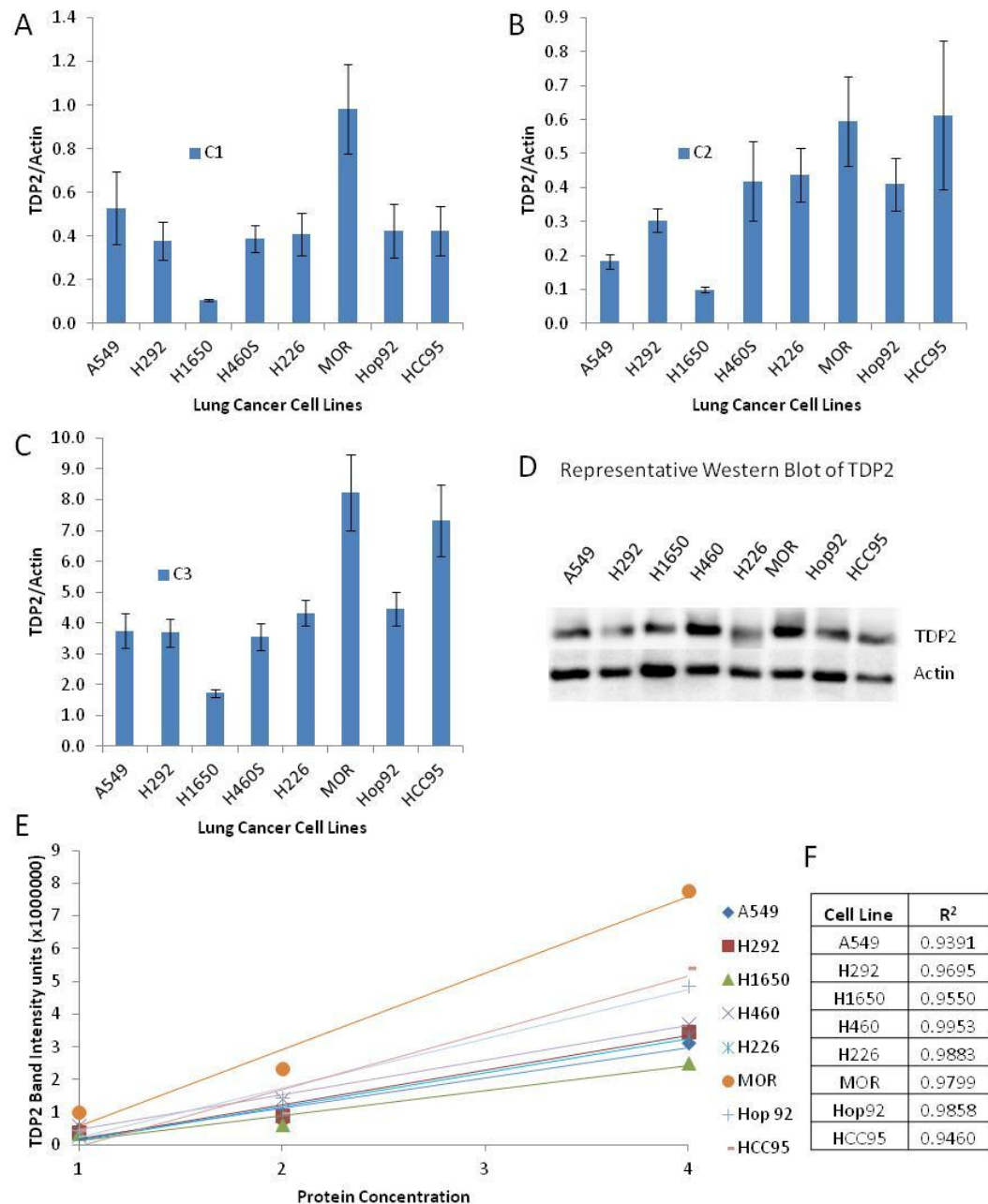
Figure 3.4: TDP2 protein levels in lung cancer cell lines. Panels A-C; Different amounts of cell extract protein from the indicated cell lines were fractionated by SDS-PAGE and subjected to Western blotting with anti-TDP2 and anti-Actin antibodies. TDP2 band intensities were quantified by densitometry and normalised for differences in loading by dividing by the band intensity for Actin. Data in Panels A-C are the mean of three independent experiments (+/- SEM). Gels employed in Panel B and Panel C contained 2x and 4x the amount of cell extract protein as that employed for the gels in Panel A. Panel D; Representative Western blots for TDP2 and Actin. Panel E; Plots of the TDP2 band intensities from Panels A-C. Panel F; R^2 values of the plots in panel E, demonstrating the linear relationship between TDP2 band intensity and amount of protein extract employed in panels A-C.

Figure 3.5: TOP2A protein levels in lung cancer cell lines. Panels A-C; Different amounts of cell extract protein from the indicated cell lines were fractionated by SDS-PAGE and subjected to Western blotting with anti-TOP2A and anti-Actin antibodies. TOP2A band intensities were quantified by densitometry and normalised for differences in loading by dividing by the band intensity for Actin. Data in Panels A-C are the mean of three independent experiments (+/- SEM). Gels employed in Panel B and Panel C contained 2x and 4x the amount of cell extract protein as that employed for the gels in Panel A. Panel D; Representative Western blots for TOP2A and Actin. Panel E; Plots of the TOP2A band intensities from Panels A-C. Panel F; R^2 values of the plots in panel E, demonstrating the linear relationship between TOP2A band intensity and amount of protein extract employed in panels A-C.

Figure 3.6: TOP2B protein levels in lung cancer cell lines. Panels A-C; Different amounts of cell extract protein from the indicated cell lines were fractionated by SDS-PAGE and subjected to Western blotting with anti-TOP2B and anti-Actin antibodies. TOP2B band intensities were quantified by densitometry and normalised for differences in loading by dividing by the band intensity for Actin. Data in Panels A-C are the mean of three independent experiments (+/- SEM). Gels employed in Panel B and Panel C contained 2x and 4x the amount of cell extract protein as that employed for the gels in Panel A. Panel D; Representative Western blots for TOP2B and Actin. Panel E; Plots of the TOP2B band intensities from Panels A-C. Panel F; R^2 values of the plots in panel E, demonstrating the linear relationship between TOP2B band intensity and amount of protein extract employed in panels A-C.

3.3 Discussion

The RT-qPCR and western blot data in this chapter reveal there is a clear variation in mRNA and protein levels in TDP2, TOP2A and TOP2B. The results obtained by RT-qPCR and western blotting, however, were not the same. It is not unusual that protein levels do not reflect mRNA levels, since there are many mechanisms that act after translation that have an impact on protein production and stability (Vogel and Marcotte, 2012, Stark et al., 2006, Dickson et al., 2007, Taquet et al., 2009). Nevertheless, both provide information in regard to how variant TDP2, TOP2A and TOP2B mRNA and protein levels are in different lung cancer cell lines.

TDP2 has already been linked to cancer development and progression. In addition, it is believed to protect cancer cells from entering apoptosis, a hallmark of cancer (Gomez-Herreros et al., 2013). A previous study, the only one that had investigated TDP2 protein levels in cancer cells before the start of this project, showed that in non-small cell lung cancer cell lines the levels of TDP2 varied, with H460, A459 and H226 demonstrating medium to high levels, and H292 and H1650 showing almost no TDP2 (Li et al., 2011a). Furthermore, in the same study, knocking down TDP2 in H460 lung cancer cells led to apoptosis (Li et al., 2011a). The TDP2 antibody that was used for this study is different to the one used in this project and side by side comparisons showed it was non-optimal, but it still showed that TDP2 protein levels vary in lung cancer cell lines, which is what the results of this chapter also reveal.

For TDP2, MOR and HCC95 possess the highest levels of TDP2 protein, followed by A549, H292, H460, H226 and Hop92 with intermediate levels, and H1650 with the lowest levels. For TOP2A, A549, H226, and MOR possess the highest protein levels, followed by H292, H460 and HCC95 with intermediate levels, and H1650 and Hop92 with the lowest. For TOP2B, H226 and MOR possess the highest protein levels, followed by A549 and H292 with intermediate levels and, H1650, H460, Hop92 and HCC95 with the lowest levels. It is clear that the levels of these three proteins do not directly correlate with each other. Just because a cell line has very high levels

of TDP2, doesn't mean it will also have very high levels of TOP2A and TOP2B, although that doesn't mean that it can't, MOR being an example of that. But what is interesting to explore is what impact high levels of TDP2 might have in addition to low levels of TOP2A and TOP2B or the opposite, in how a cell line responds to treatment with a drug that targets TOP2. Would high levels of TOP2A and TOP2B and low levels of TDP2 translate to low resistance to a TOP2 poison? Would TDP2 be overwhelmed with the high volume of TOP2-induced DSBs made abortive by the interference of the drug, leading the cell to enter apoptosis? And on the other hand, would high levels of TDP2 in combination with low levels of TOP2A and TOP2B result in high resistance to a TOP2 drug?

The variation in protein among the three different concentrations of protein used in TDP2, TOP2A and TOP2B, can probably be attributed to the affinity of antibodies used. It is clear that the TDP2 antibody has high affinity, while TOP2A has affinity to a lesser extent. The TOP2B antibody is the least capable out of the three, but nevertheless still demonstrates a variation in protein levels within the lung cancer cell line panel.

The results of this chapter mean that potentially TDP2 could be a biomarker of sensitivity to TOP2 targeting agents, at least in lung cancer cell lines. Cell lines with low levels of TDP2 and high levels of TOP2A and TOP2B, such as H226 and A549, would be the ideal target for treatment with a TOP2 poison. High levels of TOP2A and TOP2B could result in elevated TOP2-mediated abortive DSBs within a cell, while the low levels of TDP2 would be unable to repair these breaks, resulting in hypersensitivity. On the other end of the spectrum, cell lines with high levels of TDP2 and low levels of TOP2A and TOP2B, such as HCC95 and Hop92, would be the least desirable candidates, as the situation would be reversed with an abundance of TDP2 free to deal with fewer TOP2-mediated DSBs. Table 3.1 summarizes the protein levels for TDP2, TOP2A and TOP2B in the cell lines of the lung cancer panel.

Next, I repeated these experiments to explore the relative expression levels of TDP2, TOP2A and TOP2B in breast cancer cell lines.

Table 3.1: Rank order of TDP2, TOP2A and TOP2B protein levels in lung cancer cell line panel. 8 corresponds to the highest protein levels and 1 corresponds to the lowest levels.

	TDP2	TOP2A	TOP2B
A549	2	7	5
H292	3	5	6
H1650	1	1	1
H460	4	3	3
H226	5	8	8
MOR	8	6	7
Hop92	6	2	4
HCC95	7	4	2

4. CHAPTER FOUR - Characterisation of TDP2, TOP2A and TOP2B mRNA and protein levels in breast cancer cell lines

4.1 Introduction

As demonstrated in the previous chapter, TDP2, TOP2A and TOP2B mRNA and protein levels vary in different lung cancer cell lines. The next step was to explore mRNA and protein levels in a panel of breast cancer cell lines, to further examine the possible utility of TDP2 as a biomarker of sensitivity to TOP2 poisons in this type of cancer, as discussed in the previous chapter.

There is an emerging model which suggests that inducing transcription by simulating breast cancer cells with estrogen leads to the formation of DSBs. These DSBs are mediated by TOP2B, which is recruited with the estrogen receptor to regulatory sites on target genes for efficient activation of those genes (Haffner et al., 2011, Ju et al., 2006, Williamson and Lees-Miller, 2011). Failure to repair those DSBs, due to a TOP2B-targeting drug for example, could lead to apoptosis. This could potentially be more effective in breast cancer cell lines with low levels of TDP2 and higher levels of TOP2B, as this would mean there would not be enough TDP2 to repair the TOP2B-induced DSBs, increasing the chance of success of such a treatment.

Breast cancer is the most common cancer type in women worldwide (Siegel et al., 2015). Around 20-30% of breast cancer cases will become metastatic (O'Shaughnessy, 2005). In addition, about 6% of non pre-existing breast cancer cases begin as metastatic, also known as Stage 4 (Program, 2015). Etoposide is a commonly employed drug that is regaining popularity for patients that suffer from metastatic breast cancer and have received previous treatment (Yuan et al., 2015, Valabrega et al., 2015). Therefore, establishing a TDP2 profile for these patients could potentially predict the possibility of success of such a treatment, following the

hypothesis explained previously. A panel consisting of eight breast cancer cell lines was utilised for my thesis to further explore this potential.

4.1.1 Aims of this chapter

The aim of this chapter was to examine the possible utility of TDP2 as biomarker for therapeutic sensitivity to TOP2-poisons, by measuring its mRNA expression and protein levels in a panel of breast cancer cell lines.

4.2 Results

4.2.1 TDP2, TOP2A and TOP2B mRNA expression levels vary significantly in breast cancer cell lines.

To demonstrate a potential variation in TDP2 expression levels in a panel of breast cancer cell lines, TDP2 mRNA levels were measured using RT-qPCR. Initially six breast cancer cell lines were used, which were MDA-MB-453 (MM453), MDA-MB-361 (MM361), T47D, MCF7, ZR-75-1 (ZR751), BT-474 (BT474), LCC9 and MLET5. MM453 and MCF7 were added to this panel shortly after, to expand the types of breast cancer cell lines to include an ER negative cell line (MM453) and because MCF7 is the cell line from which LCC9 and MLET5 are derived from. MCF7 is ERα positive and grows in an estrogen-dependent manner, but both LCC9 and MLET5 whilst ERα positive, grow in an estrogen-independent manner

Cell pellets were collected, immediately frozen in LN_2, and stored at -20^oC until use. RNA was extracted and RT-qPCR was performed in triplicate to obtain the average Ct values for each sample, as with the lung cancer cell lines. Human RPLP0 (large ribosomal protein) was employed as a "housekeeping" control for the calculation of the relative expression of TDP2. As mentioned in the previous chapter, the RPLP0 cDNA nucleotide sequence is consisted of highly conserved regions in the 5-prime end of its open reading frame and its expression levels remain the same across different tissue types (Lyng et al., 2008, Liu et al., 2005).

TDP2 expression levels, relative to the RPLP0 control, are shown in two groups in Figures 4.1A

and B, with a representative amplification profile for TDP2 per group shown in Figures 4.1C

and D. In the amplification curve, the highest level of expression is that for which the

amplification fluorescence crosses the threshold value in the smallest number of cycles; in this

case 27 cycles. In the first group of breast cancer cell lines examined, T47D exhibited the

lowest TDP2 mRNA levels and LCC9 the highest, with a 7-fold difference between them. In the

second set, comprised of just MM453 and MCF7, the latter exhibited the lowest level, at 2-fold

lower than MM453. It is not possible to directly compare the two groups to each other and

one can observe that the scales of the Fold Expression graphs in Figures 4.1A and B are very

different. That can be attributed to several different factors. The two series of experiments

were done at different times during this project and the cells for each group were grown in

different labs. In addition, all subsequent methods used to acquire the RT-qPCR results were

also done in different labs. The materials used were also different for each group, such as the

master mix, the pipettes etc. Although there are differences in the scales, the error bars

suggest that the results of the second group are consistent.

As with the lung cancer cells, the expression levels of TOP2A and TOP2B were also measured

using RT-qPCR. These results are shown in Figure 4.2 and Figure 4.3, respectively. These

experiments revealed a 5-fold difference in TOP2A mRNA levels between the cell lines with the

lowest levels (exhibited by MLET5, T47D and MM361), and the cell line with the highest TOP2A

mRNA level (LCC9). In the second group, TOP2A mRNA levels in MCF7 and MM453 were not

significantly different. Levels of TOP2B mRNA were generally similar in terms of variation

across the cell lines to those observed for TOP2A (Fig.4.3). For example, similar to TOP2A

mRNA levels, LCC9 possessed the highest level of TOP2B mRNA, with T47D and MLET5

amongst the lowest and more than 10-fold lower than LCC9.
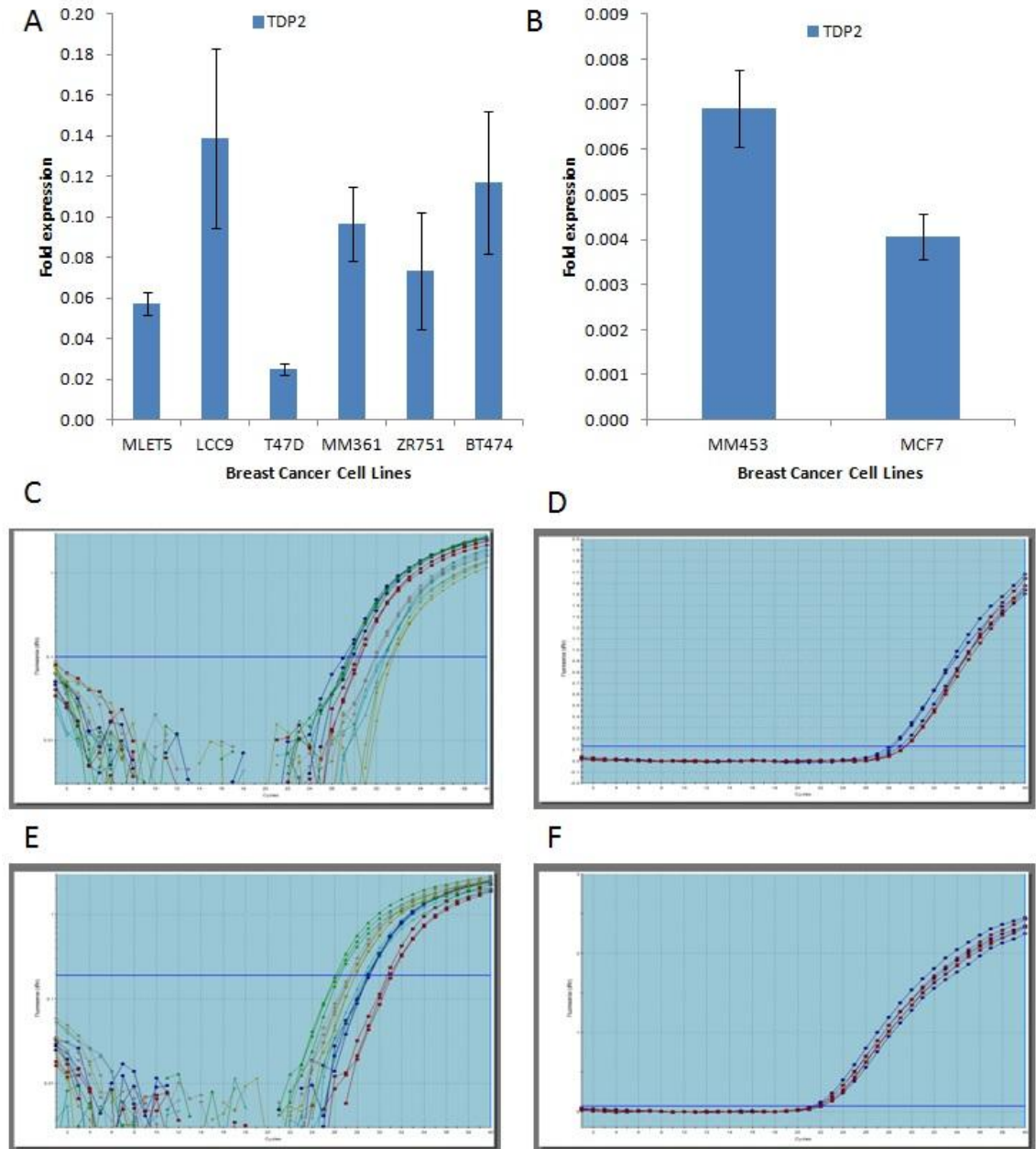
Figure 4.1: TDP2 mRNA levels in breast cancer cell lines. A,B) TDP2 mRNA expression levels were measured using RT-qPCR in the indicated breast cancer cell lines, and are expressed relative to the expression level of the RPLP0 housekeeping gene. Relative expression (fold-change) was calculated from the equation relative expression=$2^{-\Delta Ct}$, where ΔCt is calculated by subtracting the Ct value of RPLP0 from the Ct value of TDP2. Data are the mean (+/- SEM) from three independent experiments. C) Representative RT-qPCR amplification plots of TDP2 for MLET5, LCC9, T47D, MM361, ZR751 and BT474. Amplification plots represent the accumulation of product over the duration of the RT-qPCR experiment. Rn is the fluorescence of the TDP2 primer divided by the fluorescence of the internal reference dye (ROX). dRn is Rn minus the baseline (adjusted automatically by MXPro for qPCR software),and is plotted against PCR cycle number. D) Representative RT-qPCR amplification plots of TDP2 for MM453 and MCF7. E) Representative RT-qPCR amplification plot of the housekeeping gene RPLP0 for MLET5, LCC9, T47D, MM361, ZR751 and BT474. F) Representative RT-qPCR amplification plot of the housekeeping gene RPLP0 for MM453 and MCF7.
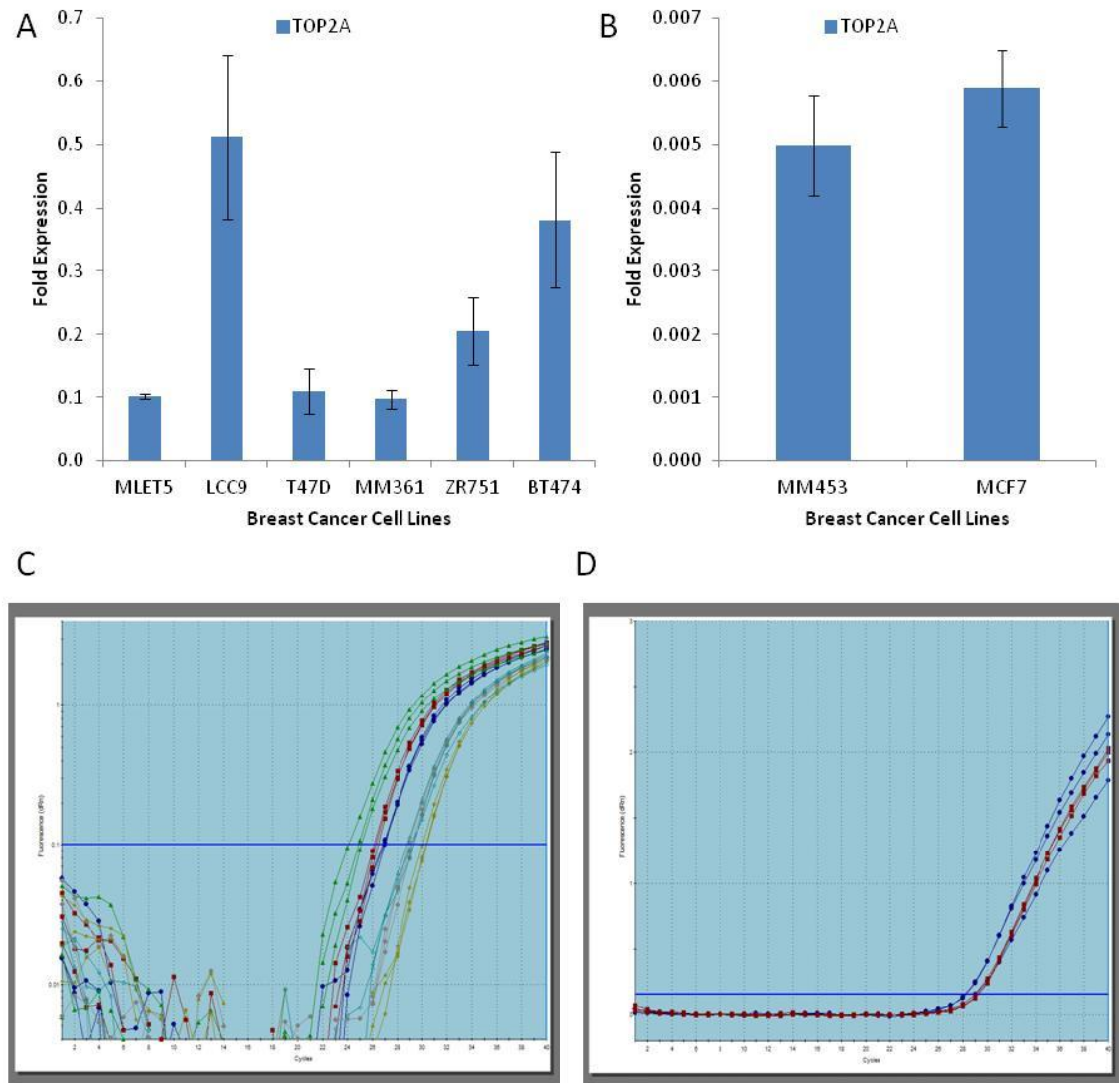
Figure 4.2: TOP2A mRNA levels in breast cancer cell lines. A,B) TOP2A mRNA expression levels were measured using RT-qPCR in the indicated breast cancer cell lines, and are expressed relative to the expression level of the RPLP0 housekeeping gene. Relative expression (fold-change) was calculated from the equation relative expression=$2^{-\Delta Ct}$, where ΔCt is calculated by subtracting the Ct value of RPLP0 from the Ct value of TOP2A. Data are the mean (+/- SEM) from three independent experiments. C) Representative RT-qPCR amplification plots of TOP2A for MLET5, LCC9, T47D, MM361, ZR751 and BT474. Amplification plots represent the accumulation of product over the duration of the RT-qPCR experiment. Rn is the fluorescence of the TOP2A primer divided by the fluorescence of the internal reference dye (ROX). dRn is Rn minus the baseline (adjusted automatically by MXPro for qPCR software),and is plotted against PCR cycle number. D) Representative RT-qPCR amplification plots of TDP2 for MM453 and MCF7.
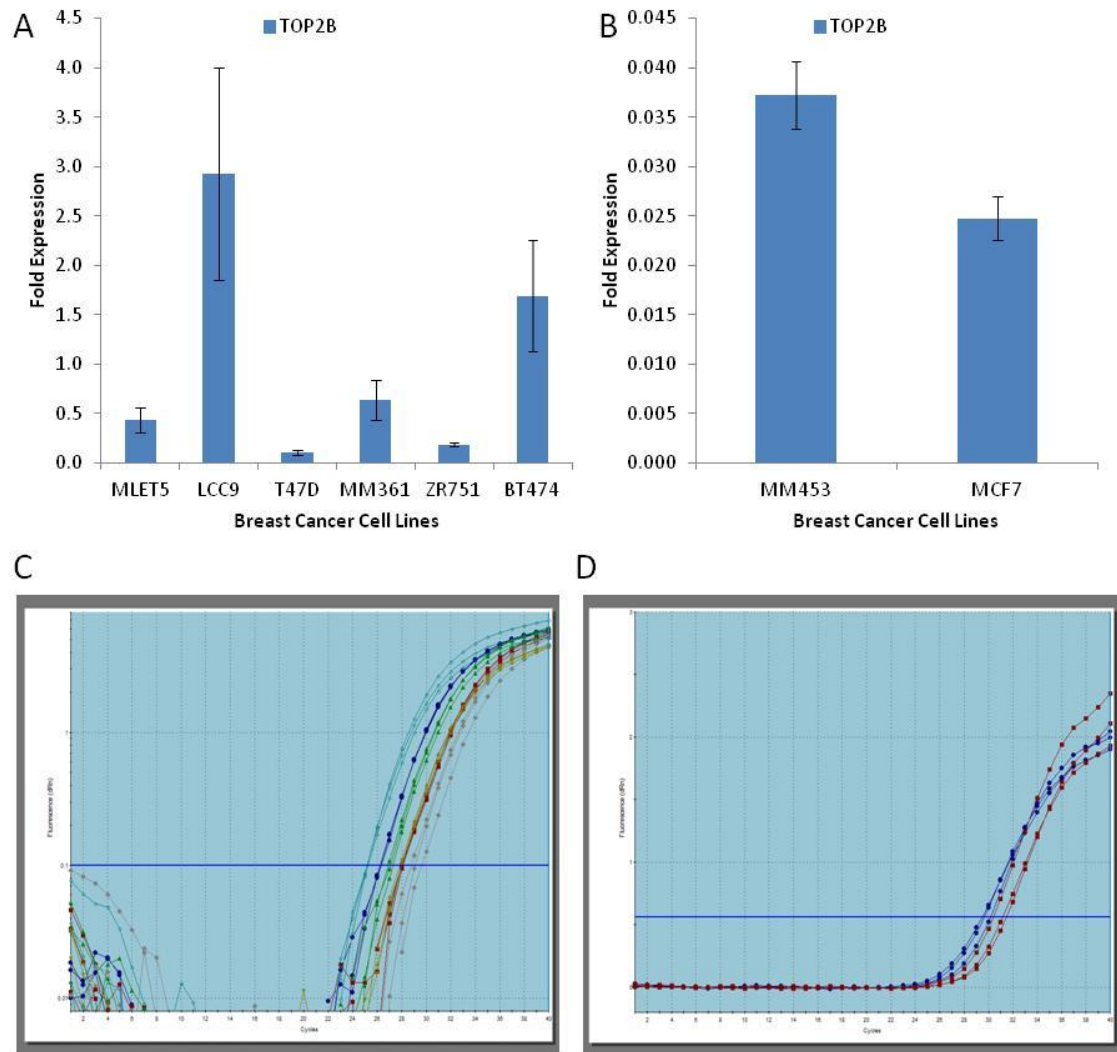
Figure 4.3: TOP2B mRNA levels in breast cancer cell lines. A,B) TOP2B mRNA expression levels were measured using RT-qPCR in the indicated breast cancer cell lines, and are expressed relative to the expression level of the RPLP0 housekeeping gene. Relative expression (fold-change) was calculated from the equation relative expression=$2^{-\Delta Ct}$, where ΔCt is calculated by subtracting the Ct value of RPLP0 from the Ct value of TOP2B. Data are the mean (+/- SEM) from three independent experiments. C) Representative RT-qPCR amplification plots of TOP2B for MLET5, LCC9, T47D, MM361, ZR751 and BT474. Amplification plots represent the accumulation of product over the duration of the RT-qPCR experiment. Rn is the fluorescence of the TOP2B primer divided by the fluorescence of the internal reference dye (ROX). dRn is Rn minus the baseline (adjusted automatically by MXPro for qPCR software),and is plotted against PCR cycle number. D) Representative RT-qPCR amplification plots of TDP2 for MM453 and MCF7.

4.2.2 TDP2, TOP2A and TOP2B protein levels vary significantly in breast cancer cell lines.

As before, TDP2, TOP2A and TOP2B, protein levels were also measured, because mRNA levels do not necessarily reflect the level of protein expression.

The same panel of breast cancer cell lines as the one for RT-qPCRs was used, but the experiments were done with all the cell lines at the same time, unlike the RT-qPCRs. Whole cell extracts prepared from each of the eight cell lines were analyzed by western blotting using antibodies against TDP2, TOP2A, TOP2B, and Actin. Actin was used as a control, to normalise the TDP2, TOP2A and TOP2B signals for differences in sample loading. Three different cell extract protein concentrations were employed for each sample, enabling me to ensure that quantification of the western blot signal was linearly related to protein quantity and was thus accurate. The first protein concentration for each sample is denoted C1, the second is denoted C2 and was twice that of C1, and the third is denoted C3 and was four-fold that of C1.

There was a good correlation between TDP2 band intensity and amount of extract protein included in the western blots in Figure 4.4 (Fig.4.4E). Whilst the relative level of TDP2 protein detected in the different cell lines differed slightly, depending on the amount of cell extract employed, ZR751 consistently exhibited the lowest levels of TDP2 protein. However, the difference between the lowest and highest expressing cell lines was no more than 2-3-fold, and thus much less than the difference observed with mRNA levels. Moreover, there was very little correlation between the level of TDP2 mRNA and the level of TDP2 protein, in this panel of cell lines.

TOP2A immunoblots revealed that MLET5, ZR751 and BT474 possessed the highest levels of TOP2A protein (Fig.4.5). However, there was considerable inconsistency in the TOP2A blots, as indicated by the lack of similarity in relative expression levels across the cell lines at different protein extract concentrations. The results for TOP2B were more consistent, with LCC9

possessing highest TOP2B levels at all cell extract concentrations employed, and up to ~8-fold more than the lowest expressing cell lines, MM361 and MM453 (Fig.4.6).
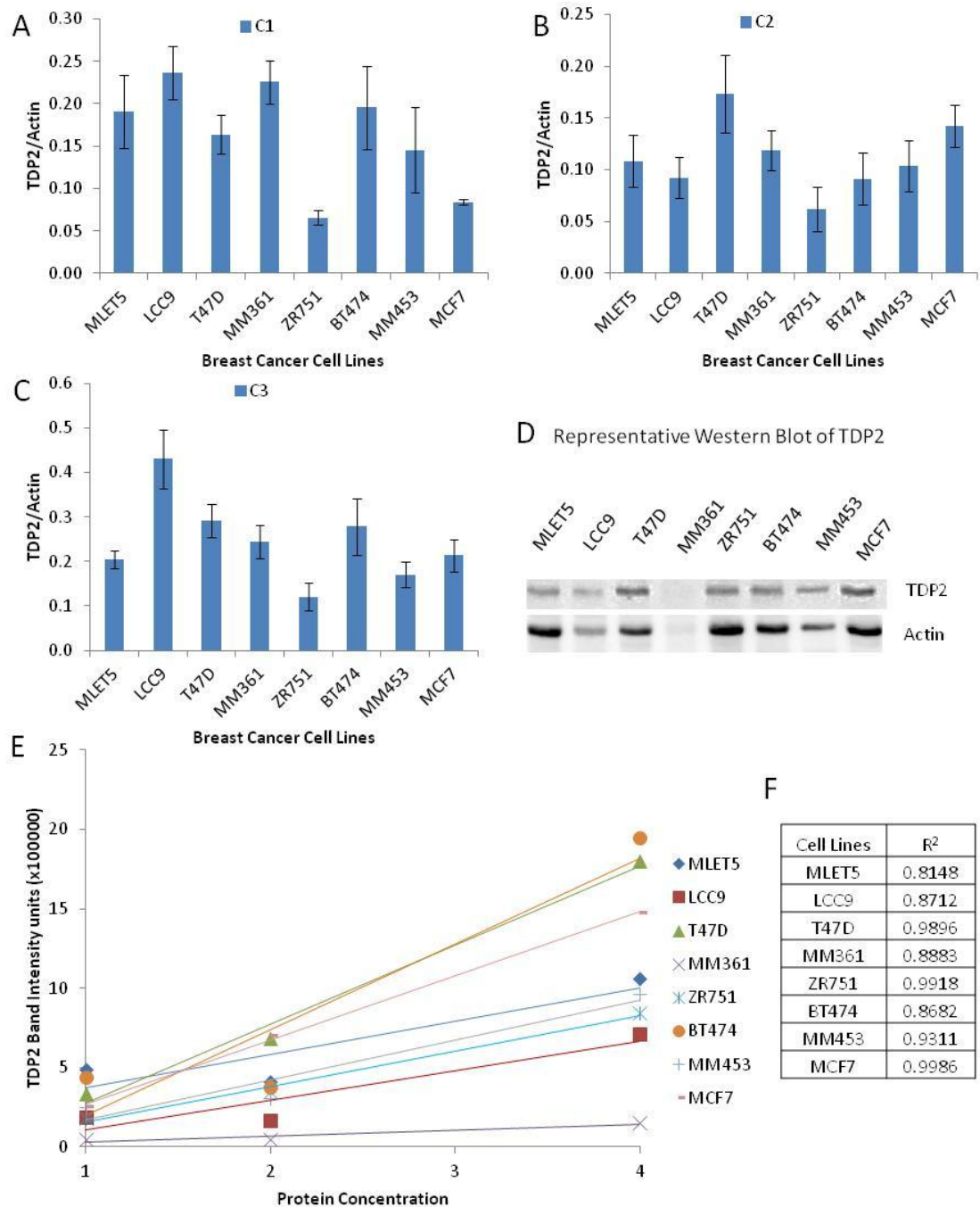
Figure 4.4: TDP2 protein levels in breast cancer cell lines. Panels A-C; Different amounts of cell extract protein from the indicated cell lines were fractionated by SDS-PAGE and subjected to Western blotting with anti-TDP2 and anti-Actin antibodies. TDP2 band intensities were quantified by densitometry and normalised for differences in loading by dividing by the band intensity for Actin. Data in Panels A-C are the mean of three independent experiments (+/-SEM). Gels employed in Panel B and Panel C contained 2x and 4x the amount of cell extract protein as that employed for the gels in Panel A. Panel D; Representative Western blots for TDP2 and Actin. Panel E; Plots of the TDP2 band intensities from Panels A-C. Panel F; R^2 values of the plots in panel E, demonstrating the linear relationship between TDP2 band intensity and amount of protein extract employed in panels A-C.

Figure 4.5: TOP2A protein levels in breast cancer cell lines. Panels A-C; Different amounts of cell extract protein from the indicated cell lines were fractionated by SDS-PAGE and subjected to Western blotting with anti-TOP2A and anti-Actin antibodies. TOP2A band intensities were quantified by densitometry and normalised for differences in loading by dividing by the band intensity for Actin. Data in Panels A-C are the mean of three independent experiments (+/- SEM). Gels employed in Panel B and Panel C contained 2x and 4x the amount of cell extract protein as that employed for the gels in Panel A. Panel D; Representative Western blots for TOP2A and Actin. Panel E; Plots of the TOP2A band intensities from Panels A-C. Panel F; R^2 values of the plots in panel E, demonstrating the linear relationship between TOP2A band intensity and amount of protein extract employed in panels A-C.

Figure 4.6: TOP2B protein levels in breast cancer cell lines. Panels A-C; Different amounts of cell extract protein from the indicated cell lines were fractionated by SDS-PAGE and subjected to Western blotting with anti-TOP2B and anti-Actin antibodies. TOP2B band intensities were quantified by densitometry and normalised for differences in loading by dividing by the band intensity for Actin. Data in Panels A-C are the mean of three independent experiments (+/-SEM). Gels employed in Panel B and Panel C contained 2x and 4x the amount of cell extract protein as that employed for the gels in Panel A. Panel D; Representative Western blots for TOP2B and Actin. Panel E; Plots of the TOP2B band intensities from Panels A-C. Panel F; R^2 values of the plots in panel E, demonstrating the linear relationship between TOP2B band intensity and amount of protein extract employed in panels A-C.

4.3 Discussion

The RT-qPCR and western blot data in this chapter demonstrate a difference in mRNA and protein levels in TDP2, TOP2A and TOP2B among the panel of breast cancer cell lines used. As with the lung cancer cell lines, the results measured with RT-qPCR and western blot, i.e. mRNA levels and protein levels are not the same, since mRNA levels do not correlate directly with protein levels of any gene. There is an abundance of post-transcriptional and post-translational mechanisms that have the ability to affect the amount of protein produced, as well as its half-life (Vogel and Marcotte, 2012, Stark et al., 2006, Dickson et al., 2007, Taquet et al., 2009). But as with the lung cancer cell lines, the information is still valuable, although for future work, it was decided that the project would focus on the protein levels instead of the mRNA levels, as the proteins are the active cellular components that take part in enzymatic functions in the cell.

As with the lung cancer cell lines, a variation in protein levels among the three different concentrations of protein used in TDP2, TOP2A and TOP2B can also be observed. As previously, the most probable explanation for this inconsistency can be attributed to the different potency of the antibodies used. Once again, one can observe that the TDP2 antibody has highest affinity in detecting the protein, as the relative protein levels remain consistent among the three different concentrations used. TOP2A antibody appears to be the most inconsistent, and more specifically for the second concentration, compared to the other two. Given the large error bars in the graphs of the first and third concentration for MLET5, this could also be attributed to an error during protein transfer or antibody incubation. The TOP2B antibody seems to have worked better than the TOP2A antibody, but nevertheless still demonstrates a variation in protein levels within the breast cancer cell line panel. Furthermore, even though measures were taken to maximize the accuracy of the samples loaded, there are instances where Actin levels show problematic loading. This could perhaps be attributed to the vastly different sizes of the cells among the cells lines. For example, MM361 cells are very small

compared to the majority of the other cell lines, which could explain why its Actin levels are usually significantly lower than the rest. Actin, being a major component of the cytoskeleton of a cell, could appear in fewer quantities in a smaller cell, leading to smaller and fainter Actin bands, even though the same number of cells was used in each loaded sample. These reasons demonstrate the importance of the use of three different concentrations, instead of just one, as this would allow for more accurate interpretation of the results obtained.

Combining the results obtained from the lung cancer cell lines and the ones in this chapter (obtained from the breast cancer cell lines), re-enforces the idea that TDP2 has the potential of being used as a biomarker of sensitivity to TOP2 targeting agents. As explained in the previous chapter, cell lines with low levels of TDP2 and high levels of TOP2A and TOP2B would be the ideal target for treatment with a TOP2 poison and cell lines with high levels of TDP2 and low levels of TOP2A and TOP2B would be the least desirable candidates. High levels of TOP2A and TOP2B could mean more TOP2-mediated abortive DSBs within a cell and low levels of TDP2 could mean that TDP2 could be overwhelmed by abortive DSBs, leading the cell to apoptosis. An abundance of TDP2 (high protein levels), on the other hand, would be able to resolve the fewer TOP2-mediated DSBs caused by low levels of TOP2A and TOP2B.

Among the breast cancer cell line panel, there doesn't appear to be many cell lines that can readily be described as ideal candidates for TOP2 poisons. One exception is ZR751, which is the cell line with the lowest levels of TDP2 and also one of the cell lines with the highest levels of TOP2A. On the other hand, other cell lines with high levels of TOP2A or TOP2B, such as MLET5, BT474 and LCC9, also demonstrate medium to high levels of TDP2. Finally, MM361 and MM453 could perhaps be considered as the baseline, as they demonstrated the lowest levels of TOP2B and their TDP2 levels are around medium. Table 4.1 summarizes the protein levels for TDP2, TOP2A and TOP2B in the cell lines of the breast cancer panel. With this in mind, I

proceeded to explore a possible correlation of TDP2 levels and sensitivity to the TOP2 poison

etoposide, in both lung and breast cancer cells.

Table 4.1: Rank order of TDP2, TOP2A and TOP2B protein levels in breast cancer cell line panel. 8 corresponds to the highest protein levels and 1 corresponds to the lowest levels.

	TDP2	TOP2A	TOP2B
MLET5	4	8	7
LCC9	8	4	8
T46D	3	2	3
MM361	6	1	2
ZR751	1	7	5
BT474	5	6	6
MM453	2	3	1
MCF7	7	5	4

5. CHAPTER FIVE - Relationship between TDP2 protein levels and sensitivity to TOP2 poison etoposide in lung and breast cancer cells

5.1 Introduction

Etoposide is a cytotoxic chemotherapy drug used to treat a variety of cancers. It is classified as a topoisomerase inhibitor (Hande, 1998). It is not uncommon for patients to develop resistance to etoposide and there is data that shows a link between that resistance and TDP2 (Zeng et al., 2011, Lage et al., 2000). I hypothesized that high TDP2 protein levels would translate to high resistance to etoposide, since more TDP2 might allow the DNA double-strand breaks to be repaired. In the previous two chapters, it was established that TDP2 mRNA and protein levels differed among two panels of lung and breast cancer cell lines. This in turn raises the question of whether there is a correlation between TDP2 protein levels, and possibly TOP2 levels, with resistance to the TOP2 poison etoposide.

If such a correlation were to be established, then TDP2 could be used as a predictive biomarker to etoposide resistance. Drug resistance is a severe problem in cancer treatment and due to common resistance mechanisms, cancer cells can acquire resistance to multiple structurally and functionally unrelated drugs, with enhanced DNA repair being one of those mechanisms (Alpsoy et al., 2014). As etoposide is a widely used anti-cancer agent, it would be very useful to establish a possible biomarker that could predict how responsive a patient is going to be to this treatment and how likely they are to develop resistance to etoposide.

5.1.1 Aims of this chapter

The aim of this chapter is to investigate the possible existence of a link between TDP2 protein levels in lung and breast cancer cell line panels and cellular sensitivity/resistance to etoposide.

5.2 Results

The effects of TDP2 expression levels on sensitivity to TOP2 poison etoposide were assessed using both cell viability assays and clonogenic cell survival assays.

For the cell viability assays I employed Alamar Blue, which is a resazurin based assay. Resazurin is a weakly fluorescent blue dye that is converted to resorufin, a highly red fluorescent dye in the presence of metabolically active cells. Alamar Blue was preferred to MTT (Microculture Tetrazolium Assay) for these experiments, because of its ease of use and reduced toxicity to cells. This allowed for more flexibility in experimental design. Where possible, for adherent cells that can grow in colonies, I also employed clonogenic survival assays, which are less susceptible to the short term impact of cytotoxins on cell cycle arrest and growth rate.

In Figure 5.1, the results of the Alamar Blue assays for the breast cancer cell line panel are depicted. The cells were seeded at plating concentrations established from Alamar Blue growth curves. Increasing concentrations of etoposide ranging from 0-30μM were employed and fluorescence measurements were taken at days 1,3,6,8 and 10 post-treatment. The mean fluorescence units from three experiments can be seen for each cell line. BT474, MM361 and MM453 were the most etoposide resistant, as suggested by the limited impact of increasing etoposide concentrations on their growth rate. This is evident from the plot in Figure 5.2A, in which cell doubling times (calculated from the slopes in Fig.5.1) were plotted against etoposide concentration. In Figure 5.2B, IC50 values have been calculated from the trendlines in Figure 5.2A and are shown in a table in descending order. To summarise, in these experiments T47D, LCC9 and ZR751 appear to be the most sensitive cell lines and MM453, BT474 and MM361 are the most resistant.
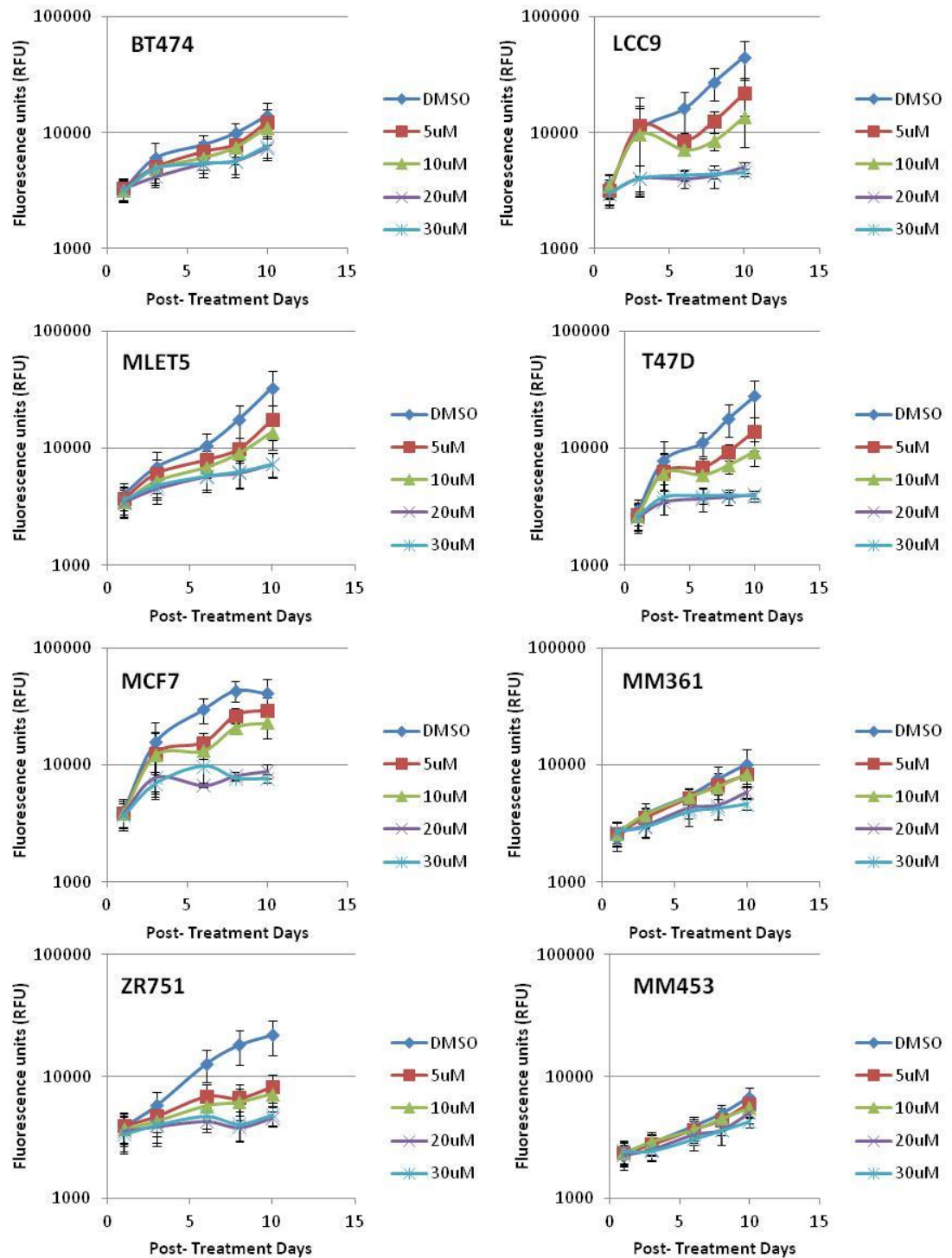
Figure 5.1: Etoposide sensitivity in breast cancer cell lines. Etoposide sensitivity was measured using Alamar Blue assay in the indicated breast cancer cell lines. Etoposide concentrations used were 0µM, 5µM, 10µM, 20µM and 30µM. Cells were treated for 3h and fluorescence measurements were taken at days 1, 3, 6, 8 and 10 post-treatment. These measurements were plotted against the corresponding post-treatment day. Data are the mean (+/- SEM) from three independent experiments.

A

B
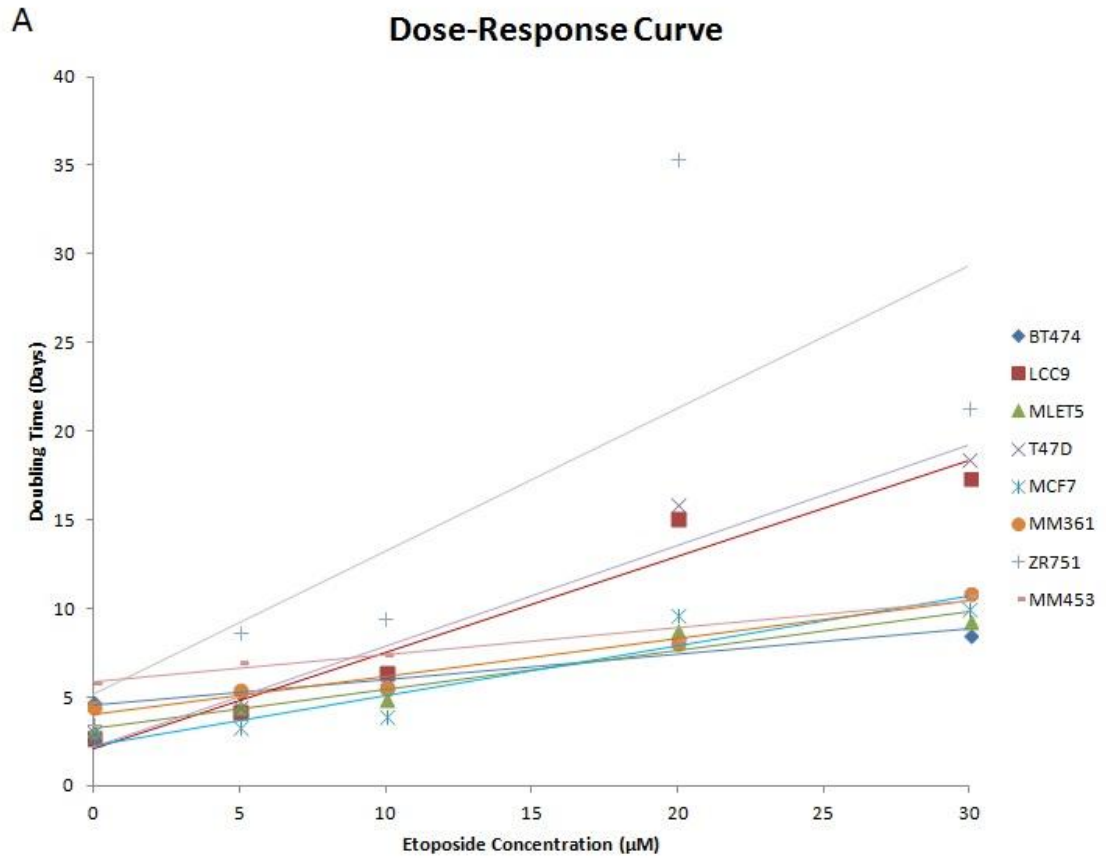
Cell Line	IC50 (µM)
MM453	38.663
BT474	32.397
MM361	19.116
MLET5	14.775
MCF7	8.235
ZR751	6.394
LCC9	3.930
T47D	3.735

Figure 5.2: Dose-Response curve for etoposide sensitivity in breast cancer cell lines A) Cell doubling times (calculated from the slopes in Fig.5.1) were plotted against etoposide concentration. B) IC50 values calculated from the trendlines in Fig.5.2 and plotted in descending order.

To further establish the Alamar Blue results for the breast cancer cell line panel, an additional experiment was designed in order to explore the sensitivity of this assay. The protocol was similar to what was described above, with the only difference being that the cells were not treated acutely on the day before fluorescent measurements started being taken, but they were subjected to a continuous treatment of lower concentrations of etoposide. This was done in order to determine if Alamar Blue is sensitive enough to detect such small differences and to establish whether it would be more suitable as a treatment option for the breast cancer cell line panel. Only two cell lines were picked for this experiment, based on the western blot results for the breast cancer cell line panel in Chapter 4. LCC9 had one of the highest levels of TDP2 and ZR751 had one of the lowest levels of TDP2. These two extremes were picked to maximise the chance of detecting a difference. Figure 5.3 depicts these results. LCC9 appears to demonstrate very similar resistance to increasing concentrations of etoposide, compared to ZR751 (Fig.5.3A). This can be verified by looking at the dose-dependent inhibition of growth graph in Figure 5.3B. The slope values for both of these cell lines are very close, with LCC9 demonstrating a slightly lower resistance to increasing etoposide concentrations, compared to ZR751. These results demonstrate that Alamar Blue is not sensitive enough to detect delicate differences that derive from continuous treatment with etoposide and therefore the acute treatment previously used, was determined as the most suitable.
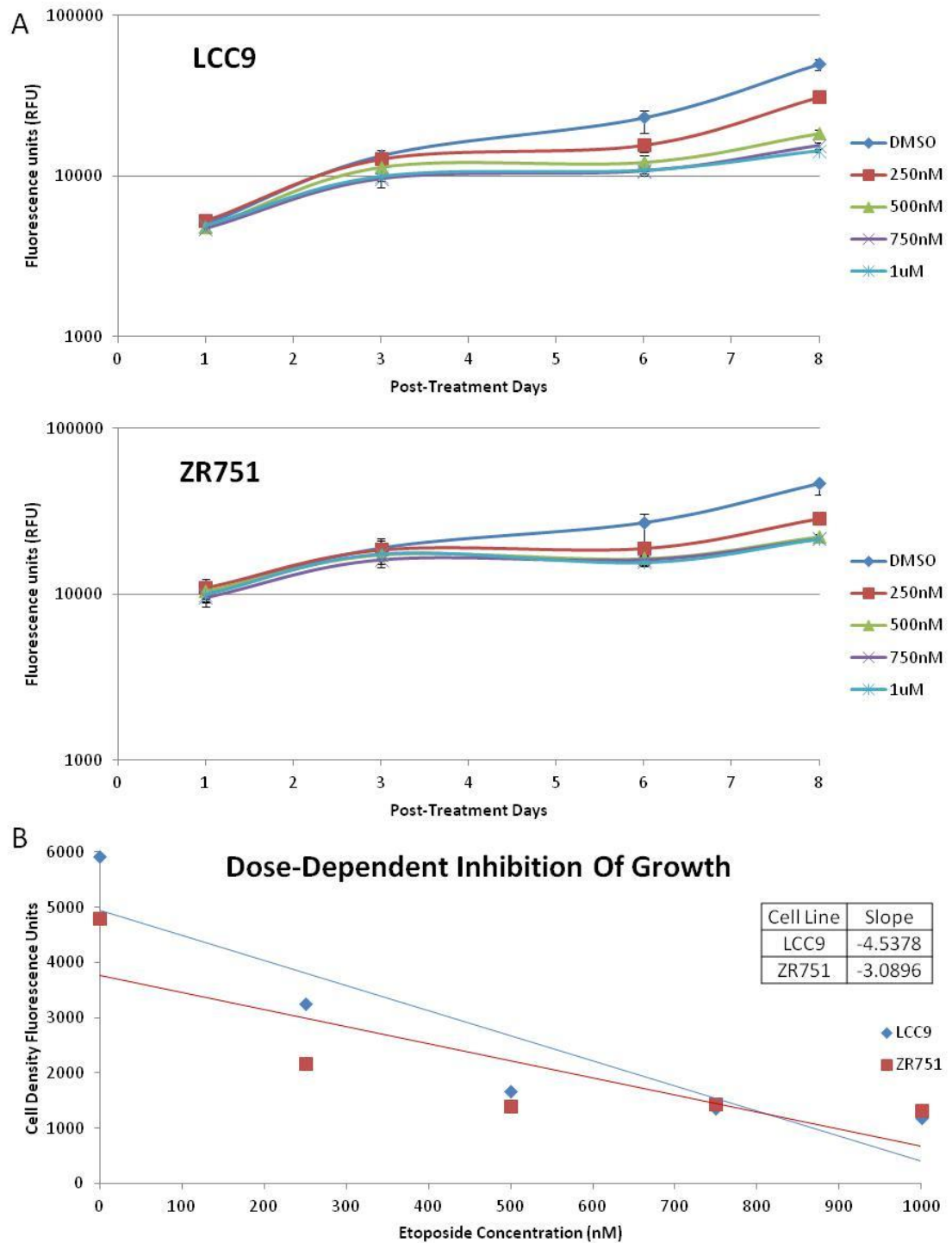
Figure 5.3: Etoposide sensitivity in breast cancer cell lines LCC9 and ZR751. Etoposide sensitivity was measured using Alamar Blue assay in the indicated breast cancer cell lines. Etoposide concentrations used were 0nM, 250nM, 500nM, 750nM and 1µM. Cells were treated continuously and fluorescence measurements were taken at days 1, 3, 6 and 8 after treatment was initiated. These measurements were plotted against the corresponding day. Data are the mean (+/- SEM) from three independent experiments.

Because the lung cancer cell lines grew efficiently in colonies as adherent cell lines, I was able to employ clonogenic cell survival assays to examine their sensitivity to etoposide. The experiments were carried out in two experimental sets, with A549 cell line included in both groups as an internal control for experimental consistency between the two sets. The first group consisted of A549, H1650, HCC95, Hop92 and MOR and the second group consisted of A549, H292, H226 and H460. The results of these experiments are shown in Figures 5.4A and 5.5A, respectively. In both groups, A549 was the cell line displaying greatest resistance to etoposide. For the first group, HCC95 was the cell line displaying greatest sensitivity. For the second group, H460 was the cell line least resistant to etoposide. To compare the relative sensitivity of the cell lines, I calculated D37 values (Fig.5.4B and 5.5B). D37 values are defined as the concentration of drug at which cell survival is 37% of the initial cell population. The relative sensitivity of the lung cancer cell lines in order of increasing sensitivity for the first group was A549, MOR, Hop92, H1650 and HCC95. For the second group, the order was A549, H292, H226 and H460, from least to most sensitive to etoposide.

A

Sensitivity to Etoposide

B
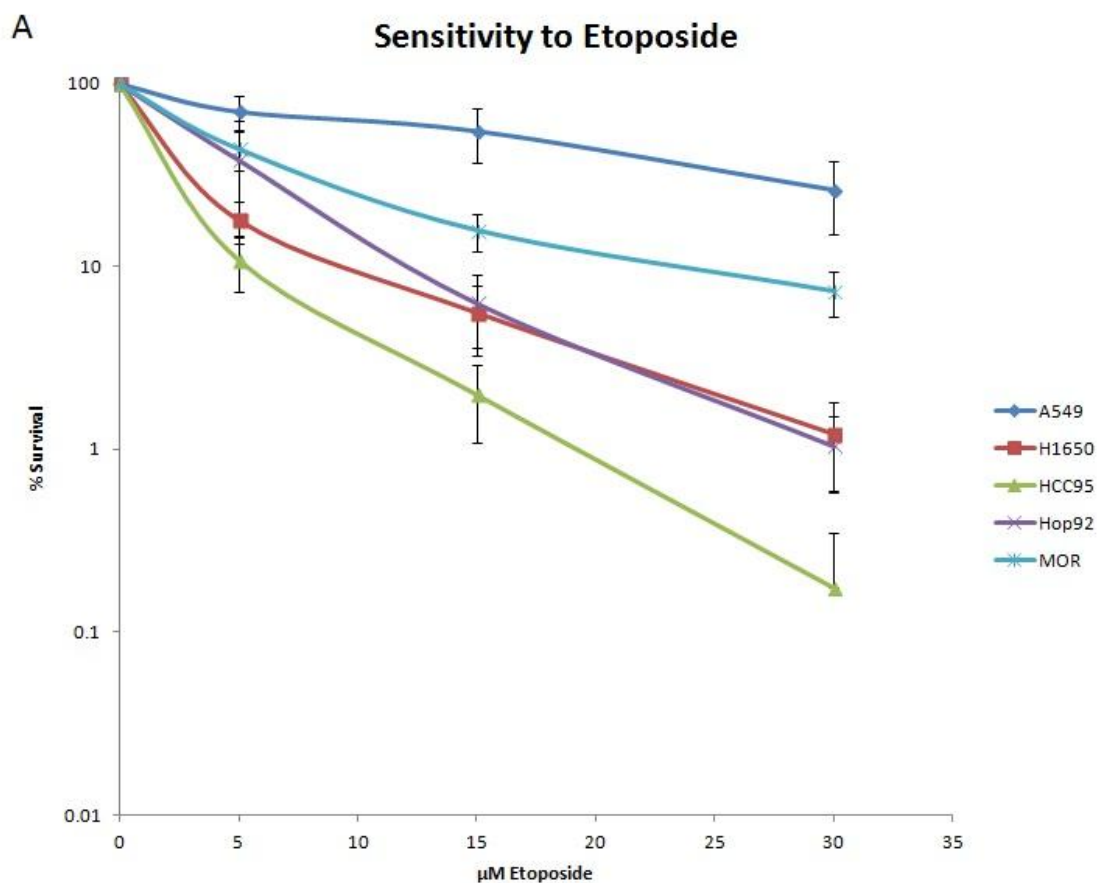
Cell Line	D37 (µM)
A549	24.04
MOR	14.28
Hop92	12.31
H1650	10.24
HCC95	9.04

Figure 5.4: Etoposide sensitivity in lung cancer cell lines A549, MOR, Hop92, H1650 and HCC95. A) Etoposide sensitivity was measured using Clonogenic Cell Survival assay in the indicated lung cancer cell lines. Etoposide concentrations used were 0µM, 5µM, 15µM and 30µM. Cells were treated for 3h and left to grow in colonies for two weeks. The colonies were counted and the percentage of survival for each concentration was calculated from the untreated control. The percentages of survival were plotted against the corresponding etoposide concentrations. Data are the mean (+/- SEM) from three independent experiments. B) D37 values calculated from the trendlines of Fig.5.4A and are depicted in descending order.

A

Sensitivity to Etoposide

B
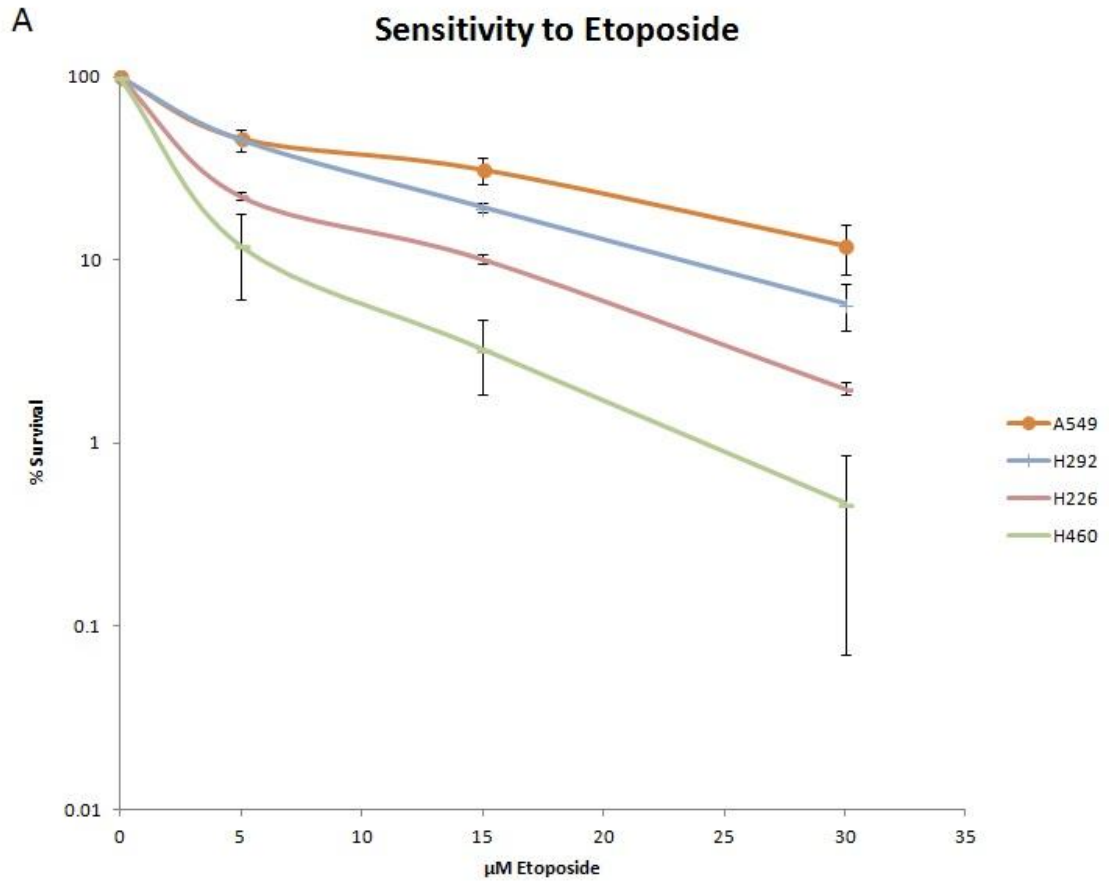
Cell Line	D37 (µM)
A549	16.62
H292	14.57
H226	11.19
H460	9.29

Figure 5.5: Etoposide sensitivity in lung cancer cell lines A549, H292, H226 and H460. A) Etoposide sensitivity was measured using Clonogenic Cell Survival assay in the indicated lung cancer cell lines. Etoposide concentrations used were 0µM, 5µM, 15µM and 30µM. Cells were treated for 3h and left to grow in colonies for two weeks. The colonies were counted and the percentage of survival for each concentration was calculated from the untreated control. The percentages of survival were plotted against the corresponding etoposide concentrations. Data are the mean (+/- SEM) from three independent experiments. B) D37 values calculated from the trendlines of Fig.5.5A and are depicted in descending order.

In order to demonstrate that Alamar Blue assay does not provide as clear results as clonogenic cell survival assay, similar experiments to Figure 5.1 were conducted for the three lung cancer cell lines MOR, HCC95 and H292 (Fig.5.6). This was done in order to further explain why it is more challenging to interpret results derived from it. These three cell lines were picked based on the preliminary western blot results for the lung cancer cell lines and they represent cell lines with low (H292) and high (HCC95 and MOR) levels of TDP2. The experiments were done in pairs, MOR with HCC95 and MOR with H292. The conditions were the same. This series of experiments also show different levels of etoposide resistance. For the first pair, MOR and HCC95, one can observe that they both have relatively high resistance to etoposide, but it is very difficult to establish which one is more sensitive (Fig.5.6A). For the second pair, MOR and H292, it can also be said that H292 shows high resistance to etoposide (Fig.5.6A). In Figure 5.6B the dose-dependent inhibition of growth is shown, which allows the reader to derive more accurate conclusions. As the slope for HCC95 is of higher value than the one for MOR, it means that HCC95 is less sensitive to etoposide than MOR (Graph on the left). H292 also has a slope of higher value than MOR, making H292 less sensitive than MOR (Graph on the right), which was not expected. In conclusion, it is clear that Alamar Blue is a useful technique for cell lines that cannot grow in colonies, but it has proven limitations.
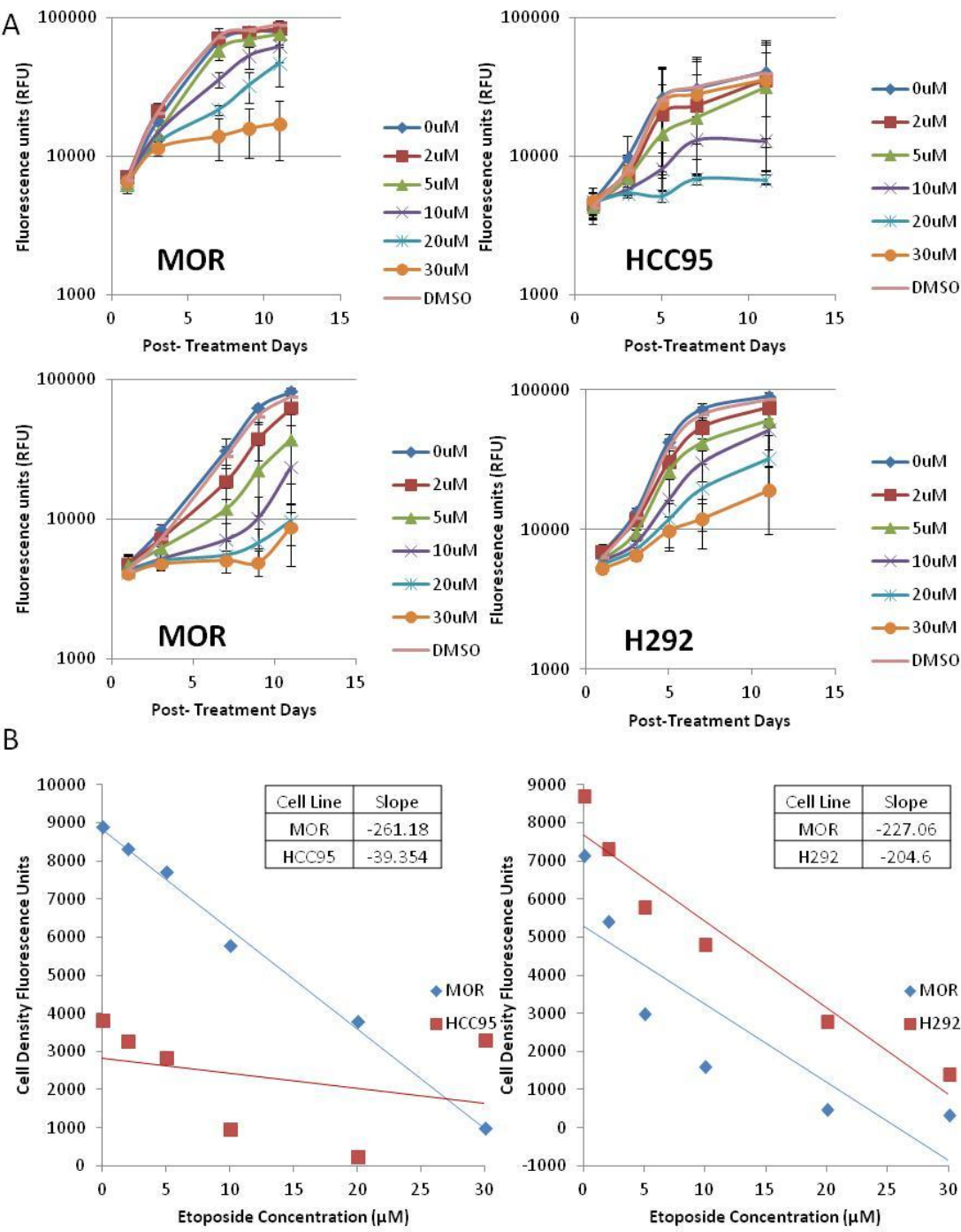
Figure 5.6: Etoposide sensitivity in lung cancer cell lines MOR, HCC95 and H292. A) Etoposide sensitivity was measured using Alamar Blue assay in the indicated lung cancer cell lines. Etoposide concentrations used were 0μM, 2μM, 5μM, 10μM, 20μM and 30μM. Cells were treated for 3h and fluorescence measurements were taken at days 1, 3, 6, 8 and 11 post-treatment. These measurements were plotted against the corresponding post-treatment day. Data are the mean (+/- SEM) from three independent experiments. B) Dose dependent inhibition of growth. Cell density fluorescence units are plotted against etoposide concentrations used. The slopes of the trendlines are depicted in tables.

In order to be able to understand the results for etoposide sensitivity in lung and breast cancer cells, the protein levels of TDP2, TOP2A and TOP2B were plotted as ratios in Figure 5.7. Figures 5.7A and B depict the TOP2A/TDP2 and TOP2B/TDP2 ratios for the lung cancer cell lines and Figure 5.7C and D depict the TOP2A/TDP2 and TOP2B/TDP2 ratios for the breast cancer cell lines. This allows the reader to visualise which cell lines have very high levels of TOP2A and TOP2B compared to TDP2, such as A549, H226, ZR751 and BT474, and which cell lines demonstrate the opposite, such as H460, HCC95, MM361 and MM453.

In Figure 5.8, one can observe two scatter plots. Figure 5.8A depicts the D37 values for the lung cancer cell lines from Figure 5.4, plotted against the TDP2/Actin protein ratio from Chapter 3. Figure 5.8B depicts the IC50 for the breast cancer cell lines from Figure 5.2, plotted against the TDP2/Actin protein ratio from Chapter 4. It is quite clear from the spread of the data points for both lung and breast cancer cell lines, that there is no correlation between etoposide resistance and TDP2 protein levels. Figures 5.9 and 5.10 depict similar results to Figure 5.8 for the TOP2A/TDP2 and TOP2B/TDP2 protein ratios of Figure 5.7, in lung and breast cancer cell lines respectively. The same D37 and IC50 values were used as in Figure 5.8. From these scatter plots, no relationship is apparent. To verify that there is no correlation between any of the phenotypes and etoposide sensitivity, the Pearson correlation coefficient or r coefficient and the corresponding p-values were calculated. These results are shown in Table 5.1. The p-values for all the r-values are less than 0.05 and are therefore not significant, which verified the conclusion from the scatter plots.
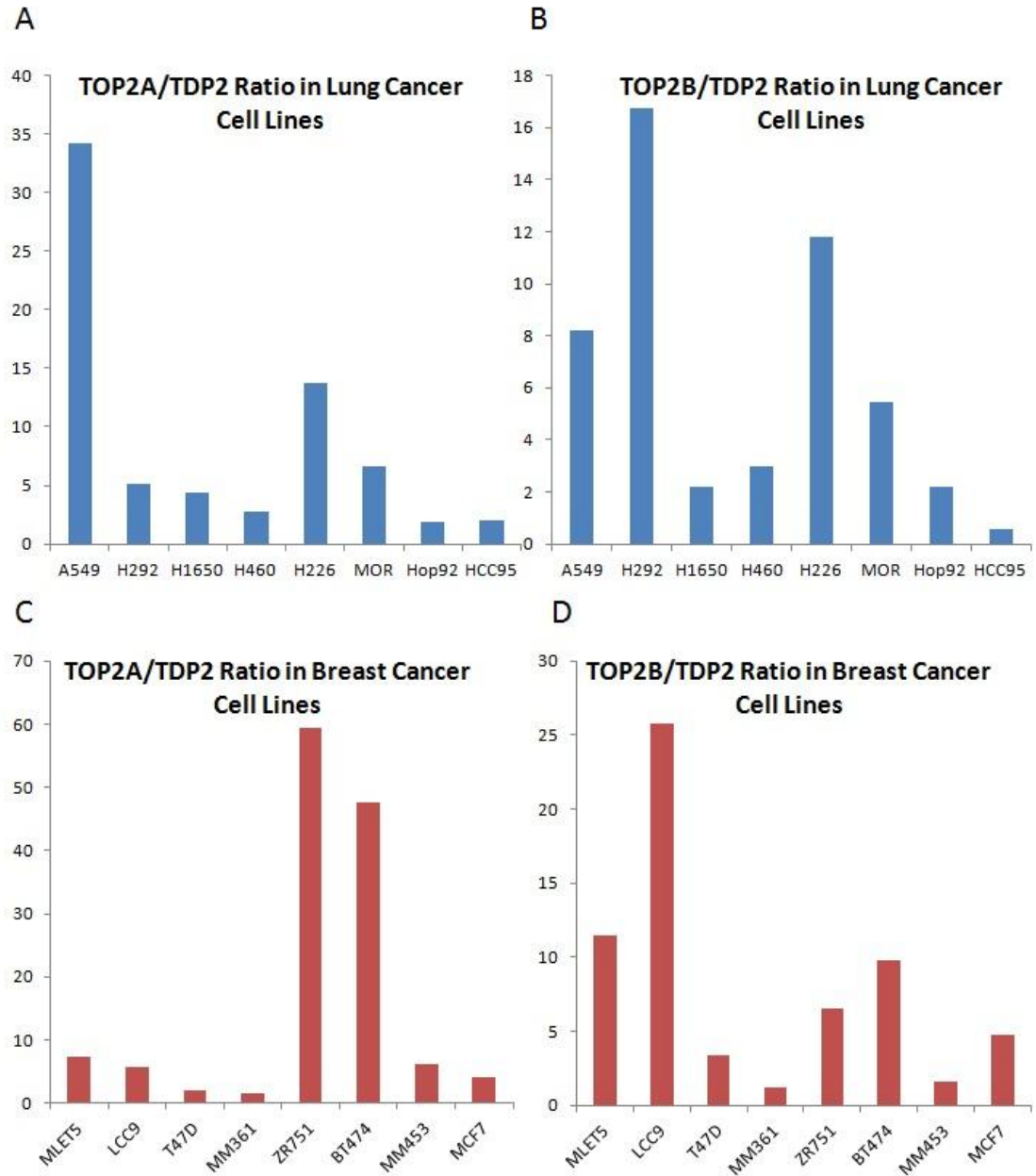
Figure 5.7: TOP2A/TDP2 and TOP2B/TDP2 protein level ratios in lung and breast cancer cell lines. TDP2, TOP2A and TOP2B protein levels were obtained from Fig.3.4, 3.5, 3.6, 4.4, 4.5 and 4.6. The averaged values for concentration C2 were used. A) TOP2A/TDP2 protein level ratio for indicated lung cancer cell lines, B) TOP2B/TDP2 protein level ratio for indicated lung cancer cell lines, C) TOP2A/TDP2 protein level ratio for indicated breast cancer cell lines, D) TOP2B/TDP2 protein level ratio for indicated breast cancer cell lines.
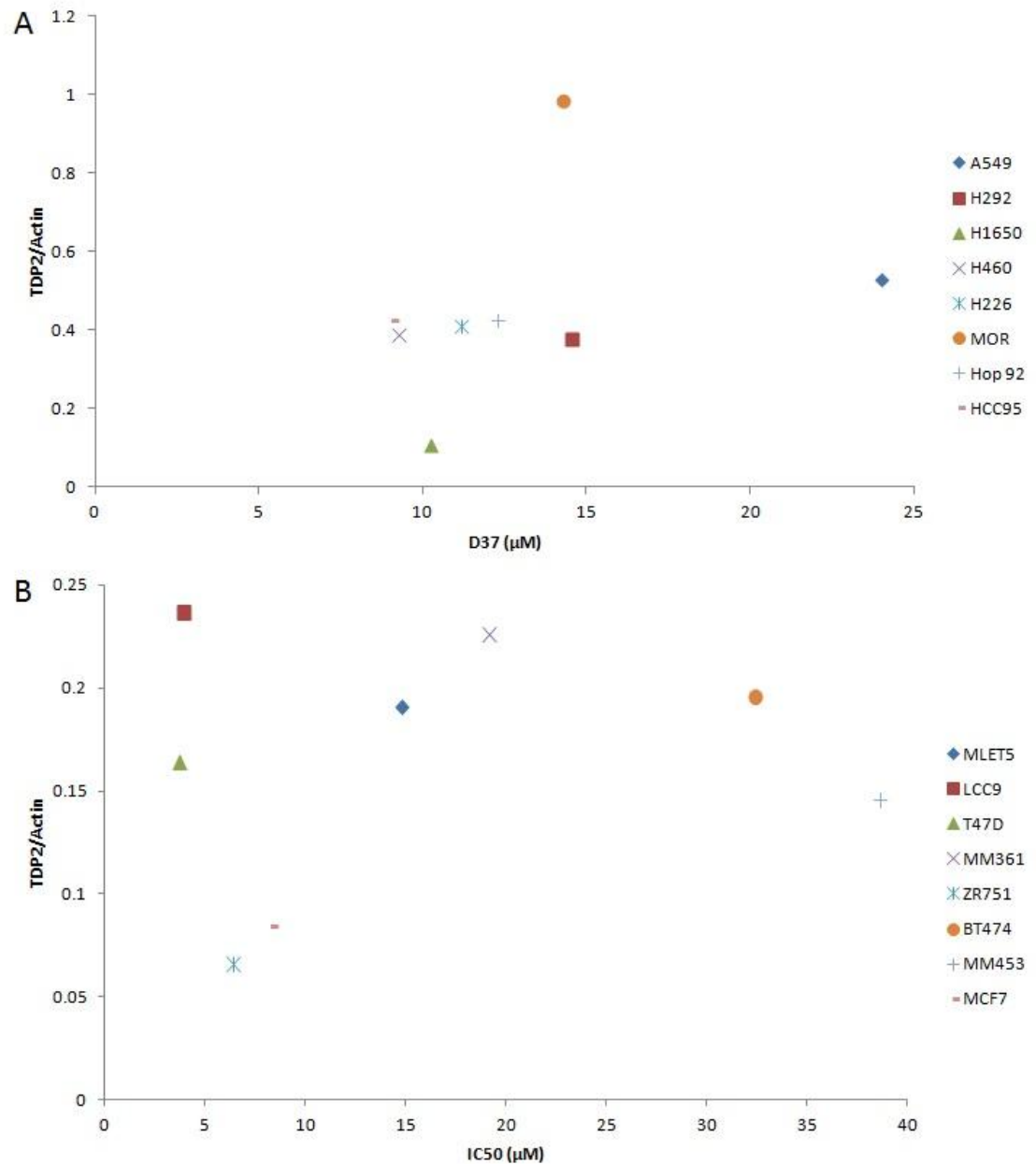
Figure 5.8: Scatter plot of TDP2 protein levels versus D37 or IC50 values of lung and breast cancer cell lines. TDP2 protein levels normalised to Actin were obtained from Figs.3.4 and 4.4. Averaged values of concentration C1 were used. A) Scatter plot for indicated lung cancer cell lines, TDP2/Actin ratio plotted against D37 values from Fig.5.4 and 5.5. B) Scatter plot for indicated breast cancer cell lines, TDP2/Actin ratio plotted against IC50 values from Fig.5.2.
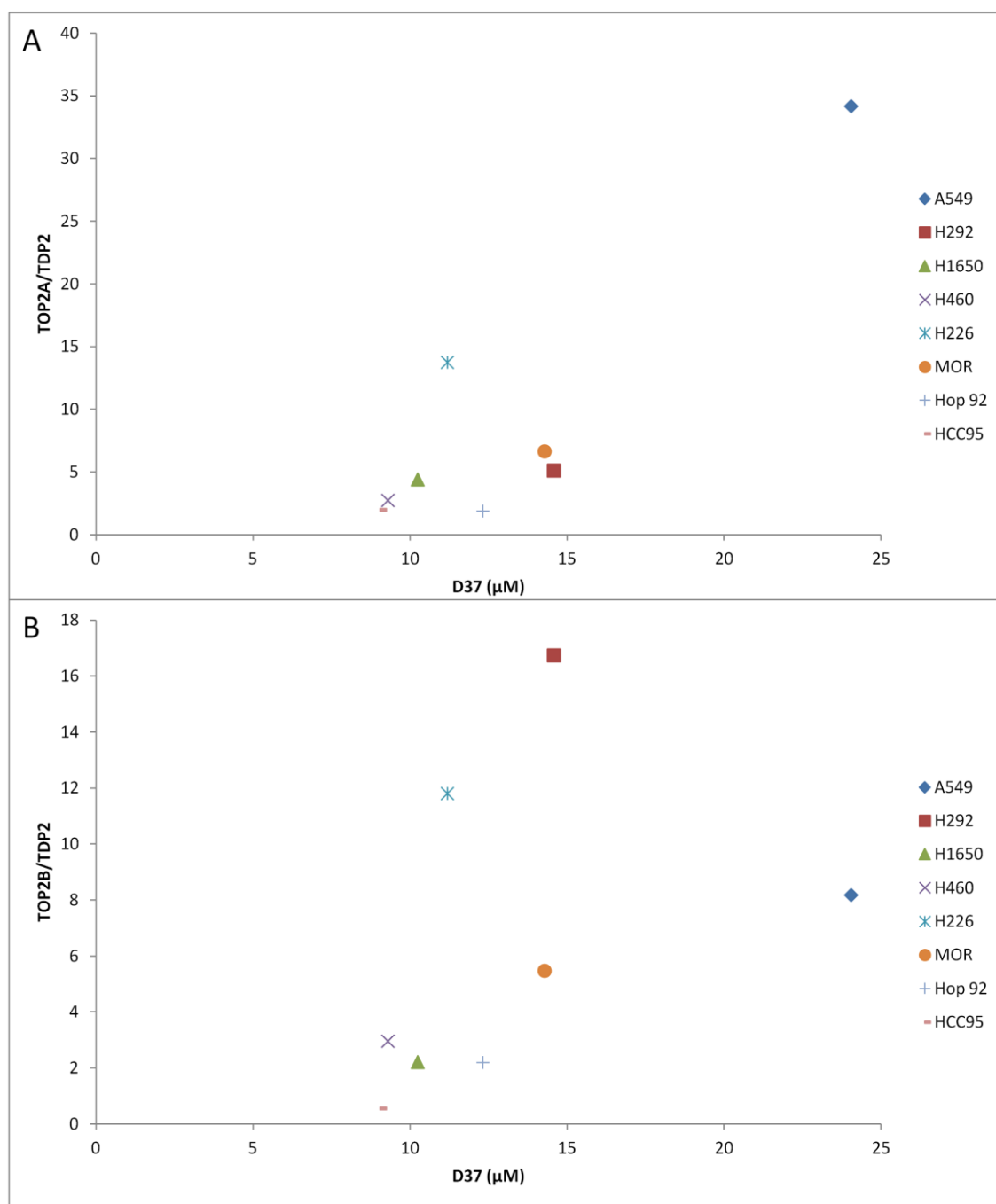
Figure 5.9: Scatter plot of TOP2A/TDP2 and TOP2B/TDP2 protein level ratios versus D37 values of lung cancer cell lines. Protein level ratios were obtained from Fig.5.7 and D37 values from Fig.5.4 and 5.5. Averaged values of concentration C1 were used. A) Scatter plot of TOP2A/TDP2 ratio for indicated lung cancer cell lines. B) Scatter plot of TOP2B/TDP2 ratio for indicated lung cancer cell lines.
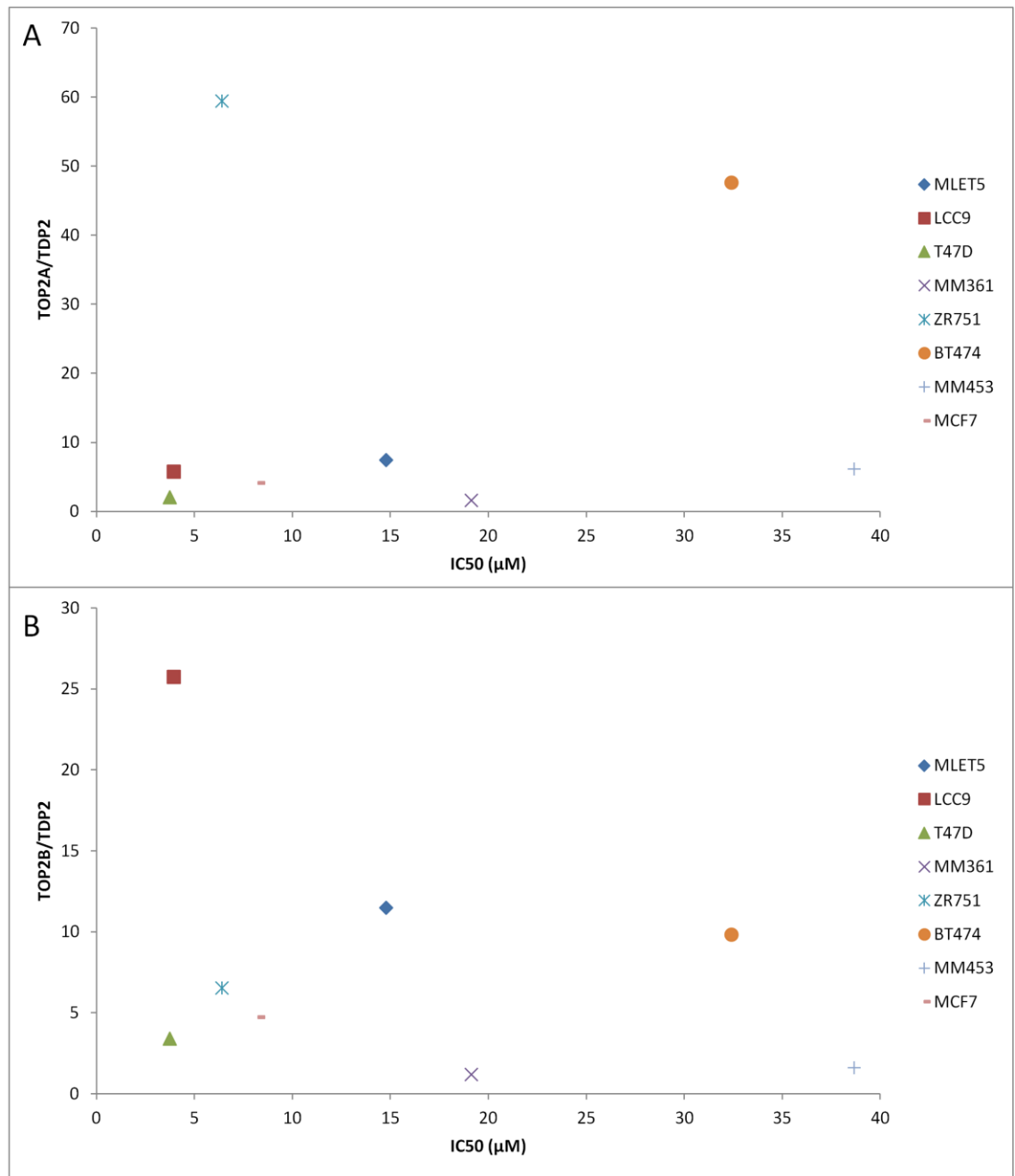
Figure 5.10: Scatter plot of TOP2A/TDP2 and TOP2B/TDP2 protein level ratios versus IC50 values of breast cancer cell lines. Protein level ratios were obtained from Fig.5.7 and IC50 values from Fig.5.2. Averaged values of concentration C1 were used. A) Scatter plot of TOP2A/TDP2 ratio for indicated breast cancer cell lines. B) Scatter plot of TOP2B/TDP2 ratio for indicated breast cancer cell lines.

Table 5.1: Statistical analysis results of TDP2 protein levels, TOP2A/TDP2 protein ratios and TOP2B/TDP2 protein ratios and their correlation with sensitivity to etoposide in lung and breast cancer cell lines (D37 or IC50 values, respectively).

	Phenotype	r-value	p -value
Lung	TDP2	0.33296	0.42030
	TOP2A/TDP2	0.58616	0.12675
	TOP2B/TDP2	0.14642	0.72936
Breast	TDP2	0.16212	0.70130
	TOP2A/TDP2	-0.21519	0.60881
	TOP2B/TDP2	-0.49259	0.21492

5.3 Discussion

A comparison of TDP2 protein levels and cell survival did not yield a positive correlation between TDP2 protein levels and sensitivity to etoposide, in neither breast nor lung cancer cell lines. It appears that protein levels do not necessarily translate to etoposide sensitivity or resistance. In addition, there seems to be no correlation between sensitivity to etoposide and TOP2A/TDP2 and TOP2B/TDP2 protein ratios either. For the lung cancer cell lines, MOR and HCC95 possess the highest levels of TDP2 protein, followed by A549, H292, H460, H226 and Hop92 with intermediate levels, and H1650 with the lowest levels (Chapter 3). On the other hand A549 was the cell line least sensitive to etoposide, followed by H292, MOR, Hop92, H226, H1650, H460 and HCC95, based on the D37 values in Figures 5.4B and 5.5B. This is not expected, as it was anticipated that cell lines with high levels of TDP2 would be resistant to etoposide and cell lines with low levels of TDP2 would be sensitive. For example, HCC95 demonstrated very high levels of TDP2 but very low levels of etoposide resistance. This cannot be explained simply by just looking at the levels of TOP2A and TOP2B, since in Figure 5.7A the ratio of TOP2A/TDP2 and TOP2B/TDP2 is very low indicating that TOP2A and TOP2B levels are very low compared to TDP2. A549 demonstrated intermediate levels of TDP2 (Chapter 3) and, while having a high ratio of TOP2A/TDP2 and intermediate ratio of TOP2B/TDP2, it demonstrated the highest resistance to etoposide, when it was expected that it would be intermediate. To further observe the lack of correlation one can look at Figure 5.8A, in which a

scatter plot depicts the D37 values of the lung cancer cell lines plotted against TDP2 protein levels normalised to Actin. It is very clear that the data points form a cloud, rather than falling on a best fit line. The same can be said for TOP2A/TDP2 and TOP2B/TDP2 protein level ratios.

A similar picture is seen for the breast cancer cell lines, as comparing the Alamar Blue results with the protein levels for each cell line, does not demonstrate a correlation between the number of viable cells in culture after etoposide treatment and the TDP2 protein levels observed previously. Although Alamar Blue has been proven to be an assay with limitations, has exhibited lower sensitivity and its results are more difficult to interpret compared to clonogenic cell survival assay, it is still the only kind of assay that can be used for many of the breast cancer cell lines in the panel used. LCC9 and T47D demonstrated the highest levels of TDP2 protein, followed by MLET5, MM361, BT474, MM453 and MCF7 with intermediate levels, and ZR751 consistently had the lowest levels of TDP2 protein (Chapter 4). From the IC50 values of Figure 5.2B, one can see that MM453 and BT474 are the most resistant to etoposide, followed by MM361, MLET5, MCF7, ZR751, LCC9 and T47D. Taking LCC9 as an example, it is not difficult to show the lack of correlation between TDP2 protein levels and sensitivity to etoposide. It is one of the cell lines with the highest levels of TDP2, but demonstrated one of the lowest levels of resistance to etoposide. In addition, its TOP2A/TDP2 ratio is very low, although its TOP2B/TDP2 ratio is relatively high (Fig.5.8B). Another example is T47D, another cell line with very high levels of TDP2 and yet it's the one with the highest sensitivity to etoposide, while having very low ratios of TOP2A/TDP2 and TOP2B/TDP2. Similar to the lung cancer cell lines, this lack of correlation is also easier to visualise in Figure 5.5B, which depicts a scatter plot of the IC50 values of the breast cancer cell lines plotted against TDP2 protein levels normalised to Actin. It is very clear again that the data points form a cloud, rather than falling on a best fit line. Similarly, no correlation can be observed in the scatter plots for TOP2A/TDP2 and TOP2B/TDP2 protein ratios.

I was able to obtain etoposide IC50 values for some of the cell lines I utilised in my experiments from a drug database, CancerDR (for selection and details, please refer to Chapter 2.5.1) and used them as a means of explaining my results. I calculated the r-values and their corresponding p-values, using my experimentally defined TDP2/Actin, TOP2A/TDP2 and TOP2B/TDP2 protein ratios, the results of which can be found in Table 5.2. Agreeing with my results, these p-values are not significant.

Table 5.2: Statistical analysis results of TDP2 protein levels, TOP2A/TDP2 protein ratios and TOP2B/TDP2 protein ratios and their correlation with sensitivity to etoposide (IC50 values), obtained from CancerDR in lung and breast cancer cell lines.

	Phenotype	r-value	p -value
Lung	TDP2	0.42630	0.47414
	TOP2A/TDP2	0.87138	0.05429
	TOP2B/TDP2	0.52240	0.36649
Breast	TDP2	-0.09065	0.88473
	TOP2A/TDP2	0.70098	0.18723
	TOP2B/TDP2	-0.30377	0.61926

All the data that has shown impact on etoposide sensitivity while manipulating TDP2 levels, either by overexpressing it or by inhibiting it, have used a closed system where the parameters were less variable. My system is more complicated with multiple parameters. Looking at intrinsic values of TDP2 on its own or as a protein ratio with other relevant proteins is more challenging, since there are a lot of factors that could be contributing to how the cell responds to etoposide, or to how TDP2 responds to the formed DSB.

A recent study similar to mine investigated whether TDP1 or TOP1 levels alone could be a predictive biomarker of sensitivity to irinotecan, a TOP1 poison, in colorectal cancer (Meisenberg et al., 2015). Their conclusion was similar to my results, that there was no correlation between intrinsic TDP1 or TOP1 levels and irinotecan, a system very similar to the TDP2-TOP2-etoposide system I was investigating.

Very recently, data has emerged that TDP2 is a novel substrate of extracellular signal-regulated kinase 3, an atypical mitogen-activated protein kinase (ERK3) (Bian et al., 2016). ERK3 was shown to phosphorylate TDP2 and promote its phosphodiesterase activity, and as a result up-regulating TDP2-mediated DNA damage response, which leads to the desensitizing of lung cancer cell lines H460 and A549 to TOP2 inhibitor-induced growth inhibition (Bian et al., 2016). Interestingly, the same paper showed that lung cancer cell lines displayed highly differential response to etoposide; H157 lung cell line showed high basal levels of γ-H2AX that were not changed after etoposide treatment with concentrations up to 20 μM, H1395 cells demonstrated undetectable γ-H2AX levels post etoposide treatment even when TDP2 was knocked down. Finally, H1437 etoposide treatment did increase γ-H2AX levels but knockdown of TDP2 did not show an obvious effect (Bian et al., 2016).

A very recent paper suggests that Mre11, which initiates DSBR by HR via a complex with Rad50 and NBs1, and TDP2 operate in separate pathways for the repair of etoposide-induced breaks (Hoa et al., 2016). Moreover, in the same paper it was shown that the heterozygous mutant $MRE11^{+/-}$, but not the heterozygous $TDP2^{+/-}$ cells, were more sensitive to etoposide than wt cells, which could imply that even small presence of TDP2 is enough to respond to etoposide-induced abortive DSBs. This observation could explain why the intrinsic variability of TDP2 protein levels in the cell lines I studied is too subtle to have a significant impact on etoposide resistance. Interestingly, in the same paper it was also demonstrated that $TDP2^{-/-}$ cells showed no significant genome instability unlike $MRE11^{-/H129N}/TDP2^{-/-}$ cells, suggesting a functional redundancy between the nuclease activity of Mre11 and TDP2 in terms of genome stability, which could be interpreted as another reason why no correlation was observed in my experiments

Nevertheless, the lack of correlation could be attributed to the levels of activity of TDP2 or perhaps even mutations that render it inactive. There is a wealth of information available for

characterised mutations in cancer cell lines. It would be very interesting and a great first step in determining if a gene mutation, other than *TDP2*, is responsible for the lack of correlation I observed between resistance to etoposide and TDP2 protein levels or TOP2A/TDP2 and TOP2B/TDP2 protein ratios. Such a study is conducted in the next chapter.

6. CHAPTER SIX – Mutational profiles of breast and lung cancer cell lines as a predictor of relationship between TDP2 and resistance to etoposide

6.1 Introduction

In the previous chapter, I investigated the possibility of a correlation between sensitivity to the TOP2 poison etoposide and TDP2 protein levels, TOP2A/TDP2 protein ratios or TOP2B/TDP2 protein ratios in a panel of lung and breast cancer cell lines. Unfortunately, I was not able to detect a simple linear correlation, where TDP2 levels or the previously mentioned ratios can accurately predict how sensitive a cell line would be to treatment with etoposide. Given that previous studies of etoposide sensitivity and TDP2 levels were manipulations within the same cell line (Cortes Ledesma et al., 2009, Gómez-Herreros et al., 2014, Zeng et al., 2011), one hypothesis for the lack of correlation is that some of the cell lines may have mutations that affect the relationship between TDP2 levels and etoposide sensitivity.

There is a wealth of information available online via several different kinds of database, but it is not easy even for trained researchers to access this information without experience. But through reviewing the literature and consulting people familiar with the field, I selected a number of sources that provided information on mutations found in the cancer cell lines I used (Futreal et al., 2004, Pavlopoulou et al., 2015, Abaan et al., 2013, Yang et al., 2015b). These were Cancer Cell Line Encyclopedia v2.17 (CCLE), IARC TP53 Database R18, Cell Lines Project v78 and COSMIC Database v78 (Barretina et al., 2012, Bouaoun et al., 2016, Forbes et al., 2015).

The Cancer Cell Line Encyclopedia (CCLE) is a collaborative project between the Broad Institute in Cambridge and the Novartis Institutes for Biomedical Research and its Genomics Institute, the Novartis Research Foundation. It is an ongoing project and currently has information on gene expression, chromosomal copy number and massively parallel sequencing data available

for about 1000 cell lines. CCLE also has pharmacological profiles available for 24 cancer drugs across 479 of the total cell lines, a collection that allows identification of genetic, lineage and gene-expression-based predictors of drug sensitivity (Barretina et al., 2012).

The IARC TP53 Database consists of TP53 mutation data from published literature since 1989. The database includes information on TP53 somatic mutations in sporadic cancers, TP53 germline mutations in familial cancers, common TP53 polymorphisms identified in human populations, functional and structural properties of p53 mutant proteins, mouse-models with engineered TP53, experimentally-induced mutations and the most relevant to this chapter, TP53 gene status in human cell-lines (Bouaoun et al., 2016).

The Cell Lines Project (CLP) of the COSMIC database exclusively displays the results of genomic analysis across a large set of common cancer cell lines used in research. At the moment the number of cell lines is 1015 but it is expected to grow to include about 500 additional cell lines. This project is maintained in a parallel but separate system along the COSMIC database (Forbes et al., 2015).

The Catalogue of Somatic Mutations in Cancer (COSMIC) is the largest and most comprehensive resource for researching the impact of somatic mutations in human cancer. High-impact genes are selected from the Cancer Gene Census, which lists hundreds of genes with substantial published evidence in oncology; this is a very high-confidence list based on good-quality publications. All publications mentioning mutations for these genes in human cancer are collected and exhaustively curated before the information is released into a new version of COSMIC (Forbes et al., 2015).

The information available in all these databases is expert-curated. For the COSMIC Database, the primary source is papers in scientific journals, therefore I was able to personally verify that the samples listed were of the same type as the cell lines I had previously used, as multiple instances of the same sample name exist as separate entries, indicating that it was not clear

for the curator whether these samples were the same. In addition, I also looked to verify that the mutated or wild type form of the genes stated in the database matched those stated in the publications. All databases were last accessed in September 2016, to include as many up-to-date entries as possible.

After mining the results, I organised them per cancer type (breast and lung) in two excel spreadsheets. Each spreadsheet listed all the genes for which I was able to find information for at least one cell line of that type, either a mutation or a wild type entry. For the breast cancer cell lines that was a total of 3686 genes and for the lung cancer cell lines 1150 genes. Out of the 8 breast cancer cell lines I studied (LCC9, MCF7, MM361, BT474, MLET5, T47D, MM453 and ZR751), I was able to find information for at least one gene for 6 cell lines: BT474, MCF7, MM361, MM453, T47D and ZR751. Out of the 8 lung cancer cell lines I studied (MOR, HCC95, Hop92, H226, H460, H292, A549 and H1650), I was able to find information for at least one gene for 7 cell lines: A549, H226, H292, H460, H1650, HCC95 and Hop92.

In order to be able to do statistical analyses for these genes, I could only use genes for which I had information about its status from at least four cell lines, but one of those cell lines had to have a different status for that gene, meaning genes with at least one of both the mutated and wild type form. This criterion reduced the number of genes to 7 for the breast cancer cell lines and 8 for the lung cancer cell lines. I used these genes for statistical modelling, in order to determine the existence of an effect of these gene mutations on the relationship between TDP2 protein levels, TOP2A/TDP2 protein ratios or TOP2B/TDP2 protein ratios and sensitivity to etoposide, and if one exists, how severe it is. For a detailed description of the data analysis, please refer to Chapter 2.5. Focusing my analysis at the level of genes allowed me to identify similarities between my cell lines rather than limiting myself by specific mutations which may not be known and whose specific effect is not understood.

6.1.1 Aims of this chapter

This chapter aims to identify genes that may affect the correlation between previously experimentally defined etoposide IC50 or D37 concentrations and TDP2 protein levels, TOP2A/TDP2 protein ratios and TOP2B/TDP2 protein ratios in a panel of lung and breast cancer cell lines, thus obscuring the correlation between the two when gene status isn't taken into account.

6.2 Results

Before exploring the possibility of an effect of TDP2 protein levels, TOP2A/TDP2 and TOP2B/TDP2 protein ratios, the possible direct effect of these gene mutations on average etoposide sensitivity (in the form of the etoposide IC50 for breast cell lines and D37 values for lung cell lines values shown in Chapter 5) in mutated cell lines versus wild type was determined first. These results are depicted in Figures 6.1 and 6.2 for breast and lung cancer cell lines respectively. The colour scheme is the same for both Figures. In red are shown the cell lines that carry a mutation for that gene and in blue are the wild type cell lines. The horizontal lines correspond to the average etoposide sensitivity for each group; the greater the distance between these two cell lines, the greater the possible effect of the mutation on etoposide sensitivity.
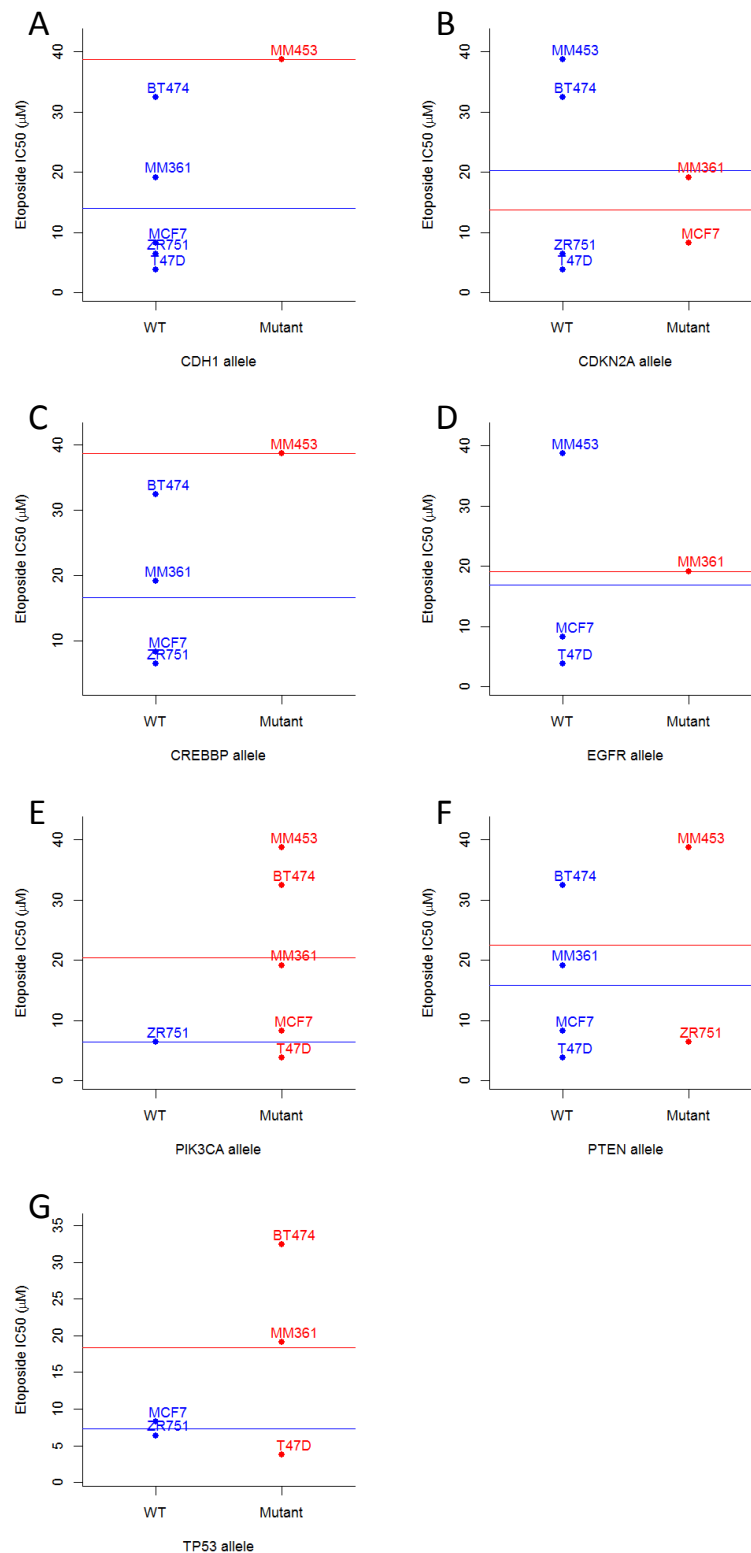
Figure 6.1: A) Graph depicting all breast cancer cell lines with available information for mutation status for gene CDH1, in red the cell lines that carry a mutation for CDH1 and in blue the cell lines that are wild type for this gene. On the y axis are shown the IC50 values for etoposide derived from experimental data discussed in Chapter 5. The red and blue lines represent the mean of the IC50 values of each cell line group and the colours correspond to the mutation status. B) Similar to A but for gene CDKN2A. C) Similar to A but for gene CREBBP. D) Similar to A but for gene EGFR. E) Similar to A but for gene PIK3CA. F) Similar to A but for gene PTEN. G) Similar to A but for gene TP53.
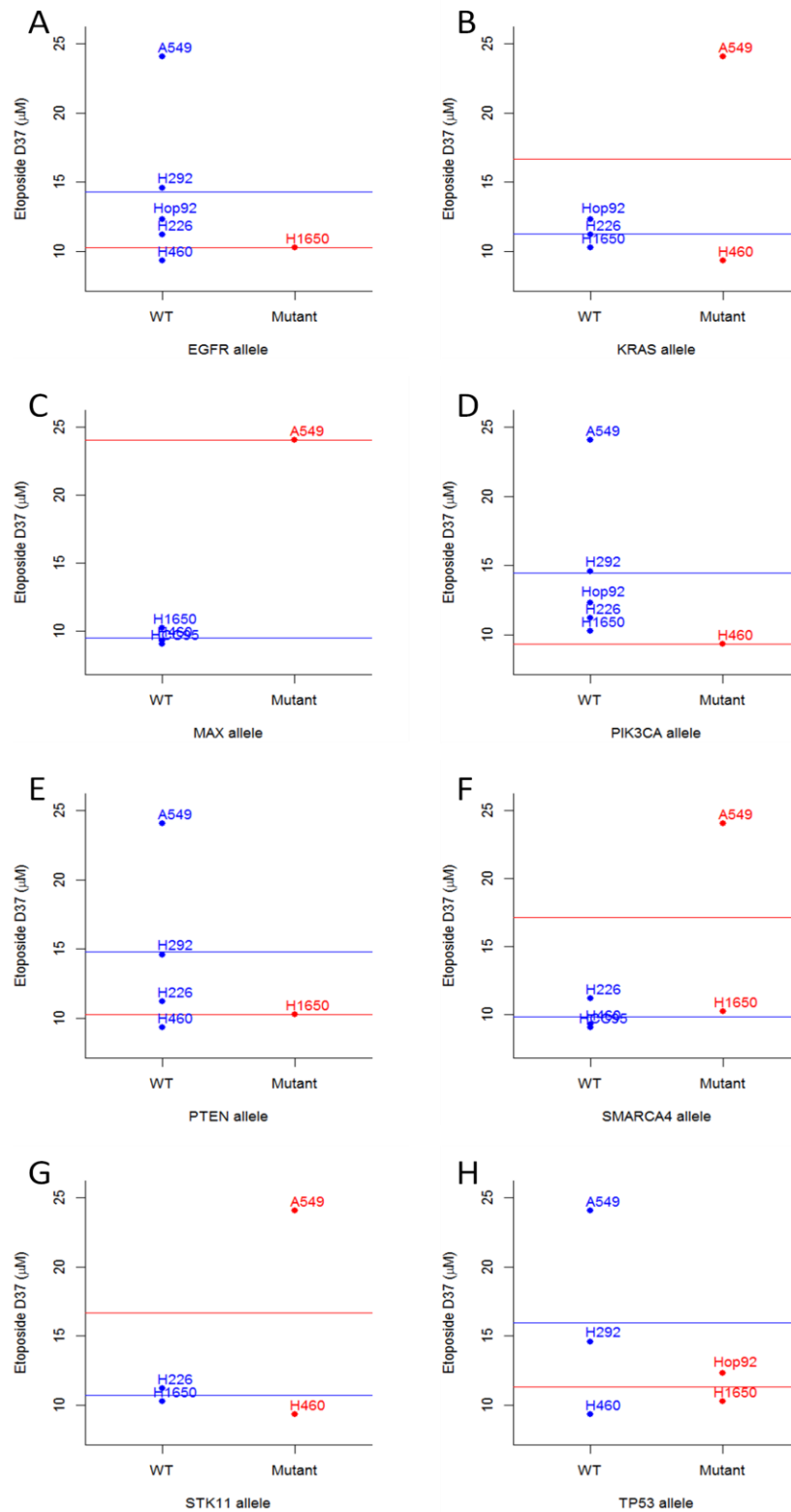
Figure 6.2: A) Graph depicting all lung cancer cell lines with available information for mutation status for gene EGFR, in red the cell lines that carry a mutation for EGFR and in blue the cell lines that are wild type for this gene. On the y axis are shown the D37 values for etoposide derived from experimental data discussed in Chapter 5. The red and blue lines represent the mean of the D37 values of each group and the colours correspond to the mutation status. B) Similar to A but for gene KRAS. C) Similar to A but for gene MAX. D) Similar to A but for gene PIK3CA. E) Similar to A but for gene PTEN. F) Similar to A but for gene SMARCA4. G) Similar to A but for gene STK11. H) Similar to A but for gene TP53.

It is immediately noticeable that not all genes appear in both Figures 6.1 and 6.2, but EGFR, PIK3CA, PTEN and TP53 are shown in both Figures. This simply means that there was information available for the mutation status of these genes for most of the cell lines that I utilised in my experiments, regardless of the tissue of origin. This results from the sheer volume of publications and studies involving these genes and the study of their mutations and their effects on cancer biology.

Table 6.1 shows statistical analysis of the plots for the breast and lung cancer cell lines. The first column for each cell line lists the relevant genes that were derived from the rigorous screening method described previously. The second column depicts the p-values that show the significance of the difference between etoposide IC50 or D37 concentrations of cell lines that carry a mutation for one of the listed genes compared to the wild type cell lines for the same gene. None of those differences are significant for the breast cancer cell lines, as all the p-values are greater than 0.05. For the lung cancer cell lines though, there is one gene, MAX, for which the difference is significant, meaning that this mutation has an effect on etoposide sensitivity independently of TDP2 protein levels or TOP2/TDP2 protein ratios.

The third column shows the estimate of the effect of each mutation on sensitivity to etoposide. If the value is positive, the mutation causes a higher IC50 or D37 value, meaning that it increases resistance to etoposide. In contrast, if the value is negative, then the mutation has the opposite effect, i.e. it causes the cell line to be more sensitive to etoposide. The absolute value depicts the numerical difference between the IC50 or D37 concentrations for the mutated cell lines versus the wild type ones (for that gene), essentially demonstrating how big the effect is. For the breast cancer cell lines, the genes that appear to have the biggest impact judging by the estimates are CDH1 and CREBBP, with EGFR having the smallest impact. All genes apart from CDKN2A appear to have a positive effect on etoposide resistance, meaning that they render the affected cell lines more resistant to etoposide. That is interesting

due to the nature of genes and the cell activities they are associated with. As mentioned though, the observed differences are not significant. For the lung cancer cell lines, the estimate of the mutation effect of the MAX gene is a positive value, thus it is likely that it decreases sensitivity to etoposide in the cell lines that it is found mutated, by increasing the etoposide D37 value by 14.5µM.

Table 6.1: Significance of the differences in the means of etoposide IC50 and D37 concentrations between mutated and wild type forms of genes for breast and lung cancer cell lines, respectively.

Breast Cancer Cell Lines			Lung Cancer Cell Lines		
Gene	p-value of mean IC50	Estimate size of mutation effect on etoposide IC50 (µM)	Gene	p-value of mean D37	Estimate size of mutation effect on etoposide D37 (µM)
CDH1	0.12981	24.68760	EGFR	0.55818	-4.04000
CDKN2A	0.65560	-6.62175	KRAS	0.40100	5.41833
CREBBP	0.19688	22.12750	MAX	0.00253	14.51667
EGFR	0.92805	2.23833	PIK3CA	0.44510	-5.18000
PIK3CA	0.44228	14.03520	PTEN	0.57994	-4.53250
PTEN	0.65380	6.65775	SMARCA4	0.25621	7.30000
TP53	0.37629	11.10150	STK11	0.50526	5.95000
			TP53	0.46534	-4.69167

The next step was looking into the relationship between the mutations, sensitivity to etoposide and protein levels of TDP2 or TOP2/TDP2 and TOP2B/TDP2 protein ratios. All protein levels and ratios for the breast and lung cancer cell lines were obtained from the experimental results described in Chapters 3 and 4. Figure 6.3 depicts the results of this analysis for the breast cancer cell lines and Figure 6.4 for the lung. The black lines represent the best fit lines for TDP2 protein levels or TOP2A/TDP2 or TOP2B/TDP2 protein ratios against etoposide IC50 values, when only the cell lines for which mutation data was available for the gene of interest are taken into consideration, but not accounting for mutation status of that gene. The red lines represent the best fit line of just the cell lines that have been found to be mutated for that gene and the blue lines show the non-mutated cell lines for that same gene. The red and blue lines are always parallel for each gene, due to the limitation imposed by the statistical analysis that requires more data points to add a slope variation.
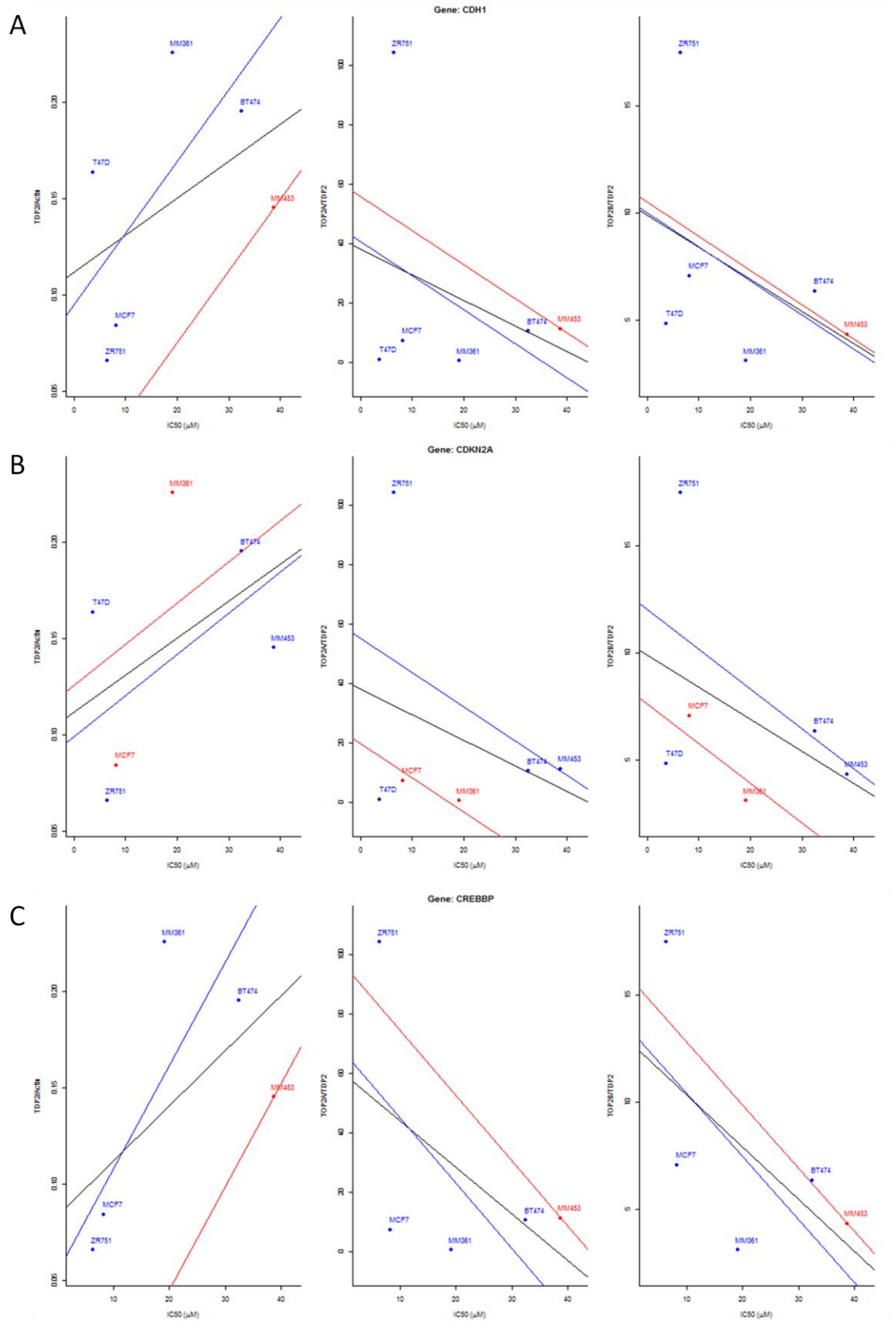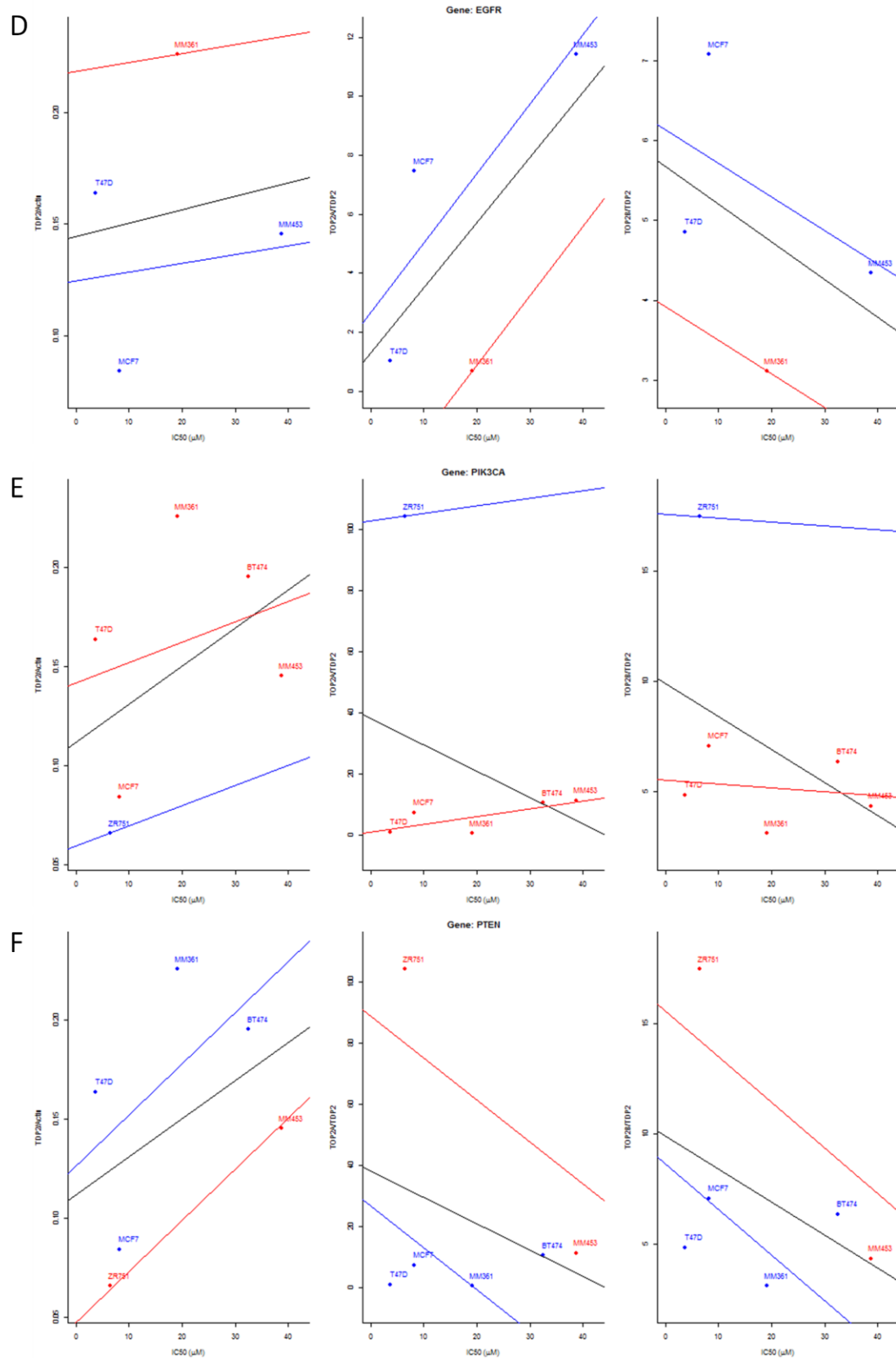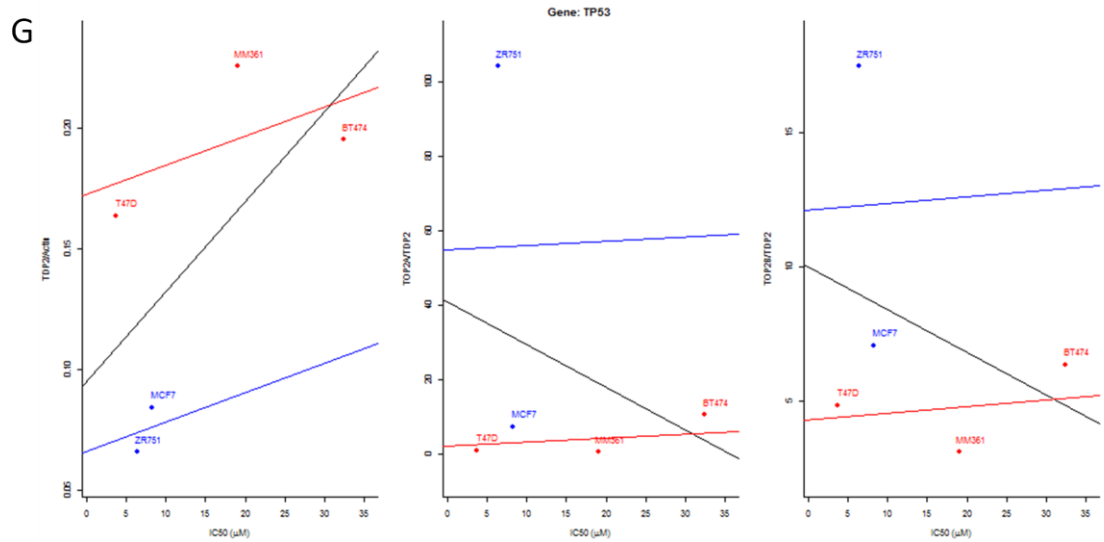
Figure 6.3: Graphs depicting best fit lines of (from left to right) the relationship between TDP2 protein levels and corresponding etoposide IC50 concentrations (µM), TOP2A/TDP2 protein ratios and corresponding etoposide IC50 concentrations (µM), and TOP2B/TDP2 protein ratios and corresponding etoposide IC50 concentrations (µM) in breast cancer cell lines for different genes. The black best fit lines are derived when all the cell lines are considered as one group independent of gene mutation status for each phenotype. The blue and the red lines are derived when the gene mutation status is taken into account, with blue being the wild type cell lines and red the mutated cell lines for that particular gene and phenotype. The genes are: A) CDH1, B) CDKN2A, C) CREBBP, D) EGFR, E) PIK3CA, F) PTEN and G) TP53.

Gene: PIK3CA

Gene: PTEN

Gene: SMARCA4
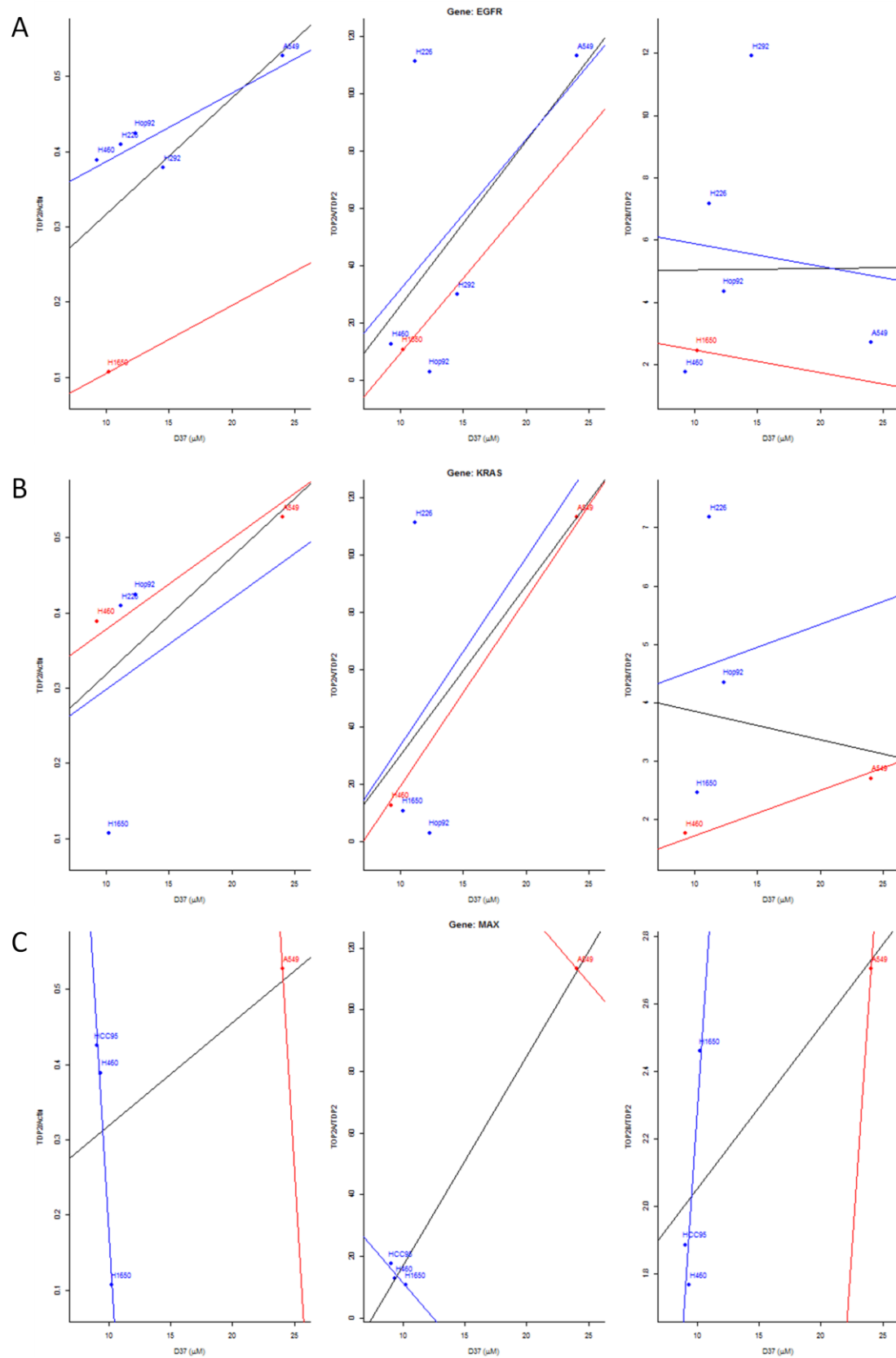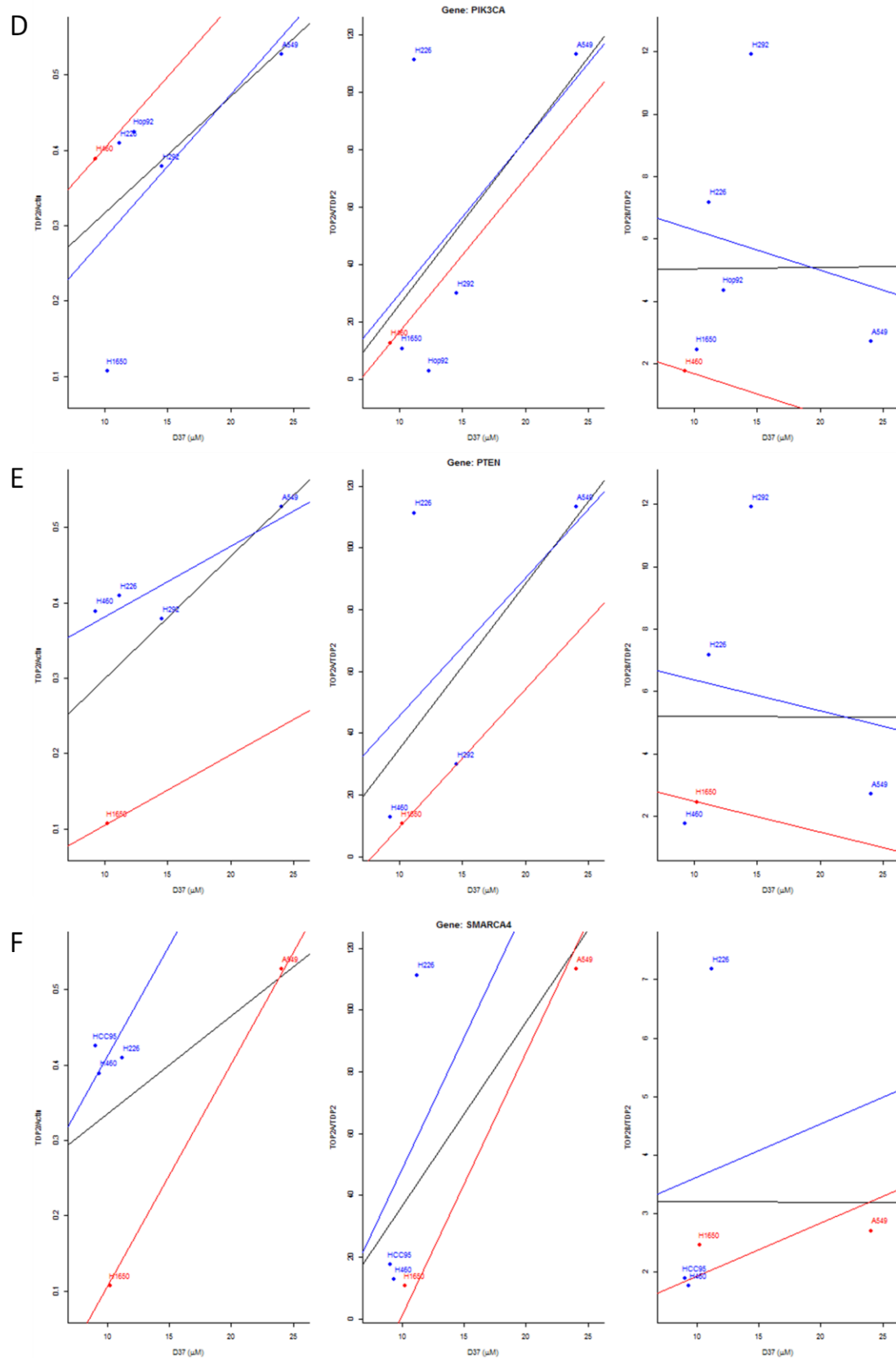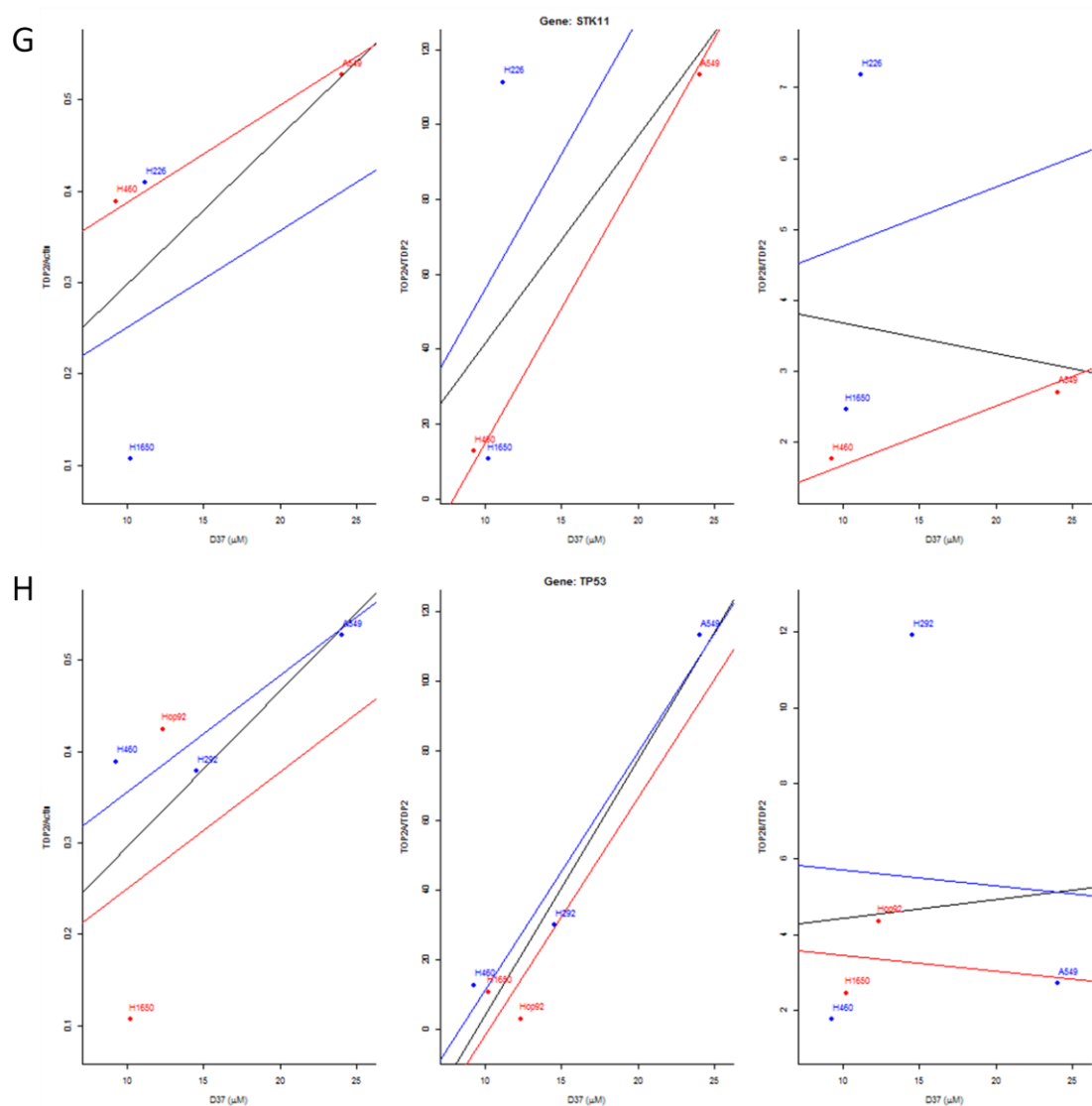
Figure 6.4: Graphs depicting best fit lines from left to right, of the relationship between TDP2 protein levels and corresponding etoposide IC50 concentrations (μM), TOP2A/TDP2 protein ratios and corresponding etoposide IC50 concentrations (μM), and TOP2B/TDP2 protein ratios and corresponding etoposide IC50 concentrations (μM) in lung cancer cell lines for different genes. The black best fit lines are derived when all the cell lines are considered as one group independent of gene mutation status for each phenotype. The blue and the red fit lines are derived when the gene mutation status is taken into account, with blue being the wild type cell lines and red the mutated cell lines for that particular gene and phenotype. The genes are: A) EGFR, B) KRAS, C) MAX, D) PIK3CA, E) PTEN, F) SMARCA4, G) STK11 and H) TP53.

What is important to notice is how far the red and blue lines are from each other and from the black line, as this indicates a greater effect of the mutation on etoposide sensitivity when taking into account TDP2 levels or ratios. Differences between the best fit lines of the mutants and the wild type cell lines can be observed in all genes. For example, in breast cancer cell lines, for most genes apart from CDKN2A, mutation appears to have had an effect on the

relationship between TDP2 protein levels and etoposide sensitivity, but not in the same way. For CDH1, CREBBP and PTEN, the red line is lower than the blue line, which means that overall, the mutated cell lines will be more resistant to etoposide treatment compared to the wild type cell lines for those genes. On the other hand, the opposite is true for the genes EGFR, PIK3CA and TP53, for which the red line is higher than the blue line, indicating mutation reduces etoposide resistance. For TOP2A/TDP2 and TOPB/TDP2 protein ratios, the least severe genes seem to be CDH1 and CREBBP, and with PIK3CA and TP53 showing the greatest differences for both ratios. In lung cancer cell lines, the striking graphs of gene MAX clearly show that the mutation of this gene has had an impact on etoposide sensitivity, rather than on the relationship between TDP2 levels or TOP2/TDP2 ratios and etoposide D37 values. For the rest of the lung cell lines, there seems to be a variable appearance of the best fit lines. EGFR and PTEN mutations appear to have the greatest impact on the relationship between TDP2 levels and D37 concentrations. Another interesting of piece information that we can get from these graphs is the best fit line slope, which implies either a negative or a positive correlation.

The data in Figures 6.3 and 6.4 are easier to interpret with the help of Tables 6.2, 6.3, 6.4 and 6.5, in which the statistical analyses of the graphs are shown. Tables 6.2 and 6.3 refer to the breast cancer cell lines and tables 6.4 and 6.5 to the lung cancer cell lines. The overall question that the analysis is attempting to answer is whether the observed differences are significant or not, how significant and what these differences mean. The first columns of all these Tables depict the genes of interest, with the second ones explaining the phenotype that is explored (TDP2 protein levels, TOP2A/TDP2 protein ratios or TOP2B/TDP2 protein ratios) for each gene.

Gene	Phenotype	p-value of single best fit line slope	Estimate of slope of single best fit line	p-values when only mutation effect is considered	Estimate effect of mt only for phenotype
Table 6.2: Statistical analysis results for the curated gene mutations in breast cancer cell lines – Part 1					
CDH1	TDP2	0.36590	0.00193	0.98322	-0.00171
CDKN2A	TDP2	0.36590	0.00193	0.84588	0.01244
PIK3CA	TDP2	0.36590	0.00193	0.17424	0.09711
PTEN	TDP2	0.36590	0.00193	0.29912	-0.06176
CREBBP	TDP2	0.28808	0.00287	0.97947	0.00249
TP53	TDP2	0.24754	0.00375	0.01561	0.12007
EGFR	TDP2	0.83665	0.00061	0.18768	0.09483
CDH1	TOP2A/TDP2	0.54642	-0.86151	0.79764	-13.39996
CDKN2A	TOP2A/TDP2	0.54642	-0.86151	0.48881	-27.76849
PIK3CA	TOP2A/TDP2	0.54642	-0.86151	0.00006	-98.07523
PTEN	TOP2A/TDP2	0.54642	-0.86151	0.13875	52.92204
CREBBP	TOP2A/TDP2	0.37084	-1.57277	0.74832	-19.34841
TP53	TOP2A/TDP2	0.61764	-1.15003	0.24937	-51.78307
EGFR	TOP2A/TDP2	0.33926	0.22156	0.42945	-5.95159
CDH1	TOP2B/TDP2	0.40775	-0.15019	0.60672	-3.44545
CDKN2A	TOP2B/TDP2	0.40775	-0.15019	0.54659	-3.16780
PIK3CA	TOP2B/TDP2	0.40775	-0.15019	0.00209	-12.34286
PTEN	TOP2B/TDP2	0.40775	-0.15019	0.25979	5.56088
CREBBP	TOP2B/TDP2	0.27042	-0.24469	0.59103	-4.17701
TP53	TOP2B/TDP2	0.58078	-0.15981	0.16205	-7.50754
EGFR	TOP2B/TDP2	0.55542	-0.04741	0.30317	-2.31174

Gene	Phenotype	p-value of mt & wt best fit lines slope	Estimate of slope of mt & wt best fit lines	p-value of mt & wt intercepts	Estimate of difference between mt & wt intecepts
Table 6.3: Statistical analysis results for the curated gene mutations in breast cancer cell lines – Part 2					
CDH1	TDP2	0.24754	0.00375	0.38881	-0.09420
CDKN2A	TDP2	0.39732	0.00215	0.69547	0.02670
PIK3CA	TDP2	0.63822	0.00103	0.32690	0.08266
PTEN	TDP2	0.19926	0.00259	0.17516	-0.07900
CREBBP	TDP2	0.19037	0.00539	0.31800	-0.11674
TP53	TDP2	0.46008	0.00122	0.06671	0.10657
EGFR	TDP2	0.88456	0.00040	0.39504	0.09394
CDH1	TOP2A/TDP2	0.61764	-1.15003	0.85286	14.99160
CDKN2A	TOP2A/TDP2	0.46349	-1.15397	0.43070	-35.40981
PIK3CA	TOP2A/TDP2	0.15958	0.24994	0.00023	-101.58324
PTEN	TOP2A/TDP2	0.22474	-1.37713	0.09351	62.09060
CREBBP	TOP2A/TDP2	0.46256	-2.20893	0.74308	29.52977
TP53	TOP2A/TDP2	0.96819	0.10791	0.41424	-52.98104
EGFR	TOP2A/TDP2	0.34496	0.23655	0.38168	-6.48107
CDH1	TOP2B/TDP2	0.58078	-0.15981	0.96041	0.49978
CDKN2A	TOP2B/TDP2	0.35478	-0.18655	0.42989	-4.40312
PIK3CA	TOP2B/TDP2	0.78558	-0.01789	0.01118	-12.09181
PTEN	TOP2B/TDP2	0.21147	-0.20786	0.15849	6.94475
CREBBP	TOP2B/TDP2	0.43196	-0.29563	0.83024	2.36463
TP53	TOP2B/TDP2	0.93401	0.02505	0.31102	-7.78563
EGFR	TOP2B/TDP2	0.62820	-0.04228	0.46571	-2.21709

Table 6.4: Statistical analysis results for the curated gene mutations in lung cancer cell lines – Part 1

Gene	Phenotype	p-value of single best fit line slope	Estimate of slope of single best fit line	p-values when only mutation effect is considered	Estimate effect of mt only for phenotype
EGFR	TDP2	0.20798	0.01554	0.00819	-0.31857
PIK3CA	TDP2	0.20798	0.01554	0.91320	0.01995
TP53	TDP2	0.23647	0.01707	0.30214	-0.16595
PTEN	TDP2	0.24889	0.01631	0.02557	-0.31891
SMARCA4	TDP2	0.36179	0.01301	0.60560	-0.09042
KRAS	TDP2	0.28404	0.01562	0.38555	0.14487
MAX	TDP2	0.44462	0.01373	0.38719	0.22047
STK11	TDP2	0.36553	0.01628	0.35120	0.20026
EGFR	TOP2A/TDP2	0.20174	5.75785	0.50541	-43.30980
PIK3CA	TOP2A/TDP2	0.20174	5.75785	0.53068	-40.91746
TP53	TOP2A/TDP2	0.00942	7.37883	0.34314	-45.19626
PTEN	TOP2A/TDP2	0.26653	5.36715	0.41385	-56.09904
SMARCA4	TOP2A/TDP2	0.18567	5.97560	0.80925	14.85592
KRAS	TOP2A/TDP2	0.25353	5.93554	0.73921	21.41217
MAX	TOP2A/TDP2	0.00763	6.81010	0.00165	99.70762
STK11	TOP2A/TDP2	0.33718	5.54654	0.97979	2.03344
EGFR	TOP2B/TDP2	0.98853	0.00547	0.52470	-3.12402
PIK3CA	TOP2B/TDP2	0.98853	0.00547	0.41182	-3.95507
TP53	TOP2B/TDP2	0.91069	0.04955	0.65988	-2.06402
PTEN	TOP2B/TDP2	0.99613	-0.00218	0.55704	-3.43446
SMARCA4	TOP2B/TDP2	0.99741	-0.00072	0.68558	-1.02972
KRAS	TOP2B/TDP2	0.82748	-0.04875	0.27185	-2.42550
MAX	TOP2B/TDP2	0.21650	0.04846	0.25930	0.66675
STK11	TOP2B/TDP2	0.87888	-0.04305	0.39524	-2.58515

Table 6.5: Statistical analysis results for the curated gene mutations in lung cancer cell lines – Part 2

Gene	Phenotype	p-value of mt & wt best fit lines slope	Estimate of slope of mt & wt best fit lines	p-value of mt & wt intercepts	Estimate of difference between mt & wt intercepts
EGFR	TDP2	0.04794	0.00910	0.00483	-0.28180
PIK3CA	TDP2	0.20920	0.01901	0.50934	0.11844
TP53	TDP2	0.45960	0.01283	0.56017	-0.10576
PTEN	TDP2	0.10423	0.00942	0.02501	-0.27623
SMARCA4	TDP2	0.01834	0.02959	0.02291	-0.30642
KRAS	TDP2	0.53083	0.01208	0.69844	0.07939
MAX	TDP2	0.06111	-0.27372	0.05806	4.19393
STK11	TDP2	0.68559	0.01065	0.66772	0.13688
EGFR	TOP2A/TDP2	0.32488	5.25292	0.73492	-22.08801
PIK3CA	TOP2A/TDP2	0.33555	5.37455	0.84753	-13.07730
TP53	TOP2A/TDP2	0.04190	6.85577	0.49307	-13.03127
PTEN	TOP2A/TDP2	0.44729	4.47282	0.63324	-35.82596
SMARCA4	TOP2A/TDP2	0.20914	8.54731	0.47385	-47.53945
KRAS	TOP2A/TDP2	0.37844	6.56555	0.84675	-14.16219
MAX	TOP2A/TDP2	0.35717	-4.70270	0.15957	167.97520
STK11	TOP2A/TDP2	0.45880	7.23343	0.68642	-41.00545
EGFR	TOP2B/TDP2	0.86969	-0.07266	0.57249	-3.41756
PIK3CA	TOP2B/TDP2	0.76643	-0.13023	0.44857	-4.62967
TP53	TOP2B/TDP2	0.94562	-0.04102	0.73295	-2.25646
PTEN	TOP2B/TDP2	0.86067	-0.09913	0.62137	-3.88378
SMARCA4	TOP2B/TDP2	0.79451	0.09086	0.68083	-1.69303
KRAS	TOP2B/TDP2	0.76045	0.07794	0.36737	-2.84780
MAX	TOP2B/TDP2	0.22908	0.54748	0.24863	-7.28090
STK11	TOP2B/TDP2	0.83433	0.08381	0.56480	-3.08383

The third column of Tables 6.2 and 6.4 represents the significance of the slope of the best fit line when both mutant and wild type cell lines are considered together for each gene, which in the graphs of Figures 6.3 and 6.4 are drawn in black. The p-value is the probability of the null hypothesis that the slope is 0, therefore lower p-values mean better support for a non zero slope. Significant p-value means that there is a significant relationship between each phenotype and resistance to etoposide regardless of mutation status for each gene. If these values are greater than 0.05, then there is no significant relationship between each phenotype

and etoposide resistance when both wt and mt forms of the gene are considered together. Most of the slopes are not significant, which is precisely what I saw in Chapter 5, when I was investigating a possible correlation between TDP2 protein levels and TOP2/TDP2 ratios, before investigating the possible role of mutations. In the lung cancer cell lines, though, for the TOP2A/TDP2 ratio, there are two genes that give a significant p-value and these are TP53 and MAX. That could just mean though, that for the cell lines for which I had data for these two mutations, there is the possibility of a relationship between that protein ratio and D37 concentrations. Which is why in the discussion, a more rounded picture for every gene will be presented. The next column of the same tables shows the estimate of the slopes of the best fit line for both mt and wt cell lines for each gene and phenotype. A positive estimate means there is a positive slope and negative means a negative slope, while a large absolute value means a steeper slope.

The next column in Tables 6.2 and 6.4 shows the p-values of the difference between TDP2 levels or TOP2/TDP2 ratios when they are considered on their own, independently of etoposide sensitivity, but when the mutation status of the genes is taken into account. These p-values represent the probability of the null hypothesis that the difference is 0. If a value is lower than 0.05, then the difference is considered significant. For most of the genes, there is no significant difference, which means that the mutations do not have an effect on the levels of TDP2 or the TOP2/TDP2 ratios. For the breast cancer cell lines there is one exception to this, gene PIK3CA, which appears to have a significant effect on both the TOP2A/TDP2 and the TOP2B/TDP2 protein ratio. For the lung cancer cell lines, gene mutations EGFR and PTEN have a significant effect on TDP2 protein levels. MAX has a significant effect on TOP2/TDP2 ratio, although that is probably the effect the mutation has on etoposide sensitivity rather than on the protein ratio. The adjacent column of Tables 6.2 and 6.4 lists the estimate of effect of the gene mutations on TDP2 levels and TOP2/TDP2 ratios independent of etoposide sensitivity. If that value is negative, the mutation's effect is a decrease of TDP2 protein or of the protein

ratio TOP2/TDP2 and if positive then the effect is the opposite. The absolute value depicts the magnitude of the effect. As one can observe, there is variety of how each mutation has affected each phenotype, with the biggest effect being the one on the TOP2A/TDP2 protein ratio for both types of cancer cell lines, where there is a significant large effect of PIK3CA in the breast cancer cell lines and MAX in lung. While EGFR and TP53 had a significant p-value for TDP2 levels, the estimate effect is small. The smallest changes are seen in TDP2 levels, so perhaps the larger levels of change for the ratio phenotypes could be due to TOP2-mutation interactions.

Moving on to Tables 6.3 and 6.5, the third column lists the p-values of the slope for the two best fit lines (red and blue in the graphs of Figures 6.3 and 6.4) that result when the mutation status for each gene is taken into consideration and the cell lines are considered separately. If these values are significant (less than 0.05), then there is a significant relationship between etoposide resistance and TDP2 levels or TOP2/TDP2 ratios when the mutation for each gene is taken into account, while the opposite is true when the p-values are not significant. There are no significant values in this list for the breast cancer cell lines, but it is interesting to observe how these p-values compare to the p-values when the mutation status of the gene is not considered. A lower p-value for the mt and wt best fit line slopes compared to the non-discriminatory best fit line slopes suggests that the separate best fit lines fit the data better, meaning that taking into account the mutation of the genes results in a stronger relationship between etoposide sensitivity and TDP2 levels or TOP2/TDP2 ratios. If the opposite is true, then taking into account the mutation status doesn't increase the relationship. There appears to be a variation, with several gene mutations seemingly fitting better to the separate best fit lines, for example CDH1, PTEN and CREBBP for TDP2 protein levels. This doesn't necessarily carry onto both TOP2A/TDP2 and TOP2B/TDP2 ratios, where for the same genes, only PTEN showed an improved relationship for TOP2A/TDP2 ratio with etoposide sensitivity when the mutation is taken into account. For the lung cancer cell lines, there is a significant p-value for

the EGFR and SMARCA4 mutations for TDP2 levels and TP53 for the TOP2A/TDP2 ratio. Similar to the breast cancer cell lines, there is a variation in how well the data all together fit a single best fit versus two best fit lines, which separate the mt and wt cell lines for each gene. The estimate of the slope of mt and wt best fit lines shows the magnitude and direction of the relationship between TDP2 and etoposide sensitivity when considering mt cell lines separate from wt cell lines. As one can observe, none of the mutations have a big effect from the absolute values, but they vary in whether they are positive or negative.

The next column of Tables 6.3 and 6.5 shows the significance of mutation status on the gene when considered in conjunction with etoposide resistance (as an effect on TDP2 levels or TOP2/TDP2 ratios). While the previous p-values measure the probability of the null hypothesis that the slope is 0 when taking into account the mutation status of each gene, the p-values of this column show how significant the difference is in TDP2 levels or TOP2/TDP2 ratios for points with the same etoposide sensitivity but different forms of the gene. This is measured by identifying if the intercept between the best fit lines for mutant and wt is different. In other words, how likely is it for a cell line with a mutation for a gene of interest to have a different response to etoposide, compared to another cell line that doesn't carry that mutation but has the same TDP2 levels or TOP2/TDP2 ratios as the mutated cell line? If this p-value is significant, then the answer to the above question is that it's likely, while if the p-value is not significant, then the answer is unlikely. While for most of the genes in breast cancer cell lines, this p-value is not significant, there is one gene that demonstrated significant p-values for both TOP2A/TDP2 and TOP2B/TDP2 ratios and that is PIK3CA, although it doesn't show this for TDP2 protein levels. This is interesting, because the same gene for the same phenotypes demonstrated a significant p-value for when only the phenotypes and not sensitivity to etoposide where considered, along with the mutation status of the genes. It is likely that the significant p-value seen for the intercepts is due to the difference in TOP2/TDP2 ratios between mutated and wild type cell lines, rather than the response to etoposide. For TP53, the

p-value is the lowest non-significant value for TDP2 protein levels. In the lung cancer cell lines, there are more genes that have a significant p-value, but that is limited to TDP2 protein levels only. These genes are EGFR, PTEN and SMARCA4. The final column of Table 6.3 and 6.5 lists the estimate of difference between the intercepts of the mt and the wt cell lines for each gene. If this value is positive, the mutation is likely to decrease resistance to etoposide for the same levels of TDP2 or TOP2/TDP2 ratios and if negative it is likely to have the opposite effect and increase resistance to etoposide. There is a variation of the effect of each mutation to the three studied phenotypes. For most of the cell lines the absolute value is far greater for TOP2A/TDP2 ratio compared to the other phenotypes, with PIK3CA demonstrating the highest change compared to all the other mutations in breast cancer cell lines and MAX in lung cancer cell lines.

6.3 Discussion

Given the complexity of the statistical model used and due to the large amount of data that needed to be organised and analysed, a summary of all the available information for each gene is essential in order to reach a conclusion of the nature of the effect a mutation can have in TDP2 protein levels, TOP2/TDP2 protein ratios and sensitivity to etoposide, when these effects are considered together or separate. Although the method used to analyse this data is complex, it is to my knowledge the simplest possible way of investigating the relationship between the 3 variables (Etoposide sensitivity, mutation status and TDP2 levels or ratios).

CDH1 is a gene that plays a role in cell proliferation by targeting multiple cell-cycle regulators for ubiquitin-dependent degradation (Eguren et al., 2014). This is a high-penetrance breast cancer gene and its mutation has been reported to be associated with hereditary breast cancer (Shiovitz and Korde, 2015). It has also been reported that cells lacking CDH1 expression display various defects, including impaired DNA repair (de Boer et al., 2016). TOP2A has been identified as a CDH1 target, with TOP2A levels showing an increase in CDH1 depleted cells, and

lack of CDH1 has been shown to result in dramatic sensitivity to etoposide (Eguren et al., 2014). My results have shown that the CDH1 mutation affects etoposide resistance in breast cancer cell lines by increasing the IC50 concentration of mutated cells by 24 μM, rendering them more resistant. That result, however, is not significant and therefore cannot be taken as a disagreement to what other people have seen. When considering the mutated cell lines separately from the wt ones in regard to a possible relationship between TDP2 levels and etoposide sensitivity, there was an increase in how well the data fit the best fit lines, but the p-values for the slopes showed than none of them were significant. The same can be said for the intercepts of the mt and wt best fit lines. There was also no significant effect of the mutation on TDP2 levels on their own. For TOP2/TDP2 levels, none of the results are significant too, but there is a difference in the slope p-values of the single best fit line versus the separate best fit lines, where the data appear to fit better to the single best fit line, as the p-value for that is lower. Due to only having data for 6 breast cancer cell lines for this gene and the fact that all the statistical analysis points towards non-significant mutation effect, I cannot draw the conclusion that CDH1 mutation has an effect on etoposide sensitivity alone or in relation to TDP2 protein levels and TOP2/TDP2 ratios, or the phenotypes on their own.

The next gene is CDKN2A, which is a gene known to cause cancer susceptibility, with increased breast cancer risk described in families with CDKN2A mutations (Maxwell et al., 2015). It is estimated to be one of the most commonly mutated genes in cancer (for example, pancreatic and melanoma) and its proteins act as tumour suppressors, by preventing premature transition from G1 to S phase (Zhen et al., 2015). My results for the breast cancer cell lines, much like CDH1, did not produce significant p-values. CDKN2A mutation does not affect etoposide sensitivity, TDP2 levels or TOP2/TDP2 ratios. The slope p-values for a single best fit line were lower than the slope for the two best fit lines for TDP2, but the opposite for both TOP2/TDP2 ratios. The difference between the intercepts for the mutated cell lines and the wt ones also didn't have a significant p-value. Therefore, I cannot conclude that a CDKN2A

mutation could have an effect on etoposide sensitivity alone or in relation to TDP2 protein levels and TOP2/TDP2 ratios, or the phenotypes studied on their own.

CREBBP is a gene involved in the transcriptional co-activation of different transcription factors, and chromosomal translocations involving CREBBP have been associated with acute myeloid leukemia, especially after adjuvant chemotherapy for breast cancer (Gupta et al., 2014). The analysis results for this gene were not significant and were very similar to the CDKN2A results. CREBBP mutation does not have an effect on etoposide IC50 concentration, TDP2 levels or TOP2/TDP2 ratios. The slope p-values for the single best fit line were lower than the slope of the two best fit lines for TDP2, but the opposite was true for both TOP2/TDP2 ratios. The difference between the intercepts for the mt cell lines and the wt ones also didn't produce a significant p-value. Therefore, I conclude that CREBBP is unlikely to have an effect on etoposide sensitivity alone or in relation to TDP2 protein levels and TOP2/TDP2 ratios, or the phenotypes studied on their own.

EGFR (epidermal growth factor receptor) is a well studied gene in relation to both breast and lung cancer. It is overexpressed in a subset of HER2-positive breast cancers, and co-expression of HER2 and EGFR has been associated with poor clinical outcome (Lee et al., 2015b). Inhibition of EGFR activation has been shown to limit proliferation in breast cancer cells (Sathya et al., 2014). Patients with NSCL cancer tumours have common and uncommon mutations at the EGFR gene (Yang et al., 2015a). Inhibition of EGFR has been shown to downregulate TOP2 expression (Qi et al., 2013). EGFR mutation did not have a significant effect in breast cancer cells. It didn't have a significant impact on etoposide sensitivity, TDP2 levels or TOP2/TDP2 ratios. The slope p-values when the mutation was taken into consideration and when it wasn't, were very similar. A much higher number of data points seems to be necessary in order to study this mutation's effect on the phenotypes and relationships I was investigating. On the other hand, even though most p-values are also not

significant in the lung cancer cell lines, the p-value for the effect of EGFR mutation on TDP2 protein levels independent of etoposide sensitivity, is significant and combined with the estimate of the effect, it appears that the mutation decreased the TDP2 level. This is likely the explanation for why there is also a significant difference between the intercepts of the best fit lines when the mt and wt cell lines are considered separately, which shows that the data fits the separate best fit lines better, as well as the slope p-value of these best fit lines. This is a very interesting result, that could be studied further in a laboratory setting, where additional lung cancer cell lines for which information about EGFR mutation status is available, can be pre-selected for measuring TDP2 protein levels, rather than using cell lines with an unavailable EGFR profile.

PIK3CA is another gene that has been shown to be involved in cancer. PIK3CA mutations are the most common in breast cancer, having been reported in over one third of cases. They are very prominent in HER2-positive breast tumours (as high as 23%), which is associated with poor prognosis (Majewski et al., 2015). They frequently co-exist with other common cancer related mutations, such as EGFR and KRAS - two genes that are relevant to this chapter, in NSCLC and they are related to poor prognosis for that type of cancer (Wang et al., 2014). PIK3CA mutations have also been found to be negatively associated with TOP2A gene amplifications, where 28% of the wt breast tumours studied had an amplified TOP2A gene, compared to 14% of mt breast tumours (Fountzilas et al., 2012). In the breast cancer cells, the PIK3CA mutation did not have a significant effect on etoposide sensitivity or TDP2 protein levels, but it did have a significant effect on TOP2A/TDP2 and TOP2B/TDP2 ratios. The effect estimate is negative for both ratios, thus a negative correlation, meaning the mutation decreases the ratio. This could be interpreted as the PIK3CA mutation having a possible negative effect on TOP2 protein levels, which agrees with the group that saw a similar effect on TOP2A expression levels (Fountzilas et al., 2012). The results for the lung cancer cell lines are not as interesting, as none of the observed differences are significant. The mutation has no

effect on etoposide sensitivity, TDP2 levels or TOP2/TDP2 ratios on their own, or etoposide sensitivity in relation to all the studied phenotypes.

PTEN is a commonly known tumour suppressor that is frequently mutated in a large number of cancers, including breast and NSCL cancers (Kechagioglou et al., 2014, Pérez-Ramírez et al., 2015). In breast cancer cell lines, none of the calculated p-values were of significance, but for all phenotypes studied, the slope p-value of the best fit lines arising when the mt cell lines are considered separately from the wt cell lines, is smaller than when all cell lines are grouped together. For the lung cancer cell lines, the PTEN mutation has a significant effect on TDP2 protein levels, but not the other phenotypes or on etoposide sensitivity. From the value of the effect estimate, one can see that the mutation reduces the levels of TDP2, as this value is negative. When looking at the relationships between the phenotypes and etoposide sensitivity, the difference seen between the intercepts of the best fit lines that arise from separating the cell lines into two groups is only significant for the TDP2 phenotype, which can be attributed to the effect the mutation has on TDP2 protein levels.

Even a quick search for TP53 and cancer will give thousands of results. There are dedicated databases that collect information on mutations solely of this gene. TP53 mutations in breast cancer have been associated with poor prognosis, but they have different clinical relevance in different breast cancer subtypes (Silwal-Pandit et al., 2014). TP53 mutations are also very common in NSCLC, but their role as a prognostic marker is not clear due to the large variation (Govindan and Weber, 2014). TDP2 has been found to be a transcriptional target of mt p53 and a mt p53 knock-down reduced both the mRNA and protein levels of TDP2 (Do et al., 2012). In the same publication it was also shown that suppressing of TDP2 sensitizes cells that express mt p53 to etoposide, and that mt p53 and TDP2 are frequently overexpressed in lung cancer (Do et al., 2012). In the breast cancer cell lines, I indeed observed a significant effect of TP53 mutation on TDP2 protein levels, with a positive estimate effect, which means that the

mutation increased TDP2 levels. I didn't find a similar effect in the lung cancer cells, which could be attributed to lack of enough data points. On the other hand there was a significant slope for both the single and double best fit lines.

KRAS is a gene that is normally involved in normal tissue signalling, but its mutation predicts an increased risk of several cancers, including NSCLC and triple negative breast cancer, especially for women (McVeigh et al., 2015). KRAS mutant lung cancers comprises of 25-30% of lung adenocarcinomas (subtype of NSCLC) without an effective treatment available (Mortazavi et al., 2015). However any differences observed in the lung cancer cell lines, were unfortunately non-significant.

MAX encodes a protein that interacts with the Myc protein, an oncoprotein implicated in cell proliferation, differentiation and apoptosis (Cascón and Robledo, 2012). Mutations of MAX have been associated with hereditary pheochromocytoma, which is a rare neural crest cell tumour that localises mainly within the adrenal medulla and rarely metastasizes (Cascón and Robledo, 2012). This gene was found to have a significant effect on both etoposide sensitivity (the highest observed in lung cancer) and TOP2A/TDP2 protein ratio in lung cancer cell lines. With all statistical analysis, there is always a chance that a significant result is a false positive and it's likely that this is true for this gene, as it has been shown to be involved in a rare form of cancer and not in breast or lung cancer. The gene did not, however, demonstrate a significant effect when these two factors were considered together, likely because there is very little variation in either the mt or wt group for etoposide IC50 values.

SMARCA4, also known as BRG1, has been proposed as a possible tumour suppressor, as it has been found to induce cell growth arrest after ectopic expression in tumour cells lacking SMARCA4, as well as its possible interaction with tumour suppressor proteins, such as BRCA2 (Medina et al., 2008). In lung cancer cell lines for TDP2 protein levels, the mt cell lines and the wt cell lines, when considered separately, fit their respective best fit lines significantly better

that when considered separately, with the slope being of significance too. These results suggest that this gene may be of interest when studying the effect of TDP2 levels.

STK11, also known as LKB1, is somatically inactivated through point mutations and large deletions in lung cancers, which are very common in non-small cell lung cancer, especially adenocarcinomas and colorectal cancer (Rodriguez-Nieto and Sanchez-Cespedes, 2009, Launonen, 2005). Unfortunately, the statistical analysis for this gene in lung cancer cell lines has shown that any differences are not significant.

In summary, this chapter has revealed some very interesting results in its attempt to explain why no correlation was found between sensitivity to etoposide and TDP2 protein levels or TOP2/TDP2 ratios. After a rigorous process of elimination of available data from different database sources and statistical modelling, mutations in certain genes were shown to have played a possible role. EGFR mutation in lung cancer cell lines shows a significant positive correlation with TDP2 protein levels, and could possibly affect the relationship between TDP2 protein levels and sensitivity to etoposide, although that could be through its effect on TDP2. PIK3CA mutation in breast cancer cell lines had a significant negative effect on TOP2A/TDP2 and TOP2B/TDP2 ratios, but not on TDP2 protein levels. This is in agreement with a group that has shown that the PIK3CA mutation decreased TOP2 expression levels (Fountzilas et al., 2012). PTEN mutation in lung cancer cell lines had a significant negative effect on TDP2 protein levels and it was also shown to have a possibly significant effect on the relationship between TDP2 protein levels and sensitivity to etoposide, although similar to EGFR that could be due to its effect on TDP2 levels. TP53 mutation in breast cancer cell lines shows a significant positive correlation with TDP2 protein levels. SMARCA4 mutation in lung cancer cell lines appears to have a significant effect on the relationship between TDP2 protein levels and etoposide sensitivity, when the mutation is taken into account.

In conclusion, I identified three mutations (EGFR, PTEN and TP53) that could possibly have an effect on TDP2 protein levels, something that would be interesting to explore further, especially how this effect arose. Mutant TP53 has already been shown to target TDP2 transcription and mt p53 knock-down was shown to reduce both the mRNA and protein levels of TDP2, which agrees with the positive correlation I discovered (Do et al., 2012). I also identified a mutation (PIK3CA) that appears to negatively affect TOP2 protein levels, which also agrees with published work (Fountzilas et al., 2012). Finally, I identified a mutation (SMARCA4) that has a significant effect on the relationship between TDP2 protein levels and etoposide sensitivity. These 5 mutations could all be contributing to the lack of correlation I observed in Chapter 5, between etoposide sensitivity and either TDP2 protein levels or TOP2/TDP2 protein ratios.

All this data serves to show how certain biological relationships that might appear in a hypothesis as straightforward, simple and logical, are actually very difficult to verify in a laboratory setting. It is not surprising that many research groups have moved on to exploring TDP2 inhibitors, rather than looking at intrinsic levels of TDP2. As I have shown here the naturally occurring variation in two types of cancer cell lines, is not enough to significantly affect etoposide sensitivity.

This chapter has also made a contribution in the mutation profiles of several breast and lung cancer cell lines that are available for use in the Caldecott lab and as an extension, the whole Genome Damage and Stability Centre, as well as other collaborators. This extensive list of thousands of mutations is available in Appendices 6 and 7. This chapter demonstrated the importance of being aware of the mutation profiles of the cell lines that a researcher is considering using for an experiment, as certain cell lines are not used as extensively in scientific experiments and therefore a lot less information is available to help with data interpretation. Having this information available for the cell lines used can serve to explain

certain unexpected results. In some cases potentially interesting mutations may have been missed simply because information was unavailable for enough cell lines.

7. CHAPTER SEVEN – Potential use of TDP2 as a biomarker of sensitivity to a panel of ant-cancer drugs

7.1 Introduction

As it has been discussed in previous chapters, TDP2 is an important factor in TOP2-mediated DSBs that has received little attention in terms of its possible relationship with non-TOP2 related drugs. There are a number of anti-cancer drugs or compounds currently in clinical trials, and so it stands to reason that some of them may have unknown interactions with TDP2 or a possible correlation with TDP2 protein levels. I have access to TDP2 protein levels in 8 lung cancer cell lines and 8 breast cancer cell lines from my experimental results from Chapters 3 and 4, and there are several databases dedicated to anti-cancer drugs. Therefore I decided to mine data and take a first step towards identifying other relevant compounds for which TDP2 might be a biomarker for sensitivity or clinical outcome based on the phenotype profiles I have experimental data for. Using statistical tests, I was able to organise and analyse the data collected, combined with my own data to make a shorter list that could be tested experimentally in the future.

The first step was to collect the appropriate data. The database with information about the most drugs and their pharmacological profiling across the largest number of cancer cell lines was CancerDR (Kumar et al., 2013). CancerDR is a database that has collected information for 148 anti-cancer drugs and their profile in 952 cancer cell lines from the COSMIC and CCLE databases and has organised it in an accessible manner. Out of these drugs, 36 are FDA approved drugs, 48 are or have been in clinical trials and 64 are experimental drugs. Figure 7.1 summarises the curation methods applied to the data collected from COSMIC and CCLE databases.
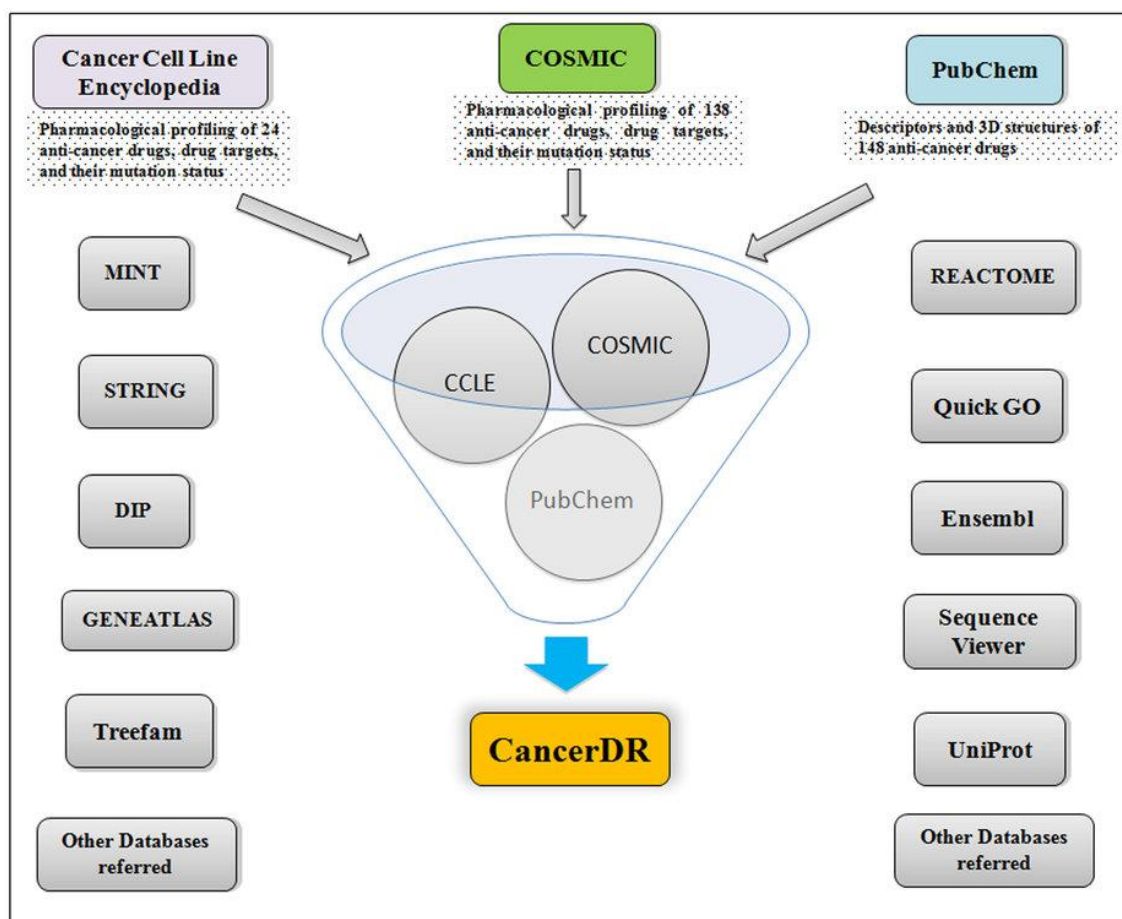
Figure 7.1: Schematic representation of the curation procedure followed by CancerDR (Kumar et al., 2013).

CancerDR is composed of a curated version of the release 2 of 'Genomics of drug sensitivity in cancer' (a COSMIC project), where 138 anti-cancer drugs targeting a wide range of therapeutic targets were screened in 714 cancer cell lines, and a study by CCLE, where 24 drugs were screened in 503 cancer cell lines (Barretina et al., 2012, Garnett et al., 2012). Further information about the anti-cancer drugs is mined from PubChem (Bolton et al., 2008). I also utilised PubChem as an information source to acquire information about the relevant anti-cancer drugs I studied in this chapter. In addition I used DrugBank, a large bioinformatics and cheminformatics resource that combines chemical, pharmacological and pharmaceutical drug data with sequence, structure, and pathway information for the drug target (Wishart et al., 2006). To access further information about certain drugs' clinical trial status, I also referred to ClinicalTrials.gov, a clinical trial database maintained by the U.S. National Library of Medicine

at the National Institutes of Health that currently lists 230,427 studies with locations in 194 countries (Zarin et al., 2016).

7.1.1 Aims of this chapter

This chapter aims to use existing information on drug sensitivities to assess whether TDP2 protein levels could be used as a biomarker for sensitivity to drugs other than etoposide.

7.2 Results

The first step of selection was to search for anti-cancer drugs in CancerDR, for which there was information on IC50 values for the breast and lung cancer cell lines I had studied in the previous chapters and had characterised for TDP2 protein levels and TOP2/TDP2 protein ratios. As with the mutations in the previous chapter, I was not able to find this information for each drug in every cell line I had utilised. In order to obtain meaningful statistical results, I required drugs for which I had information for a minimum of 5 cancer cell lines for each type, with at least one variation in IC50 values. The statistical analysis method is described in detail in Chapter 2.6. This limitation reduced the analysed drugs to 113 for the breast cancer cells and 194 for the lung cancer cells.

For these two sets of drugs the Pearson correlation coefficient or r coefficient and the corresponding p-values were calculated for the three phenotypes. r coefficients show how good the correlation is between drug resistance and the studied phenotypes. Negative values mean that there is a negative correlation (when one value increases, the other decreases) and positive values mean that there is a positive correlation (when one value increases, the other increases too in a similar manner). The values can only be between -1 and 1, with the extreme values showing a perfect correlation. The question it is trying to answer is, do TDP2 protein level values (or the other two phenotype values) and drug IC50 concentrations co-vary, i.e. change in a co-dependent manner, when compared with the variance in TDP2 levels and drug resistance for the available cell lines? The drugs that produced at least one significant p-value

for any of the three phenotypes were shortlisted. These results are shown in Tables 7.1 and 7.2

for breast and lung cell lines respectively. The complete data for all drugs 113 for the breast

cancer cells and all 194 for the lung cancer cells can be found in Appendices 8 and 9,

respectively.

Drug	Original source	TDP2 r-value	TOP2A/ TDP2 r-value	TOP2B/ TDP2 r-value	TDP2 p-value	TOP2A/ TDP2 p-value	TDP2B/ TDP2 p-value
ABT888	COSMIC	-0.508	0.934	0.513	0.382	0.020	0.376
AEW541	CCLE	0.967	-0.571	-0.749	0.007	0.314	0.145
AKT inhibitor VIII	COSMIC	-0.859	0.434	0.896	0.062	0.466	0.039
AP24534	COSMIC	0.972	-0.059	-0.782	0.006	0.925	0.118
Axitinib	COSMIC	-0.113	0.984	0.062	0.856	0.002	0.921
AZD6482	COSMIC	-0.826	0.020	0.911	0.085	0.975	0.032
BIBW2992	COSMIC	-0.931	0.277	0.616	0.022	0.652	0.269
FTI277	COSMIC	-0.724	0.125	0.924	0.166	0.841	0.025
Nilotinib	COSMIC	-0.343	0.947	0.432	0.572	0.015	0.468
NU7441	COSMIC	-0.240	0.895	-0.014	0.697	0.040	0.982
Nutlin3	CCLE	0.643	-0.995	-0.922	0.242	0.000	0.026
PD0332991	COSMIC	0.050	0.901	-0.193	0.936	0.037	0.755
TKI258	CCLE	0.862	-0.932	-0.959	0.060	0.021	0.010
Vorinostat	COSMIC	-0.751	0.146	0.972	0.143	0.815	0.006
VX702	COSMIC	-0.601	0.892	0.467	0.283	0.042	0.428
ZM447439	COSMIC	-0.165	0.968	0.041	0.791	0.007	0.948

Table 7.1: Statistical analysis results of TDP2 protein levels, TOP2A/TDP2 protein ratio and TOP2B/TDP2 protein ratio and their correlation with sensitivity (IC50 values) to each curated anti-cancer drug in breast cancer cell lines.

Table 7.2: Statistical analysis results of TDP2 protein levels, TOP2A/TDP2 protein ratio and TOP2B/TDP2 protein ratio and their correlation with sensitivity (IC50 values) to each curated anti- cancer drug in lung cancer cell lines.

Drug	Original source	TDP2 r-value	TOP2A/ TDP2 r-value	TOP2B/ TDP2 r-value	TDP2 p-value	TOP2A/ TDP2 p-value	TDP2B/ TDP2 p-value
AG014699	COSMIC	-0.860	-0.324	-0.390	0.028	0.531	0.445
AKT inhibitor VIII	COSMIC	0.251	0.972	0.213	0.684	0.005	0.731
AS601245	COSMIC	0.358	0.607	0.949	0.554	0.277	0.014
Bicalutamide	COSMIC	0.960	0.539	0.469	0.010	0.349	0.426
Bleomycin	COSMIC	0.351	0.953	0.186	0.563	0.012	0.765
BMS708163	COSMIC	-0.852	-0.486	-0.497	0.031	0.329	0.316
BMS754807	COSMIC	-0.639	-0.239	-0.917	0.245	0.699	0.028
BX795	COSMIC	-0.920	-0.220	-0.272	0.009	0.675	0.602
Camptothecin	COSMIC	-0.919	-0.313	-0.403	0.010	0.546	0.428
CEP701	COSMIC	-0.931	-0.298	-0.327	0.007	0.567	0.527
CHIR99021	COSMIC	0.440	0.922	0.111	0.458	0.026	0.859
CI1040	COSMIC	-0.939	-0.371	-0.411	0.005	0.469	0.419
DMOG	COSMIC	0.361	0.906	0.172	0.551	0.034	0.782
FTI277	COSMIC	0.293	0.975	0.561	0.632	0.005	0.326
Nilotinib	COSMIC	0.310	0.054	-0.836	0.550	0.919	0.038
Obatoclax Mesylate	COSMIC	-0.983	-0.329	-0.418	0.003	0.589	0.484
PF562271	COSMIC	-0.064	0.918	0.183	0.918	0.028	0.768
QS11	COSMIC	-0.938	-0.288	-0.442	0.018	0.639	0.456
RDEA119	COSMIC	-0.924	-0.247	-0.292	0.008	0.637	0.575
SL01011	COSMIC	0.302	0.943	-0.049	0.561	0.005	0.927
Thapsigargin	COSMIC	-0.994	-0.356	-0.367	0.001	0.556	0.543
Tipifarnib	COSMIC	0.170	0.964	0.205	0.785	0.008	0.741
TW37	COSMIC	-0.873	-0.102	-0.171	0.023	0.848	0.746
Vinblastine	COSMIC	-0.889	-0.514	-0.562	0.018	0.297	0.246
AZD6482	COSMIC	-0.862	-0.203	-0.487	0.027	0.700	0.327

There is a large variety in the results of both r-values and p-values. When an r-value is close to either -1 or 1, it is highly likely that its p-value will be significant. p-values can also depend on other factors, such as the number of data points. For the breast cancer cell lines, the drugs that showed a significant (p-value < 0.05) correlation between how resistant the different cell lines were based on their TDP2 protein levels, were AEW541 (positive), AP24534 (positive) and

BIBW2992 (negative). Based on the breast cell lines' TOP2A/TDP2 protein ratios, the drugs that showed significant results were ABT888 (positive), Axitinib (positive), Nilotinib (positive), NU7441 (positive), Nutlin3 (negative), PD0332991 (positive), TKI258 (negative), VX702 (positive) and ZM447439 (positive). For the TOP2B/TDP2 protein ratios for the same type of cell lines, significant correlation results were found for AKT inhibitor VIII (positive), AZD6482 (positive), FTI277 (positive), Nutlin3 (negative), TKI258 (negative) and Vorinostat (positive).

For the lung cancer cell lines, the drugs that showed a significant correlation (p-value < 0.05) between how resistant the different cell lines were based on their TDP2 protein levels, were AG014699 (negative), Bicalutamide (positive), BMS708163 (negative), BX795 (negative), Camptothecin (negative), CEP701 (negative), CI1040 (negative), Obatoclax Mesylate (negative), QS11 (negative), RDEA119 (negative), Thapsigargin (negative), TW37 (negative), Vinblastine (negative) and AZD6482 (negative). The drugs that showed significant results for the TOP2A/TDP2 ratio phenotype were AKT inhibitor VIII (positive), Bleomycin (positive), CHIR99021 (positive), DMOG (positive), FTI277 (positive), PF562271 (positive), SL01011 (positive) and Tipifarnib (positive). The significant results for TOP2B/TDP2 ratios were AS601245 (positive), BMS754807 (negative) and Nilotinib (negative).

Figures 7.2 and 7.3 depict the scatter plots for the drugs listed in Tables 7.1 and 7.2 for breast and lung cancer cell lines respectively. The x axis shows the IC50 concentrations for each drug and the y axis each of the studied phenotypes (TDP2 levels and TOP2/TDP2 ratios). Next to each drug name, the original source of the information is also stated (COSMIC or CCLE). These scatter plots can be used to better visualise the results in Tables 7.1 and 7.2, which will be analysed in the next few paragraphs.

D

Drug: AP24534 Source: COSMIC

E

Drug: Axitinib Source: COSMIC

F

Drug: AZD6482 Source: COSMIC

Drug: TKI258 Source: CCLE

Drug: Vorinostat Source: COSMIC

Drug: VX702 Source: COSMIC
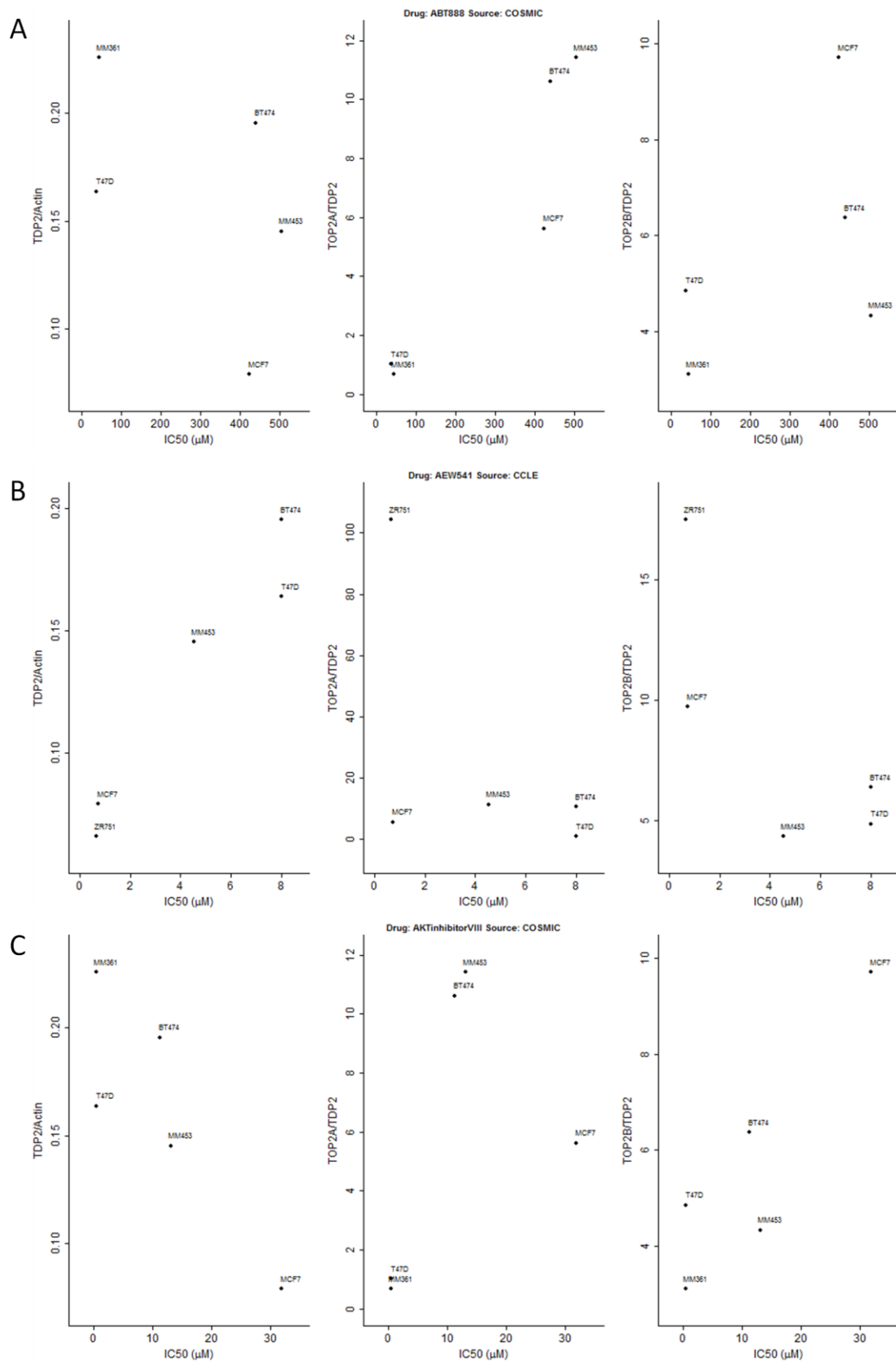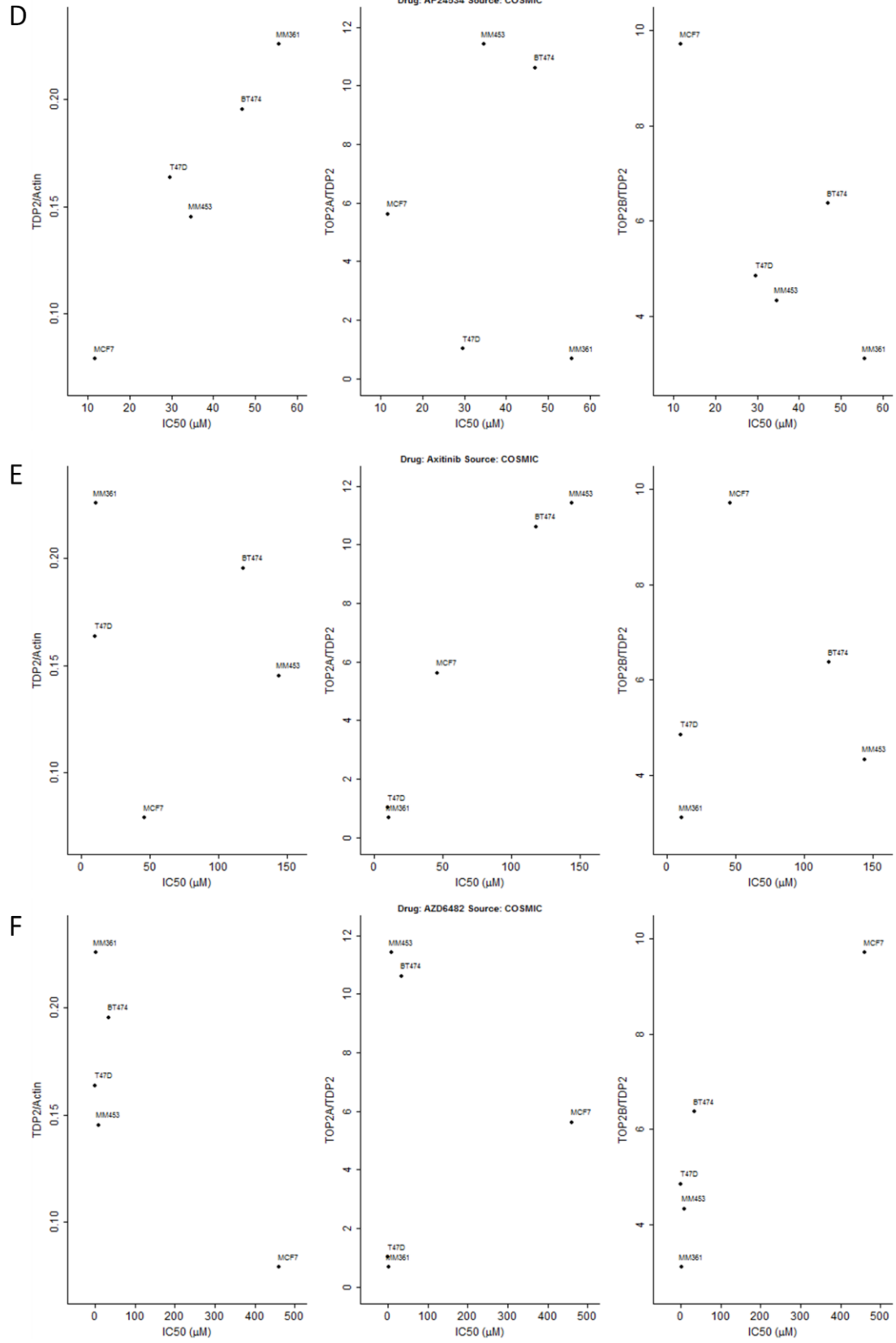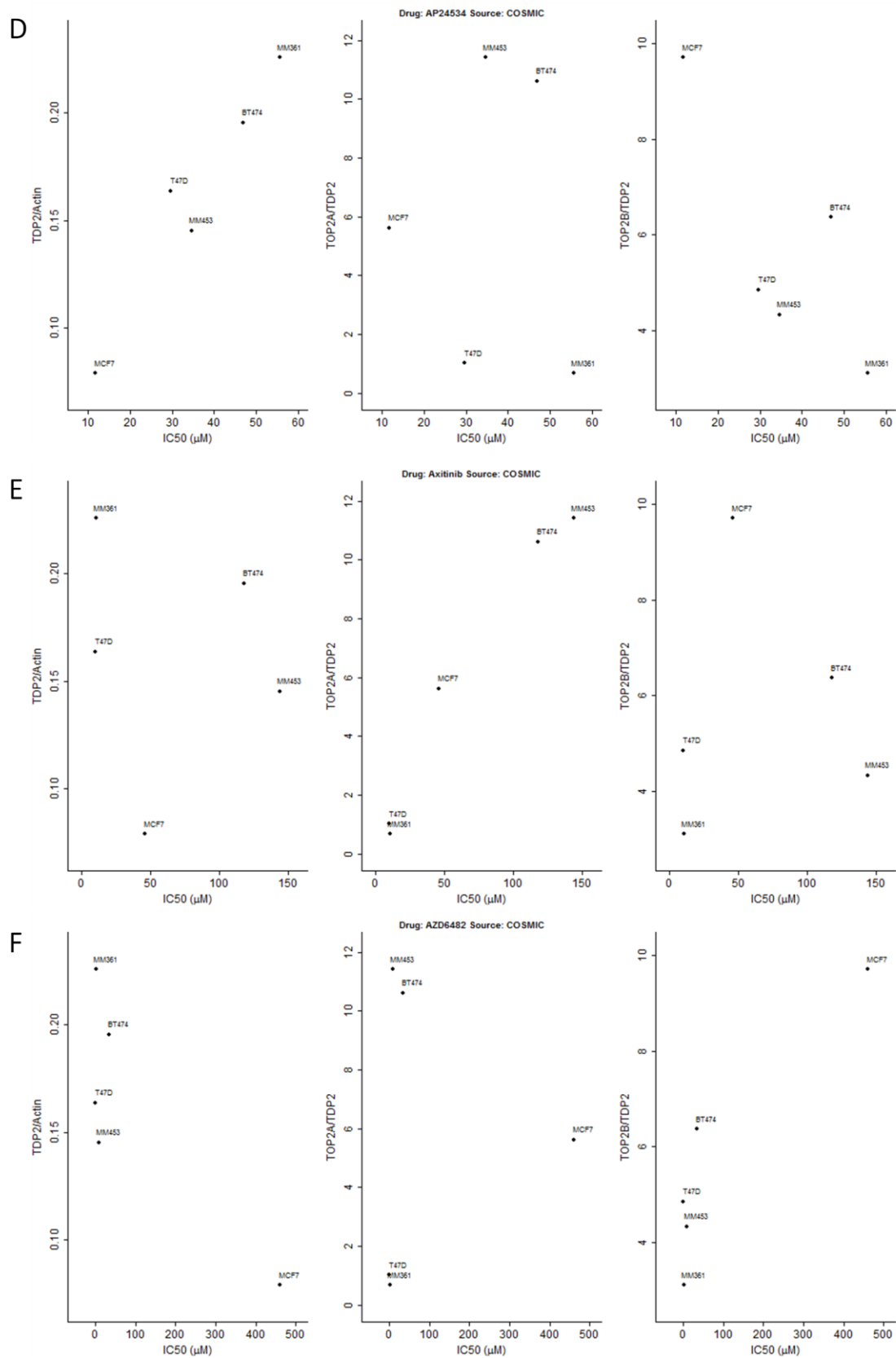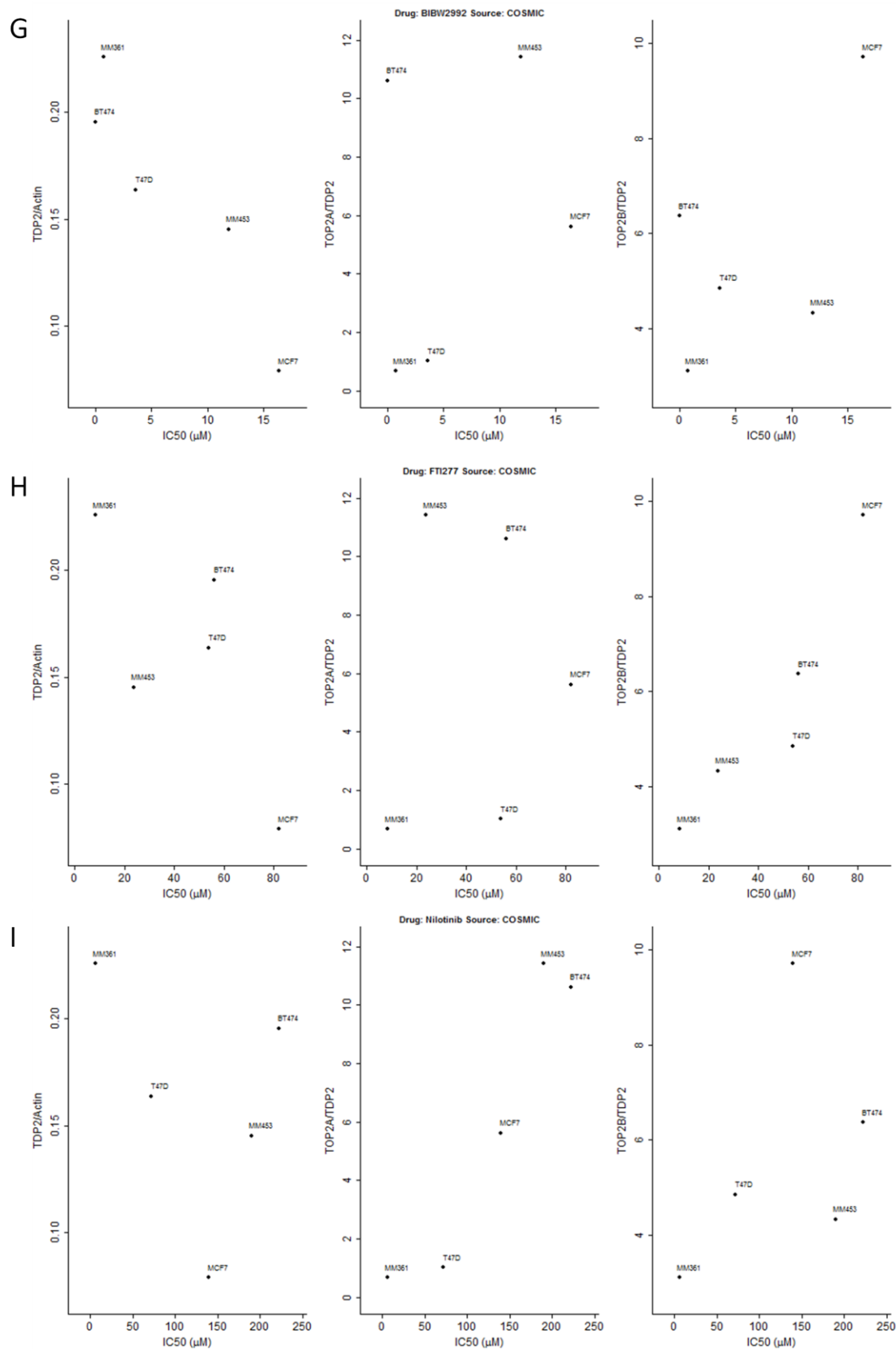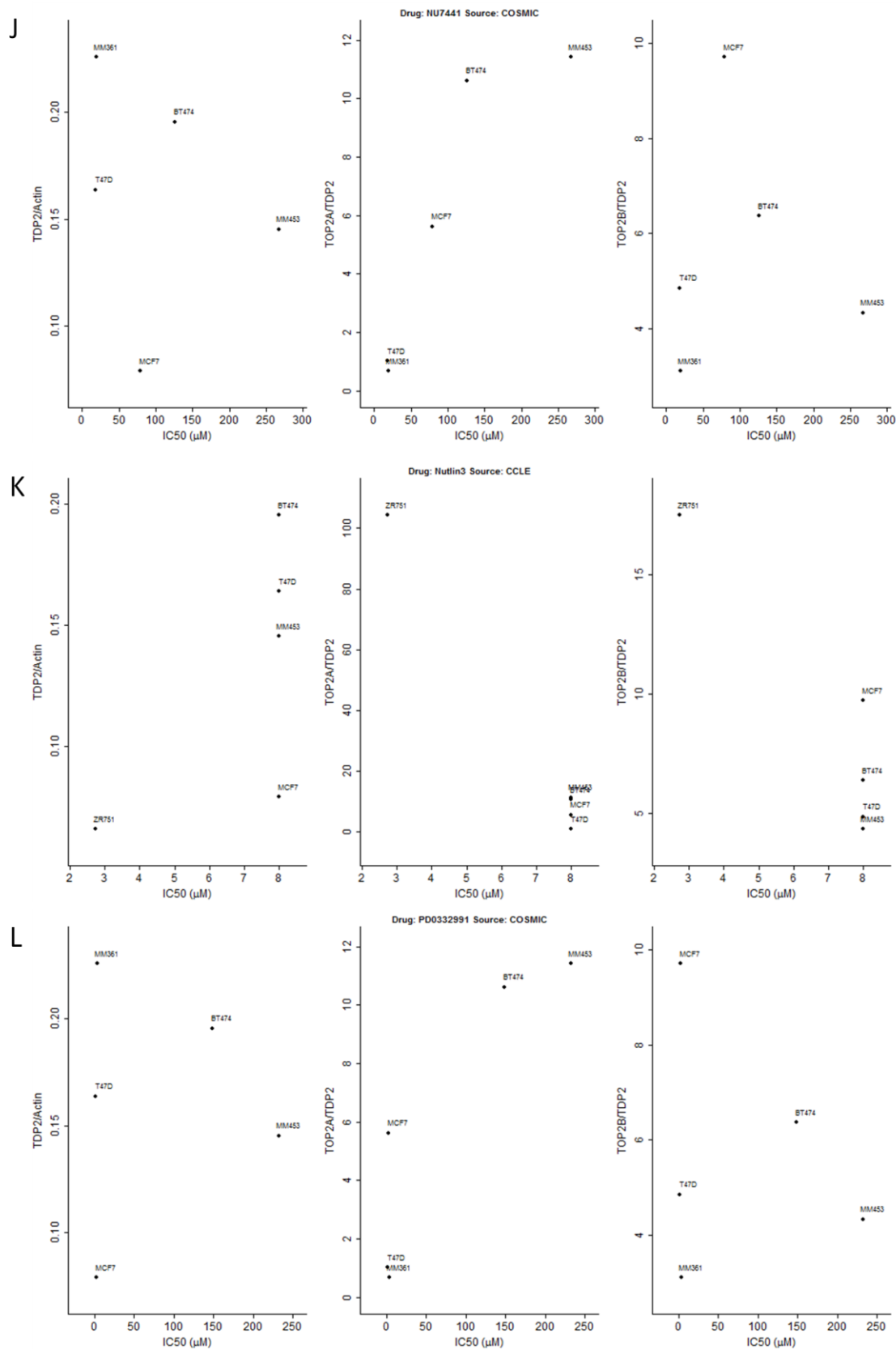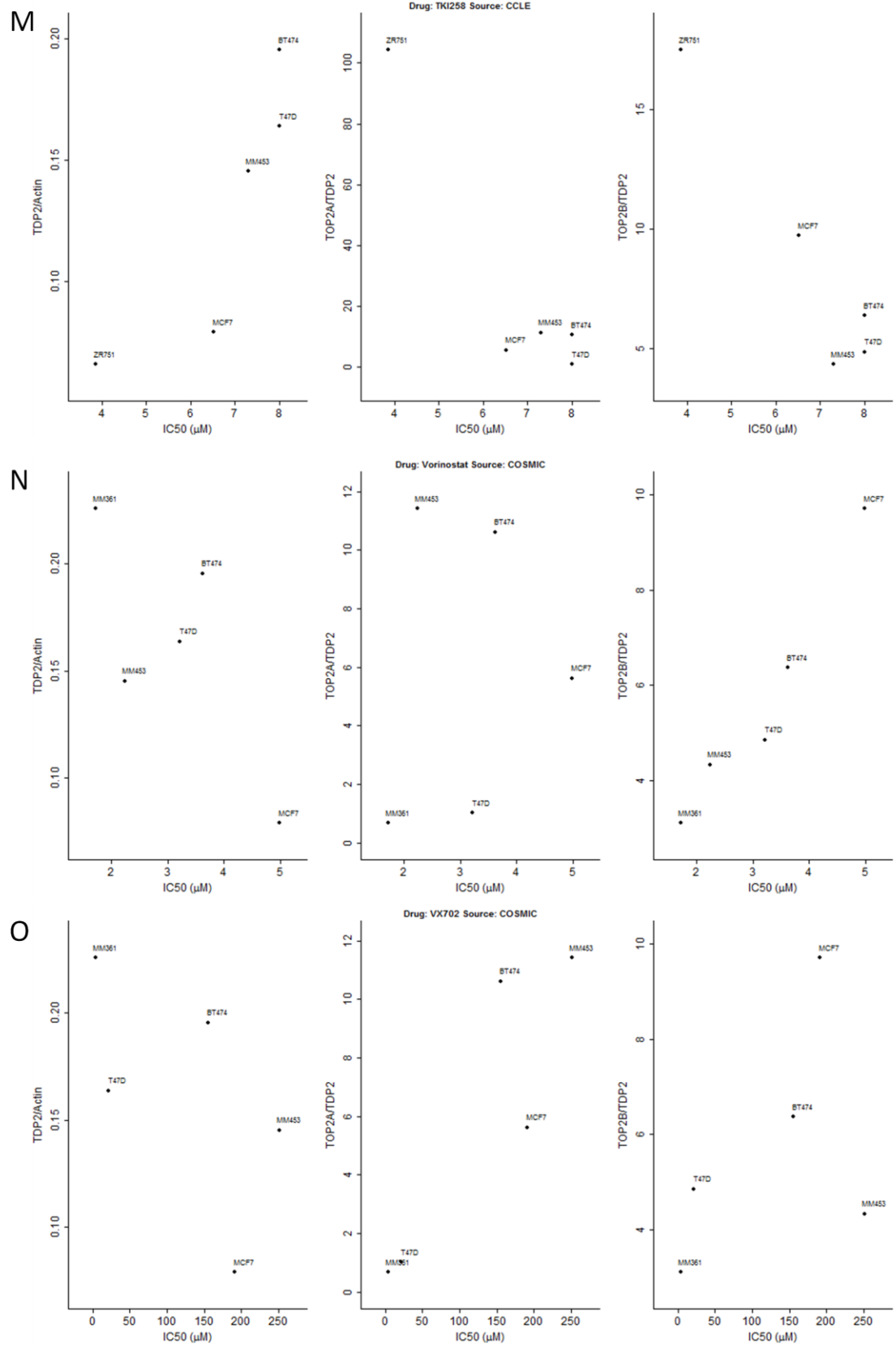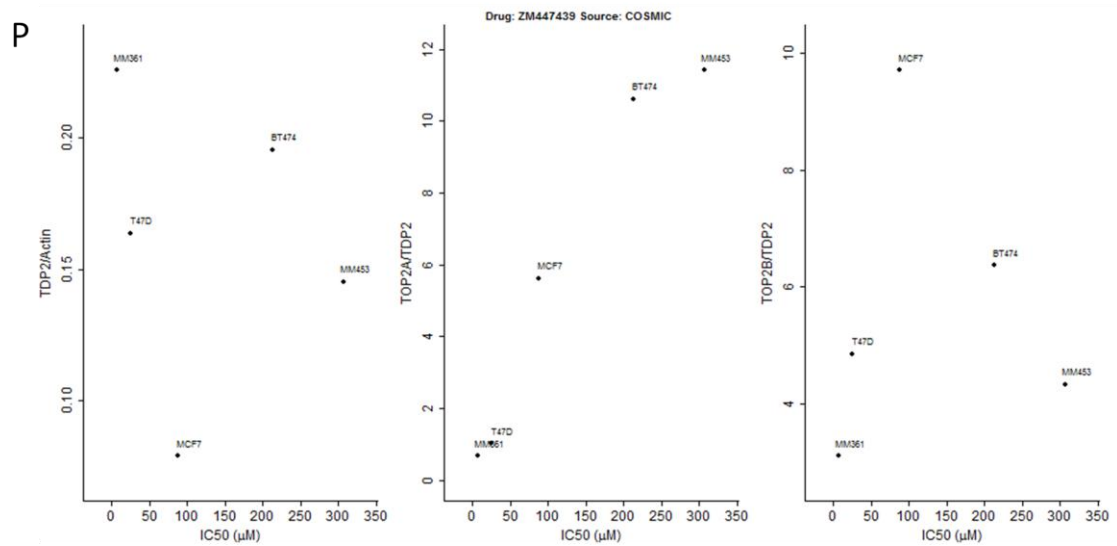
Figure 7.2: Breast cancer cell lines: Scatter plots for curated anti-cancer drugs depicting IC50 concentrations for each drug obtained from CancerDR database, plotted against the three experimentally studied phenotypes, TDP2 protein levels, TOP2A/TDP2 protein ratios and TOP2B/TDP2 protein ratios. Each set of plots also shows the original source of the database data. A) Drug: ABT888, Source: COSMIC. B) Drug: AEW541, Source: CCLE. C) Drug: AKT Inhibitor VIII, Source: COSMIC. D) Drug: AP24534, Source: COSMIC. E) Drug: Axitinib, Source: COSMIC. F) Drug: AZD6482, Source: COSMIC. G) Drug: BIBW2992, Source: COSMIC. H) Drug: FTI277, Source: COSMIC. I) Drug: Nilotinib, Source: COSMIC. J) Drug: NU7441, Source: COSMIC. K) Drug: Nutlin3, Source: CCLE. L) Drug: PD0332991, Source: COSMIC. M) Drug: TKI258, Source: CCLE. N) Drug: Vorinostat, Source: COSMIC. O) Drug: VX702, Source: COSMIC. P) Drug: ZM44739, Source: COSMIC.

Drug: AG014699 Source: COSMIC

Drug: AKTinhibitorVIII Source: COSMIC

Drug: AS601245 Source: COSMIC

G

Drug: BMS754807 Source: COSMIC

H

Drug: BX795 Source: COSMIC

I

Drug: Camptothecin Source: COSMIC

M

Drug: DMOG Source: COSMIC

N

Drug: FTI277 Source: COSMIC

O

Drug: Nilotinib Source: COSMIC

Drug: ObatoclaxMesylate Source: COSMIC

Drug: PF562271 Source: COSMIC

Drug: QS11 Source: COSMIC

Figure 7.3: Lung cancer cell lines: Scatter plots for curated anti-cancer drugs depicting IC50 concentrations for each drug obtained from CancerDR database, plotted against the three experimentally studied phenotypes, TDP2 protein levels, TOP2A/TDP2 protein ratios and TOP2B/TDP2 protein ratios. Each set of plots also shows the original source of the database data. A) Drug: AG014699, Source: COSMIC. B) Drug: AKT Inhibitor VIII, Source: COSMIC. C) Drug: AS601245, Source: COSMIC. D) Drug: Bicalutamide, Source: COSMIC. E) Drug: Belomycin, Source: COSMIC. F) Drug: BMS708163, Source: COSMIC. G) Drug: BMS754807, Source: COSMIC. H) Drug: BX795, Source: COSMIC. I) Drug: Camptothecin, Source: COSMIC. J) Drug: CEP701, Source: COSMIC. K) Drug: CHIR99021, Source: COSMIC. L) Drug: CI1040, Source: COSMIC. M) Drug: DMOG, Source: COSMIC. N) Drug: FTI277, Source: COSMIC. O) Drug: Nilotinib, Source: COSMIC. P) Drug: Obatoclax Mesylate, Source: COSMIC. Q) Drug: PF562271, Source: COSMIC. R) Drug: QS11, Source: COSMIC. S) Drug: RDEA119, Source: COSMIC. T) Drug: SL01011, Source: COSMIC. U) Drug: Thapsigargin, Source: COSMIC. V) Drug: Tipifarnib, Source: COSMIC. W) Drug: TW37, Source: COSMIC. X) Drug: Vinblastine, Source: COSMIC. Y) Drug: AZD6482, Source:

The drug targets, alternative names, their targets and their present use can be found in Table 7.3.

Table 7.3: A list of all the curated drugs discussed in this chapter with their alternative names, targets and present use.

Drug	Also known as	Target	Use	References
ABT888	Veliparib	PARP1/2	Phase 1/2/3 clinical trials for cancer therapy	(Donawho et al., 2007)
AEW541		IGF-IR	Commercially available for laboratory use as inhibitor	(García-Echeverría et al., 2004)
AG014699	Rucaparib	PARP	Phase 1/2/3 clinical trials for cancer therapy	(White et al., 2000)
AKT inhibitor VIII		Akt1/2	Commercially available for laboratory use as inhibitor	(Lindsley et al., 2005)
AP24534	Ponatinib	BCR-ABL wt/mt	Cancer therapy	(O'Hare et al., 2009)
AS601245		JNK	Commercially available for laboratory use as inhibitor	(Carboni et al., 2004)
Axitinib		VEGFR 1/2/3, PDGFRβ, c-Kit	Cancer therapy	(Rugo et al., 2005)
AZD6482		PI3Kβ	Has been in clinical trials for its antiplatelet effect	(Nylander et al., 2012)
BIBW2992	Afatinib	EGFR, HER2/4	Cancer therapy	(Reid et al., 2007)
Bicalutamide	ICI 176,334	Androgen receptor	Cancer therapy	(Furr et al., 1987)
Bleomycin		DNA	Cancer therapy	(Umezawa, 1965)
BMS708163		γ-Secretase	Has been in clinical trials for Alzheimer's	(Gillman et al., 2010)
BMS754807		IGF-IR	Has been in clinical trials for cancer	(Wittman et al., 2009)
BX795		TBK1, IKKε, PDK1	Commercially available for laboratory use as inhibitor	(Clark et al., 2009)
Camptothecin	Topotecan, Irinotecan	TOP1	Derivatives used in cancer therapy	(Wall et al., 1966)
CEP701	KT-5555	FLT3, JAK2, TrkA/B/C	Has been in clinical trials for cancer	(Miknyoczki et al., 1999)

CHIR99021		GSK-3	Commercially available for laboratory use as inhibitor	(Tighe et al., 2007)
CI1040	PD184352	MAPK	Commercially available for laboratory use as inhibitor	(Sebolt-Leopold et al., 1999)
DMOG		Prolyl 4-hydroxylase	Commercially available for laboratory use as inhibitor	(Asikainen et al., 2005)
FTI277		Farnesyltranserase	Commercially available for laboratory use as inhibitor	(Lerner et al., 1995)
Nilotinib	AMN107	BCR-ABL, c-kit, PDGF	Cancer therapy	(Weisberg et al., 2005)
NU7441		DNA-PK	Commercially available for laboratory use as inhibitor	(Leahy et al., 2004)
Nutlin3		MDM2-p53 interaction	Commercially available for laboratory use as inhibitor	(Vassilev et al., 2004)
Obatoclax Mesylate		Bcl-2	Has been in clinical trials for cancer	(Nguyen et al., 2007)
PD0332991	Palbociclib	CDK4/6	Cancer therapy	(Fry et al., 2004)
PF562271		FAK	Commercially available for laboratory use as inhibitor	(Roberts et al., 2008)
QS11		ARFGAP1, Wntβ-catenin pathway	Commercially available for laboratory use as inhibitor	(Zhang et al., 2007)
RDEA119	Refametinib, Bay 86-9766	MEK1/2	Commercially available for laboratory use as inhibitor	(Iverson et al., 2009)
SL01011		RSK	Commercially available for laboratory use as inhibitor	(Smith et al., 2005)
Thapsigargin		SERCA	Commercially available for laboratory use as inhibitor	(Hakii et al., 1986)

Tipifarnib		Farnesyltransferase	Phase 1/2/3 clinical trials for cancer therapy	(End et al., 2001)
TKI258	CHIR-258, Dovitinib	Multiple kinases	Phase 1/2/3 clinical trials for cancer therapy	(Trudel et al., 2005)
TW37		Bcl-2	Commercially available for laboratory use as inhibitor	(Zeitlin et al., 2006)
Vinblastine		Tubulin	Cancer therapy	(Jordan and Wilson, 2004)
Vorinostat		Histone kinases	Cancer therapy	(Marks and Breslow, 2007)
VX702		p38 MAPKs	Has been in clinical trials for rheumatoid arthritis	(Kuliopulos et al., 2004)
ZM447439		Aurora B	Commercially available for laboratory use as inhibitor	(Ditchfield et al., 2003)

7.3 Discussion

The drugs that I will be focusing on in the discussion are drugs currently in use for cancer therapy, or are presently in clinical trials for their potential use in cancer therapy. A summary of these drugs and the correlation (positive or negative) of TDP2 protein levels, TOP2A/TDP2 protein ratios and TOP2B/TDP2 protein ratios of the breast and lung cancer cell lines used in this chapter, with their resistance to these drugs (database data of IC50 values) can be seen in the following Table (Table 7.4).

Table 7.4: Relevant drugs and the correlation of different phenotypes of breast and lung cancer cell lines used in this chapter, with resistance to these drugs (database data).

Phenotype	Breast cancer cell lines		Lung cancer cell lines	
	Positive correlation	Negative correlation	Positive correlation	Negative correlation
TDP2 protein levels	AP24534	BIBW2992	Bicalutamide	AG014699
				Camptothecin
				Vinblastine
TOP2A/TDP2 protein ratios	ABT888	TKI258	Bleomycin	
	Axitinib		Tipifarnib	
	Nilotinib			
	PD0332991			
TOP2B/TDP2 protein ratios	Vorinostat	TKI258		Nilotinib

Anti-cancer drugs generally liaise with the cell cycle stages and fall under multiple categories, some of which are relevant to DNA damage and repair. Given how precious research funding is, a hypothesis is worth exploring when it is based on sound reasoning. My analysis looked at a large number of drugs with a variety of targets and thus some of the results, even though their IC50 values' correlation was statistically significant with the phenotypes studied, could be false positives. That means that the phenotypes might not actually be affecting the IC50 concentrations of these drugs found in the available cell lines. Drugs that fall under this category are AP24534, that targets Bcr-Abl, a tyrosine kinase; BIBW2992, whose targets EGFR and HER2 are also tyrosine kinases; Axitinib, whose targets VEGFR 1/2/3, PDGFRβ and c-Kit are

receptor protein kinases; PD0332991, that targets CDK4/6, both cyclin-dependent kinases; Vorinostat, whose targets are histone kinases; and finally Tipifarnib, that targets farnesyltransferase, an enzyme protein that modifies Ras proteins, which play a role in cellular signal transduction.

TKI258 produced two significant correlations, unlike other drugs whose target does not seem relevant to TDP2, which is why I do not include it in that group. TKI258, which targets receptor tyrosine kinases that participate directly or indirectly in tumour growth, survival, angiogenesis, and vascular development, is currently in clinical trials for different types of cancer, including breast and NSCLC (Lee et al., 2015a). In the breast cancer cell lines, I found a negative correlation between resistance to TKI258 and both TOP2A/TDP2 and TOP2B/TDP2 protein ratios. It is interesting, because TKI258 has been shown to also target TOP2A by stabilizing the enzyme-cleavage complex and inducing DSBs (Hasinoff et al., 2012). This makes the correlation I saw very interesting and worth of further exploration *in vivo*.

The only drug that both types of cell lines have in common is Nilotinib. Nilotinib is used as treatment for chronic myeloid leukaemia and has been designed to increase its affinity and specificity for the oncogenic tyrosine kinase Bcr-Abl (Ostendorf et al., 2014). Tyrosine kinases are responsible for the activation of many proteins by signal transduction cascades. Patients with chronic myeloid leukaemia have shown resistance to TOP2 inhibitors due to low TOP2 expression (Valkov and Sullivan, 1997, Silber et al., 1989). I was not able to find further data about whether this meant TOP2A or TOP2B. Even though Nilotinib has been identified as having significant results for both types of cell lines, it is not used for treatment of breast or lung cancers. I believe that it would be interesting to further explore TDP2, TOP2A and TOP2B protein levels and Nilotinib resistance, as for the lung cancer cell lines I found that high TOP2B/TDP2 ratios mean lower resistance to Nilotinib. It is likely that TDP2 does not play a role, as a significant correlation was not found.

Vinblastine is a mitotic inhibitor that targets tubulins, which polymerise into microtubules, a major component of the cytoskeleton of eukaryotic cells. It has been shown that Vinblastine along with etoposide induced the formation of apoptotic TOP1 cleavage complexes (Sordet et al., 2006). TDP2 has been shown to promote the repair of TOP1-mediated DNA damage in TDP1 knock-out cells (Zeng et al., 2012). But I propose that the negative correlation between resistance to Viblastine and TDP2 protein levels that I found in the panel of lung cancer cell lines, can be explained by the increased proliferation in NSCLC observed when TDP2 is overexpressed (Li et al., 2011a). As high levels of TDP2 mean that there is increased proliferation, that also means that there is increased mitosis too, which is the target of Vinblastine. Therefore, its target binding-target tubulin is available in larger numbers, making the drug more effective. This is a very interesting and unexpected correlation, which merits further investigation.

ABT888 and AG014688 are both drugs that target PARP1 and are currently in multiple clinical trials at different stages, varying from stage 1 to stage 3 and for different types of cancer. The importance of PARP1 has been discussed in the introduction of this thesis extensively, as well as the upsurge of interest in PARP inhibitors. AG014688 was the first PARP inhibitor to enter clinical trials (Plummer et al., 2008). It is the phosphate salt of AG14447 and has improved aqueous solubility, but this name might be of use if one is interested further in this inhibitor's background (Thomas et al., 2007). In lung cancer cell lines, I found a negative correlation between AG014688 resistance and TDP2 protein levels, which means that when TDP2 levels decrease there is higher resistance to AG014688. In breast cancer cells, I saw that when TOP2A/TDP2 protein ratios increase, then the resistance to ABT888 increases, which could be attributed to high TOP2A levels or low TDP2 levels. It would be very interesting to further investigate this result in a larger panel of cell lines. It has been shown that TOP2B associates with Ku70 and PARP1 during DSBR in primary neurons after peroxide mediated damage; DSBs in neurons follow two repair pathways, a Ku70 dependent and a PARP1 dependent (Mandraju

et al., 2011). Activation of gene transcription has been shown to involve TOP2B-mediated DSBs and following that PARP1 enzymatic activity (Ju et al., 2006, Ju and Rosenfeld, 2006). In addition, another group has shown that inhibition of PARP1 with the NU1025 inhibitor sensitized HeLa cells to the TOP2 inhibitor C-1305 (Sabisz et al., 2010). There is evidence that alludes to an interesting relationship between PARP1 and TOP2, and since TOP2 is so closely related to TDP2, I believe there is a possibility for TDP2 or its ratio to TOP2 to be a candidate for sensitivity to PARP1 inhibitors.

Bleomycin was first characterised as an antibiotic, but due to its cytotoxic effects it is used as an anti-cancer drug. It works by binding to the DNA and introducing DSBs and SSBs, and their ratio in addition to absolute values, can determine which pathway will be followed for cell death: apoptosis, mitotic cell death, or pseudoapoptosis (Tounekti et al., 2001). It has also been shown that Bleomycin resistance is characterised by reduced DNA damage, shown via COMET and γH2AX assays in resistant cell lines (Wang et al., 2013). My results showed that TOP2A/TDP2 protein ratios had a positive correlation with Bleomycin resistance and given the role of both of these proteins in DSBR, Bleomycin could be a candidate for further exploring how its effect on cancer cells can be affected by TDP2 and TOP2.

Camptothecin is a TOP1 inhibitor; TDP2-depleted cells have shown lack of hypersensitivity to it, compared to etoposide (Zeng et al., 2011). Further studies that look into intrinsic levels of TDP2 have not been published and as TDP2 has been shown to step in when TDP1 is absent and promote the repair of TOP1-mediated DNA damage in its place, it is not unlikely that the protein levels of TDP2 might have an effect on how a cancer cell will respond to treatment with camptothecin (Zeng et al., 2012). Therefore, it might be interesting to further expand on my results showing that high TDP2 protein levels had a negative impact on camptothecin resistance.

Bicalutamide, a non-steroidal anti-androgen, is a selective antagonist of the androgen receptor, preventing androgens from binding on and activating the androgen receptor; it is used primarily for the treatment of prostate cancer, as androgen receptor signalling is essential to the growth and survival of prostate cancer (Singh et al., 2000). It has been shown that androgen-deprived prostate cancer cells display transient TOP2B-dependent DSBs when exposed to androgen; the same group also showed that exposing androgen-deprived prostate cancer cells to Bicalutamide lead to increased DSB γH2AX foci formation (Coulter et al., 2015). In this chapter, I showed that there is a positive correlation between resistance to Bicalutamide and TDP2 protein levels. Given the involvement of TDP2 in resolving TOP2-mediated DSBs, I believe that it merits further exploration *in vivo*.

To conclude, I propose that TDP2 could be investigated as a possible biomarker of sensitivity for other anti-cancer drugs in a similar manner to my investigation of etoposide, as preliminary data show significant correlations of sensitivity and protein levels. Such research projects though entail a relatively high risk, as they are not mechanistic and therefore additional preliminary checks should be carried out. I would propose inclining towards investing in cell lines that are relevant to the types of cancer that these drugs are used for. But even before proceeding to obtain such cell lines for experimental data, they should be scrutinised for whether extensive information is available for their profile; for example, are IC50 values for these drugs available and do they vary? Have other researchers used those cell lines for similar models? And are their mutation profiles rich with information? These are important questions that need to be answered to ensure that risk in minimised.

8. CHAPTER EIGHT - Impact of the co-operative effects of etoposide, estradiol and PARP1 inhibitor on the survival of breast cancer cell lines

8.1 Introduction

In Chapter 5, a comparison of TDP2 protein levels and cell survival failed to show a correlation between TDP2 protein levels and sensitivity to etoposide, in neither breast nor lung cancer cell lines. In this chapter, I examined the co-operative effects of etoposide, estradiol and PARP1 inhibitor on the survival of breast cancer cell lines.

Estradiol, as mentioned in the introduction, is one of the most common estrogen hormones and is sometimes used as part of hormone cancer therapy in both women and men, especially breast cancer (Ingle, 2002). In addition, it has been implicated in higher breast and ovarian cancer risk in women that have been prescribed estradiol for menopause (Beral, 2003, Beral et al., 2007). My interest in estradiol reflects the ability of this hormone to induce TOP2-dependent transcription programmes, which in turn are predicted to generate TOP2-dependent DSBs. This may render breast cancer cells more sensitive to loss or inhibition of TDP2, either as a synthetic lethal effect or in conjunction with etoposide treatment. It has been previously shown that for successful signal-dependent activation of transcription by the estrogen receptor (ER), TOP2B needs to be recruited with the ER to regulatory sites on target genes in order for it to generate transient TOP2B-mediated DSBs (Haffner et al., 2011). Furthermore, there is evidence that shows PARP1 recruitment with TOP2B after treatment with estradiol (Ju et al., 2006). Inhibition of PARP1 decreased ERα binding to estrogen response element (short DNA sequence within ERα promoter) and prevented ERα-dependent gene transcription (Zhang et al., 2013).

PARP1 is a protein that is crucial for the repair of SSBs. If such breaks persist unrepaired until the DNA is replicated, then DSBs can form (Rouleau et al., 2010). Drugs that work as PARP1

inhibitors can induce the formation of such DSBs by interfering with the repair of SSBs. In addition, it has been shown that tumours experiencing acute or chronic hypoxia (such as solid or fast growing tumours) are sensitive to PARP1 inhibitors (Chan et al., 2010). Furthermore, there is data which suggests that certain types of breast cancer might be sensitive to PARP1 inhibitors (Weil and Chen, 2011). Whilst the generation of SSBs is part of TOP1's activity, the two subunits of TOP2 introduce DNA breaks in an independent but coordinated process and therefore can be trapped as a SSB if the activity of TOP2 on both strands is not coordinated, which normally would be a target for PARP1 (Zechiedrich et al., 1989, Nitiss, 2009b). Whether SSBs or DSBs will result from etoposide treatment has been shown to depend on the molar ratio between etoposide and TOP2 (Bromberg et al., 2003). PARP1 has also been shown to promote repair by HR by suppressing components of NHEJ, the pathway that TDP2-resolved cleavage complexes of TOP2 follow for their repair (Saberi et al., 2007, Gomez-Herreros et al., 2013). PARP1 is also required for replication fork slowing on damaged DNA (Sugimura et al., 2008). Combining both a PARP1 inhibitor and etoposide could result in an overwhelmed DNA repair system, with accumulating damage leading to apoptosis.

PARP1 inhibitors are very popular at the moment in clinical studies and a quick search in clinicaltrials.gov, the U.S. National Institutes of Health depository for past and current clinical trials in 194 countries, will return over 200 results of studies in different phases, for different types of cancer and different PARP1 drugs (either alone or in combination with other drugs). In December 2014, the FDA and the EMA approved Olaparib, a PARP1 inhibitor as monotherapy for the treatment of recurrent ovarian cancer in women with a BRCA1 or BRCA2 mutation (Frampton, 2015). Not only that, but there is currently a trial (NCT01642251) that is testing a combination of the PARP1 inhibitor veliparib, cisplatin and etoposide for extensive stage small-cell lung cancer, metastatic large cell neuroendocrine NSCLC and small-cell carcinoma of unknown primary or extrapulmonary origin. This trial is at Phase II and due to conclude in

January 2018, demonstrating that combination therapy targeting both PARP1 and TOP2 simultaneously shows promise (Hosoya and Miyagawa, 2014).

Combining a PARP1 inhibitor with estradiol and etoposide is hypothesised to have a big impact on the cancer cells' ability to repair the estradiol-induced abortive DSBs. The cancer cells would have to repair TOP2-mediated DSBs caused by induction of transcription with estradiol, and DSBs created by unrepaired SSBs due to the PARP1 inhibitor or SSBs that formed from uncoordinated function of TOP2. All of this DNA damage will have been rendered abortive with the addition of etoposide at a frequency that TDP2 cannot cope with, thus inducing apoptosis and killing the cells.

As the precise impact of hormone treatment on TOP2-dependent transcription and DSB induction in breast cancer is not clear, I adopted the approach of selecting breast cancer cell lines representative of several clinically relevant scenarios. For example, MCF7 is one of the most commonly used breast cancer cell lines in experiments and appears in many published papers (Comsa et al., 2015). It is also ERα positive and grows in an estrogen-dependent manner. That means that types of cancer with the MCF7 profile usually respond well to inhibitors of ERα activity or drugs that block estrogen synthesis. MLET5 is an MCF7 derivative, which grows in an estrogen-independent manner despite being ERα positive. It was established by over-expressing ERα in MCF7 cells using adenovirus gene transduction, which resulted in ligand-independent expression of pS2 and PR (estrogen-regulated genes) and the cell line to be able to grow in an estrogen-independent manner (Tolhurst et al., 2011). Similarly, LCC9 is an MCF7 derivative that grows in an estrogen-independent manner, but it was acquired in a different way. Brunner et al. had previously established an MCF7 derivative, LCC1, which is able to grow *in vivo* and *in vitro* independent of estrogen, but still remains sensitive to antiestrogens by cultural selection for resistance (Brunner et al., 1993a). The same group was also able to grow an MCF7/LCC1 derivative, LCC2, which in addition to estrogen

independent growth, is also resistant to 4OH-TAM, again via cultural selection for resistance, but was not cross-resistant to the steroidal anti-estrogens ICI 182,780 and ICI 164,384 (Brunner et al., 1993b). ICI 182,780 and ICI 164,384 are C7-acyl-substituted analogues of estradiol and have a higher affinity for ER than TAM (Wakeling et al., 1991, Thompson et al., 1989). Therefore, a few years later, Brunner et al published a paper in which they reveal another MCF7/LCC1 derivative, LCC9, which was selected against ICI 182,780, but through that it also acquired cross-resistance for TAM (Brünner et al., 1997).

8.1.1 Aims of this chapter

The aim of this chapter is to investigate the effects of a combination treatment of PARP1 inhibitor, estradiol and etoposide on cell viability and survival in a selection of breast cancer cell lines relevant to the clinical situation (MCF7, MLET5 and LCC9). It is hypothesised that initial treatment with PARP1 inhibitor will induce DSBs due to the unrepaired SSBs caused by the inability of the inhibited PARP1 to repair them. The addition of estradiol will induce transcription and thus TOP2B-mediated DSBs. Those DSBs will be rendered abortive with the addition of etoposide, which will target TOP2 and interfere with the re-ligation of the DNA. Finally, TDP2 will not be able to cope with all these breaks, the cells' DNA damage response system will be overwhelmed and the cells will enter apoptosis. Etoposide, olaparib and estradiol are already used in cancer therapy. Therefore, it is a very attractive idea that if the results of this chapter are positive, it could be converted into a proposal for a wider preclinical trial.

8.2 Results

The effect of the combination treatment of PARP1 inhibitor, estradiol and etoposide was assessed using cell viability (Alamar Blue) and clonogenic cell survival assays. The results for the Alamar Blue assays for MCF7 are depicted in Figure 8.1. Comparing the different treatments in Figures 8.1B-D, it is clear than none of the combinations increased the sensitivity

of MCF7 cells above that observed with etoposide alone (Fig.8.1A). Similar results were observed for MLET5 (Fig.8.2) and LCC9 (Fig.8.3).
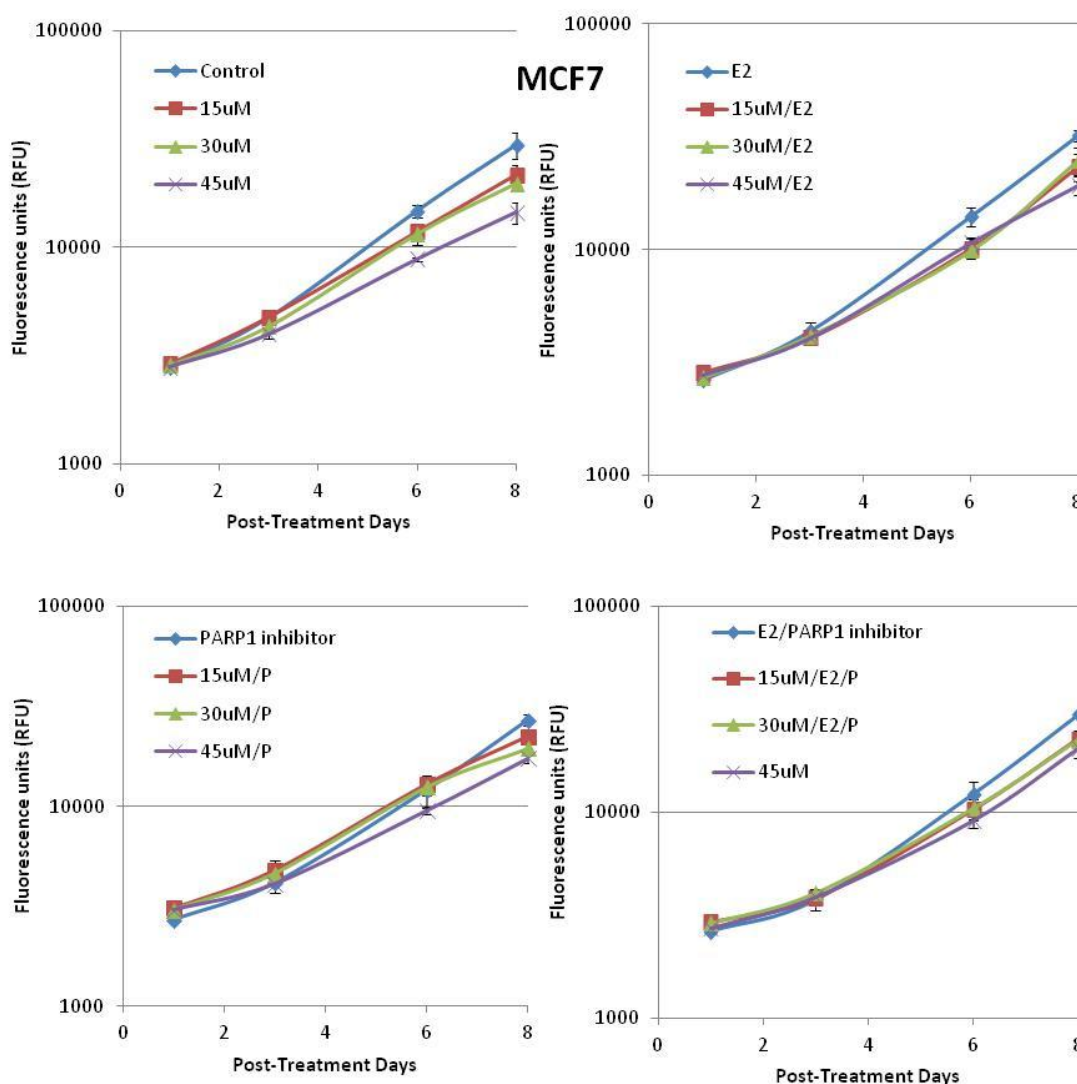
Figure 8.1: Comparison of sensitivity to etoposide, estradiol and etoposide, PARP1 inhibitor and etoposide or estradiol, PARP1 inhibitor and etoposide in breast cancer cell line MCF7. Sensitivity was measured using Alamar Blue assay in the indicated breast cancer cell line. Etoposide concentrations used were 0nM, 15μM, 30μM and 45μM. Estradiol (E2) concentration used was 100nM. PARP1 inhibitor (P) concentration used was 500nM. Cells were treated for 30min with PARP1 inhibitor or DMSO, then 3h with estradiol or EtOH and 3h with etoposide or DMSO. Fluorescence measurements were taken at days 1, 3, 6 and 8 after treatment was initiated. These measurements were plotted against the corresponding day. Data are the mean (+/- SEM) from three independent experiments.
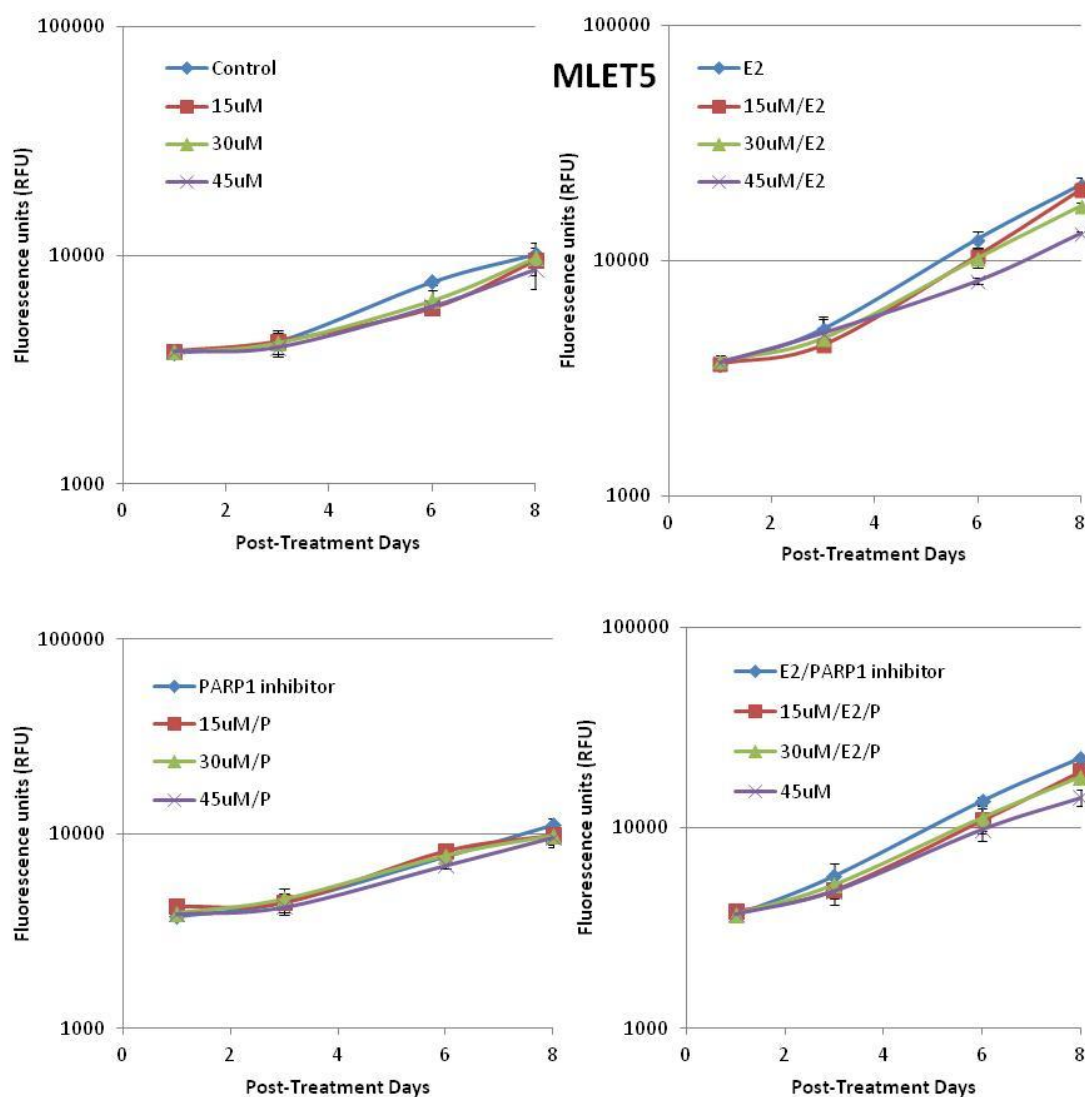
Figure 8.2: Comparison of sensitivity to etoposide, estradiol and etoposide, PARP1 inhibitor and etoposide or estradiol, PARP1 inhibitor and etoposide in breast cancer cell line MLET5. Sensitivity was measured using Alamar Blue assay in the indicated breast cancer cell line. Etoposide concentrations used were 0nM, 15μM, 30μM and 45μM. Estradiol (E2) concentration used was 100nM. PARP1 inhibitor (P) concentration used was 500nM. Cells were treated for 30min with PARP1 inhibitor or DMSO, then 3h with estradiol or EtOH and 3h with etoposide or DMSO. Fluorescence measurements were taken at days 1, 3, 6 and 8 after treatment was initiated. These measurements were plotted against the corresponding day. Data are the mean (+/- SEM) from three independent experiments.
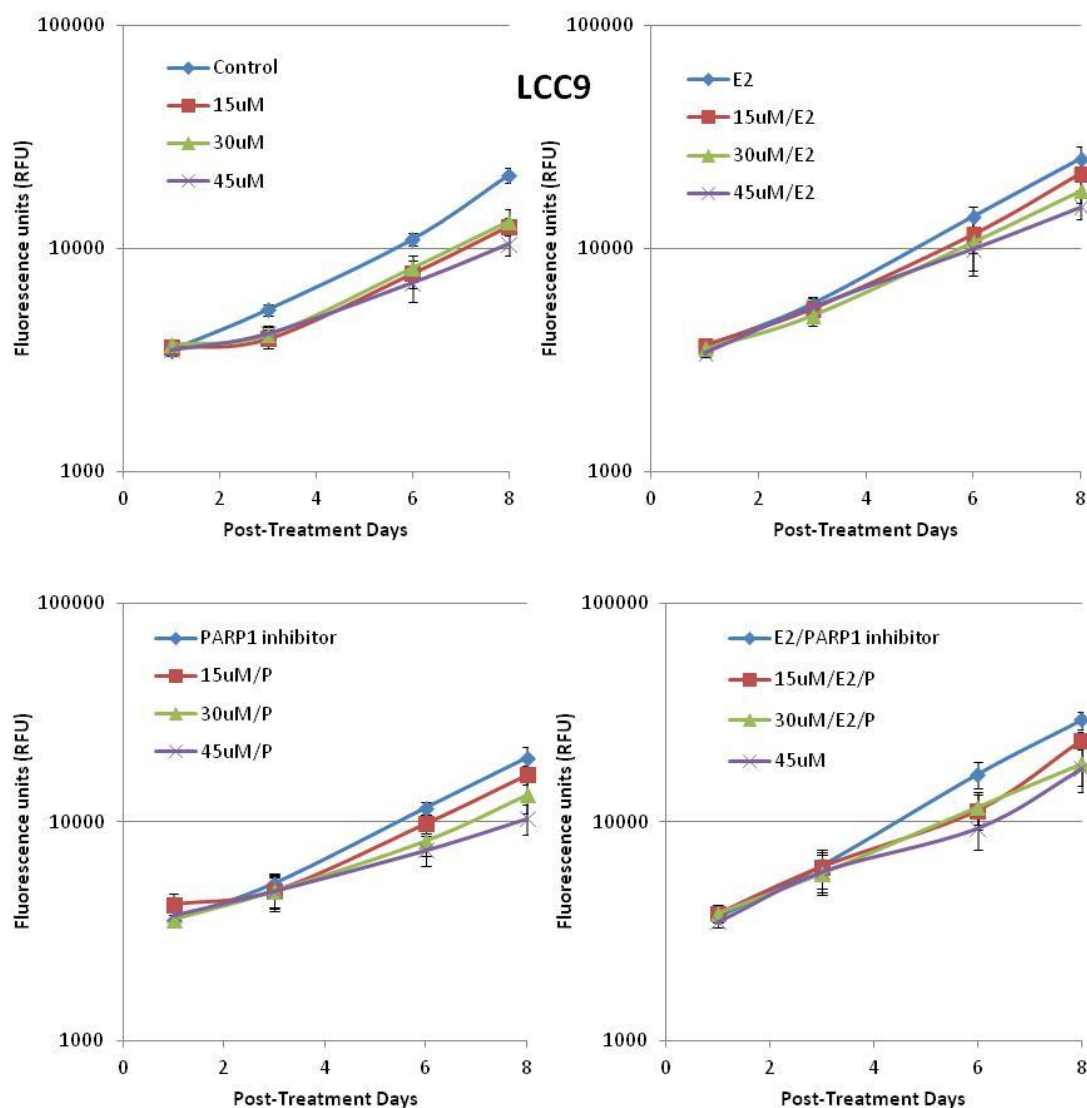
Figure 8.3: Comparison of sensitivity to etoposide, estradiol and etoposide, PARP1 inhibitor and etoposide or estradiol, PARP1 inhibitor and etoposide in breast cancer cell line LCC9. Sensitivity was measured using Alamar Blue assay in the indicated breast cancer cell line. Etoposide concentrations used were 0nM, 15µM, 30µM and 45µM. Estradiol (E2) concentration used was 100nM. PARP1 inhibitor (P) concentration used was 500nM. Cells were treated for 30min with PARP1 inhibitor or DMSO, then 3h with estradiol or EtOH and 3h with etoposide or DMSO. Fluorescence measurements were taken at days 1, 3, 6 and 8 after treatment was initiated. These measurements were plotted against the corresponding day. Data are the mean (+/- SEM) from three independent experiments.

As previously discussed, Alamar Blue has significant disadvantages when compared to clonogenic cell survival. Although Alamar Blue was the only option for MLET5 and LCC9 cells, which cannot grow colonies from single cells as they require close contact with each other to multiply, I was able to repeat these experiments for MCF7 using clonogenic cell survival assays. Figure 8.4A shows the impact of etoposide alone or in combination with PARP1 inhibitor and estradiol. As one can observe, the combination did not further sensitise MCF7 cells to etoposide. Even though there appears to be no difference between cell survival for cells treated with etoposide only and the combination treatment of etoposide/estradiol/PARP1 inhibitor, in order to verify this observation I undertook an additional experiment to compare etoposide only to treatment with etoposide and estradiol or PARP1 inhibitor or both. Figure 8.4B depicts these results. It is clear that there is no difference between these two treatments in regard to cell survival. Figure 8.4C depicts the cloning efficiency of the cell lines in etoposide vehicle DMSO alone (set at 100%) compared to DMSO plus either PARP1 inhibitor and/or estradiol, and indicates that PARP1 and estradiol had only a small but significant (Fig.8.4D) impact on survival of MCF7 cells separately or together.
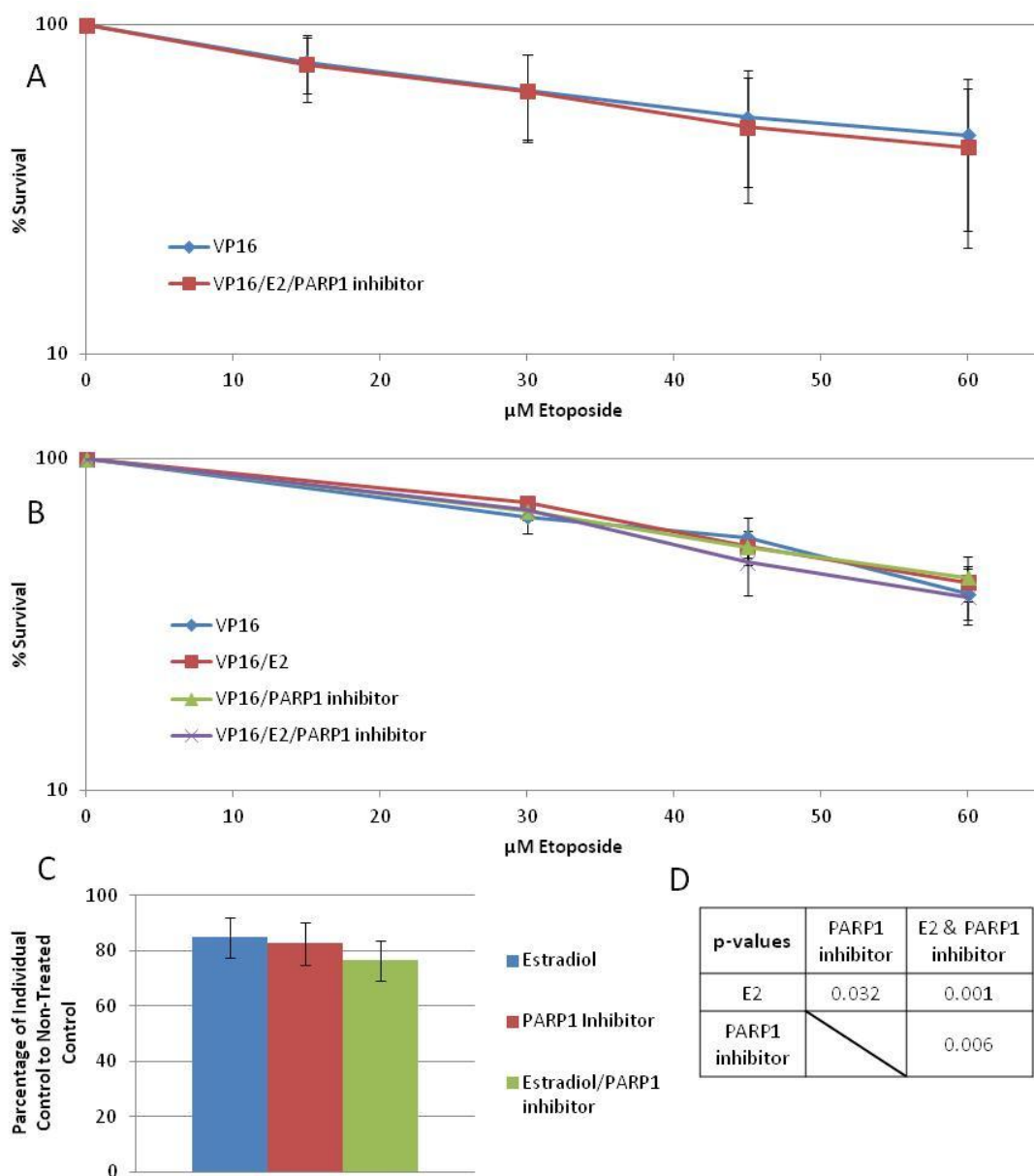
Figure 8.4: Comparison of sensitivity to etoposide, estradiol and etoposide, PARP1 inhibitor and etoposide or estradiol, PARP1 inhibitor and etoposide in breast cancer cell line MCF7. Etoposide (VP16) sensitivity was measured using Clonogenic Cell Survival assay in the MCF7 breast cancer cell lines. Etoposide concentrations varied. Estradiol (E2) concentration used was 100nM. PARP1 inhibitor concentration used was 500nM. Cells were treated for 30min with PARP1 inhibitor or DMSO, then 3h with estradiol or EtOH and 3h with etoposide or DMSO and left to grow in colonies for two weeks. The colonies were counted and the percentage of survival for each concentration was calculated from the untreated control. The percentages of survival were plotted against the corresponding etoposide concentrations. Data are the mean (+/- SEM) from three independent experiments. A) Etoposide concentrations used were 0µM, 15µM, 30µM, 45µM and 60µM. B) Etoposide concentrations used were 0µM, 30µM, 45µM and 60µM. C) Cloning efficiency of the cell line in etoposide vehicle DMSO alone (set at 100%) compared to DMSO plus either PARP1 inhibitor and/or estradiol. D) P-values were calculated with paired two-tailed t-test.

8.3 Discussion

This chapter examined whether triggering TOP2-dependent transcription by estradiol treatment resulted in cell sensitivity in the presence of PARP1 inhibitor (either by inhibition of SSBR and/or replication fork maintenance (Sugimura et al., 2008)) and/or etoposide (to convert TOP2 cleavage complexes into DSBs). However, I failed to detect such an effect.

There have been other studies in the past, in which treatment of etoposide and a PARP1 inhibitor only showed weak or no combination effect. One study used Olaparib, a PARP inhibitor currently used in cancer therapy, and etoposide, but the results were disappointing in terms of a synergistic effect, where an insignificant difference was found between single treatment and combined treatment (Murai et al., 2014). An older study showed that PARP inhibitor NU1025 did not increase DSBs or cytotoxic effect caused by etoposide alone and in addition, PARP was not activated after treatment with etoposide, even after the concentrations caused significant levels of apoptosis (Bowman et al., 2001). Finally, a different study had shown that the PARP inhibitor, 3-Aminobenzamide, when combined with etoposide had no significant effect in the human lymphoma cell line HT58 (Sebestyén et al., 1997).

In 2010, a group investigated the effects of estradiol on the cytotoxicity of etoposide in A549 lung cancer cells and did not see an increase in etoposide cytotoxicity, even though estradiol treatment showed a tendency towards increased proliferation rates (Sebestyén et al., 1997). This observation agrees with the lack of a synergistic effect in my results.

In Chapter 5, I showed that there isn't a direct correlation between sensitivity to etoposide and either TDP2 protein levels, TOP2A/TDP2 protein ratios, or TOP2B/TDP2 protein ratios. One possible explanation for the failure to detect an effect is that there is still enough intrinsic TDP2 in the cancer cells to cope with the induced abortive DSBs.

Another reason why the combination treatment was unsuccessful in increasing etoposide sensitivity in the breast cancer cell lines used, could be that the level of induction of transcription due to the estradiol treatment was not significant enough to result in many TOP2B-mediated DSBs that then could be rendered abortive via etoposide treatment.

Therefore, to test these ideas, I manipulated TDP2 protein levels in cancer cells and also measured the impact of hormone treatment on the expression of TOP2-dependent genes.

9. CHAPTER NINE – Impact of TDP2 depletion on TOP2B DSB repair in MCF7 breast cancer cells during estrogen-dependent transcription.

9.1 Introduction

Previous chapters explored the possibility that intrinsic levels of TDP2 might correlate with sensitivity of cancer cells to etoposide, however I did not find evidence for this. I also examined the possibility of combining etoposide treatment with PARP1 inhibition, testing the idea that the elevated DSBs arising in breast cancer cells due to the estrogen-induced transcription may render them hypersensitive to PARP1 inhibition. This is because the PARP1 protein is implicated in proper protection of replication forks that encounter DSBs by HR (Sugimura et al., 2008). However, my studies were unable to provide evidence to support this hypothesis.

In this chapter, I adopted a different approach. I directly manipulated TDP2 levels in cancer cells, with the view that if loss of TDP2 sensitises to etoposide-induced DSBs, or to those caused simply by inducing an estrogen-driven transcription program, then this would provide proof in principle for the development and utility of small molecule inhibitors of TDP2. There are already promising *in vitro* and *in vivo* results in that field (Raoof et al., 2013, Kankanala et al., 2016, Marchand et al., 2016, Kont et al., 2016, Kossmann et al., 2016, Hornyak et al., 2016). The cell line that was chosen for this series of experiments was the breast cancer cell line MCF7.

9.1.1 Aims of this chapter

The aim of this chapter was to demonstrate that introducing TOP2B-mediated DSBs via estradiol-induced transcription and rendering them abortive by etoposide treatment would have a larger effect on cell viability and repair kinetics in TDP2-depleted MCF7 cells than in mock-depleted MCF7 cells.

9.2 Results

I first attempted to establish a TDP2 knock-down in MCF7 breast cancer cells. This was possible with the help of two vectors: pSUPER-TDP2 in order to knock-down TDP2 or empty pSUPER for the mock knock-down and pcD2E as the selectable marker for G418 (Brummelkamp et al., 2002). This approach was successful, as indicated by the western blot in Figure 9.1A. The next step was to examine differences in cell viability after etoposide treatment in the wild type, mock-depleted, and TDP2-depleted MCF7 cell lines, using Alamar Blue assays. As indicated in Figures 9.1B-D, there was no significant difference between the mock-depleted and TDP2-depleted MCF7 cell lines for the lowest and highest concentrations of etoposide, but a small yet significant difference was observed at 500nM and 750nM of etoposide (Appendix 1). Next, I examined the TDP2-depleted MCF7 cell line by clonogenic cell survival assay. As expected, TDP2 depletion significantly decreased cellular resistance to etoposide, compared to both the mock-depleted and wild type MCF7 cells (Fig.9.1E, Appendix 2).
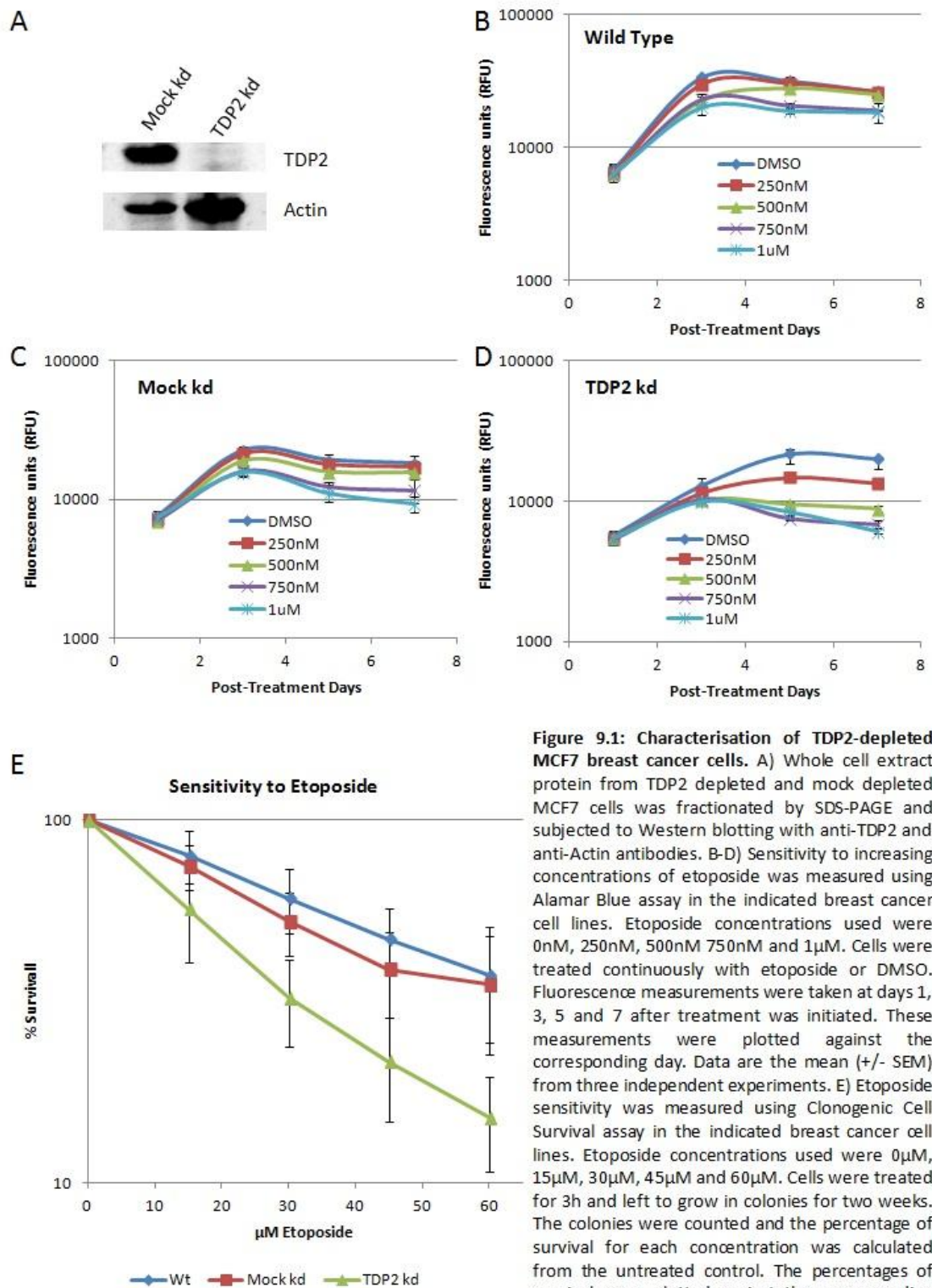
Figure 9.1: Characterisation of TDP2-depleted MCF7 breast cancer cells. A) Whole cell extract protein from TDP2 depleted and mock depleted MCF7 cells was fractionated by SDS-PAGE and subjected to Western blotting with anti-TDP2 and anti-Actin antibodies. B-D) Sensitivity to increasing concentrations of etoposide was measured using Alamar Blue assay in the indicated breast cancer cell lines. Etoposide concentrations used were 0nM, 250nM, 500nM 750nM and 1µM. Cells were treated continuously with etoposide or DMSO. Fluorescence measurements were taken at days 1, 3, 5 and 7 after treatment was initiated. These measurements were plotted against the corresponding day. Data are the mean (+/- SEM) from three independent experiments. E) Etoposide sensitivity was measured using Clonogenic Cell Survival assay in the indicated breast cancer cell lines. Etoposide concentrations used were 0µM, 15µM, 30µM, 45µM and 60µM. Cells were treated for 3h and left to grow in colonies for two weeks. The colonies were counted and the percentage of survival for each concentration was calculated from the untreated control. The percentages of survival were plotted against the corresponding etoposide concentrations. Data are the mean (+/- SEM) from three independent experiments.

Next, I examined the impact of TDP2 depletion on DSB repair rates, as measured by immunostaining for γH2AX. Figure 9.2A revealed that there is a significant increase in foci per cell in TDP2-depleted cells, at six hours after irradiation. This suggests that there is a delayed response in DSB repair in the TDP2-depleted cells. Representative images of the γH2AX foci in mock-depleted and TDP2-depleted cells are presented in Figure 9.2B.
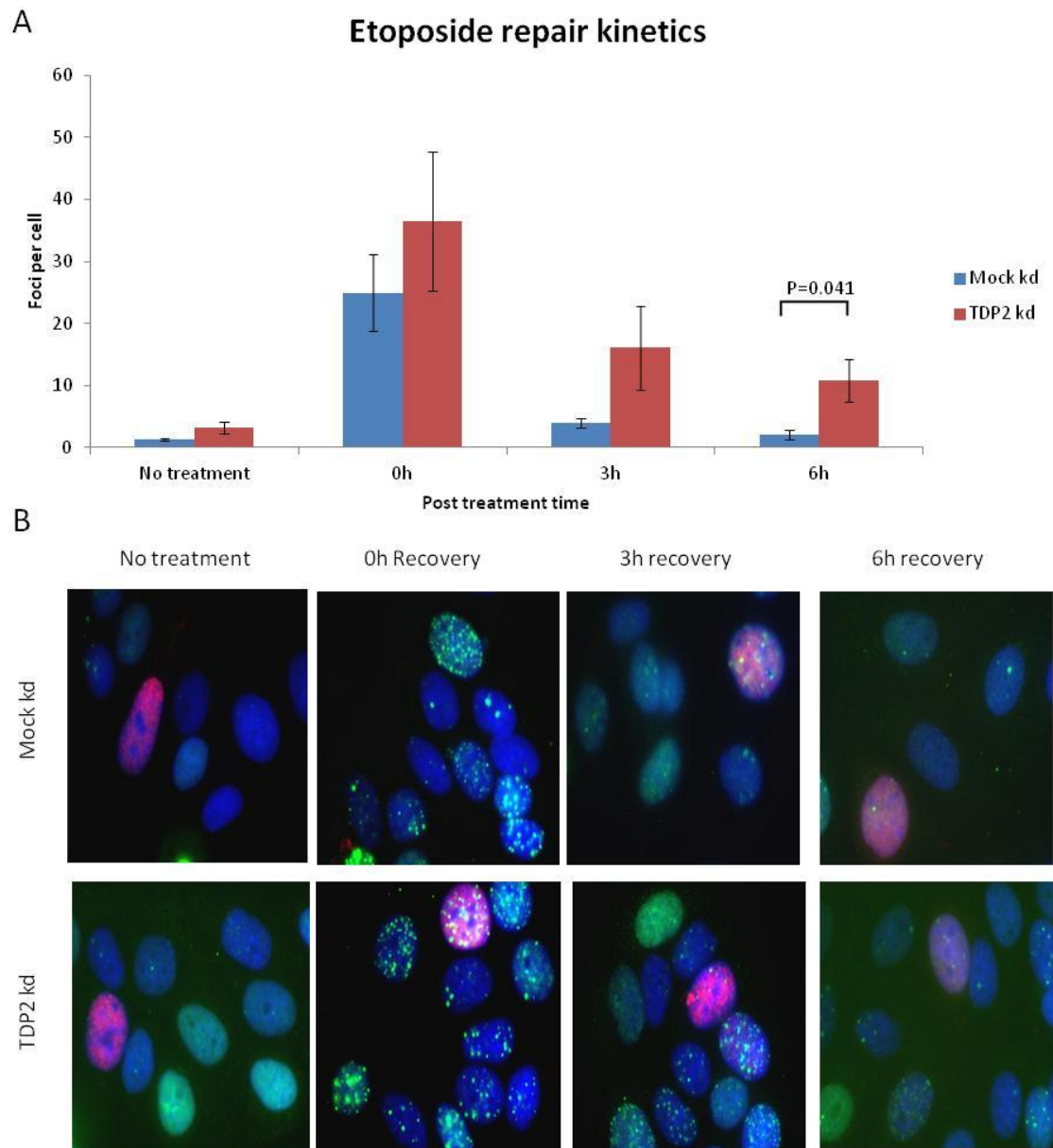
Figure 9.2: Etoposide Repair kinetics of TDP2-depleted MCF7 breast cancer cells. Measured by immunostaining for γH2AX. Cell lines used were TDP2 depleted and mock depleted MCF7 breast cancer cells. Cells were treaded with DMSO or 20μM of etoposide for 30min and left to recover for 3h and 6h. A) γH2AX per cell was counted and plotted against recovery time after treatment with etoposide. B) Representative images of the γH2AX foci.

Next, I established the conditions for estradiol induced transcription in the mock-depleted and TDP2-depleted MCF7 cells using two known TOP2-dependent genes, TFF1 and GREB1 and two control genes, Actin and TBP. TFF1 (Trefoil factor 1) is a protein expressed in breast cancer cells, among others and it contains an estrogen-inducible sequence, so that its expression can be regulated by estrogen (Pelden et al., 2013). GREB1 (Growth regulation by estrogen in breast cancer 1) is an early response gene in the estrogen receptor-regulated pathway and is expressed in hormone-responsive breast cancers (Ghosh et al., 2000). Actin is a commonly used housekeeping gene that has been proven to be reliable for use in breast cancer cell lines (Liu et al., 2015, Morse et al., 2005). In addition, TBP (TATA-binding protein) has also been determined as a suitable housekeeping gene for RT-qPCR in breast cancer cell lines (Lyng et al., 2008). From preliminary experiments (data not shown) it was determined that the best concentration for maximum induction of the TOP2-dependent genes was 100nM of estradiol over a period of 24h. From the graphs in Figure 9.3, it was evident that TFF1 and GREB1 were induced in a time dependent manner by exposure to 100nM estradiol in both mock-depleted and TDP2-depleted MCF7 cells. TDP2 depletion had no impact on the transcript levels, because any difference in the graphs was not statistically significant (Paired t-test p-values>0.05 for all samples).
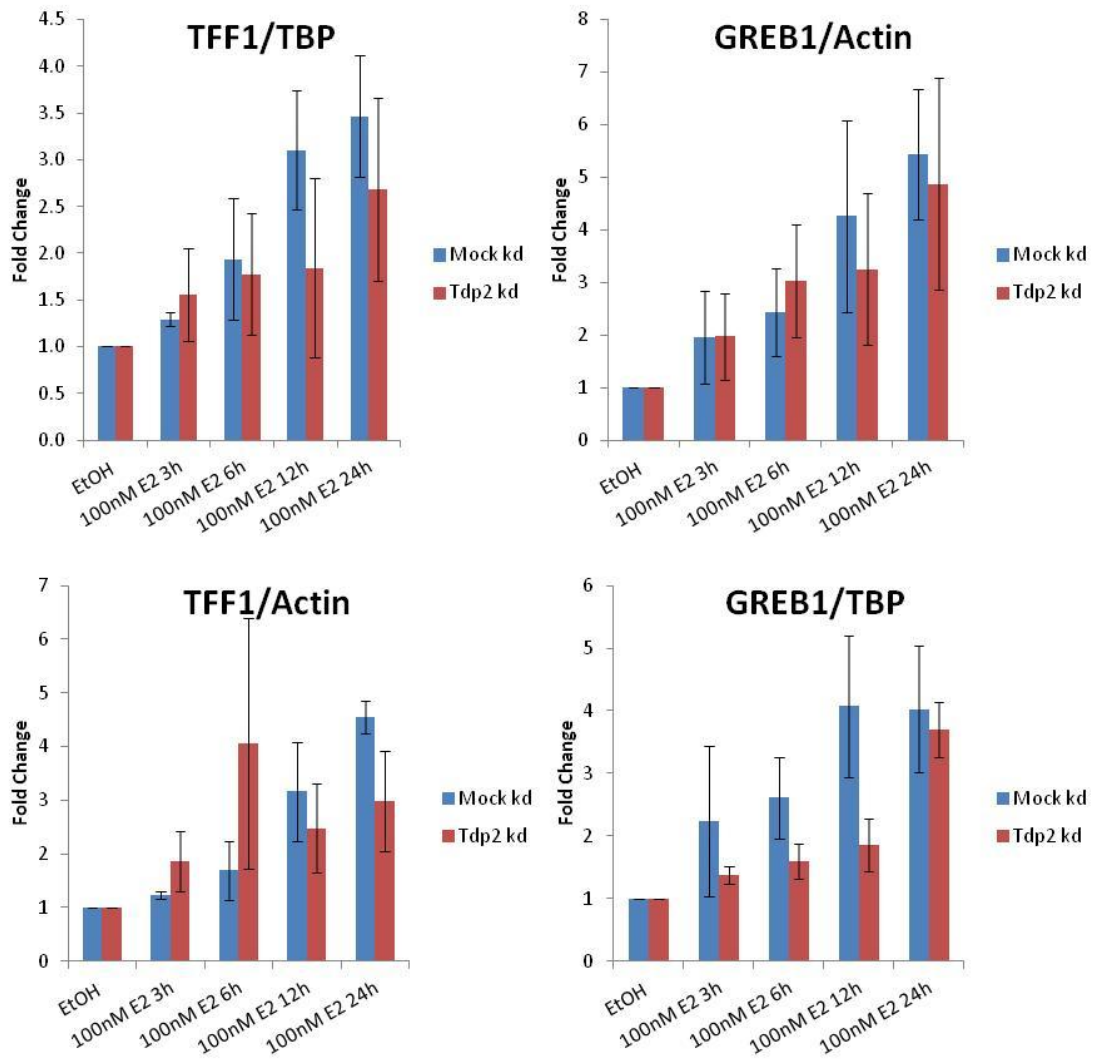
Figure 9.3: Induction of transcription following estradiol treatment. TFF1 and GREB1 mRNA expression levels were measured using SYBR Green qPCR in the indicated MCF7 breast cancer cell lines after treatment with estradiol (E2) for 3h, 6h, 12h and 24h, and are expressed relative to the expression level of the TBP or Actin housekeeping gene. Relative expression (fold-change) was calculated from the equation relative expression=$2^{-\Delta Ct}$, where ΔCt is calculated by subtracting the Ct value of the housekeeping gene from the Ct value of the gene of interest. Data are the mean (+/- SEM) from three independent experiments.

Once the induction of transcription was evaluated, I proceeded to examine the formation of DSBs as measured above. Figure 9.4A, reveals that in the TDP2-depleted cells the number of foci per cell was significantly higher than the mock-depleted and wild type cell lines after estradiol treatment. Representative images are presented in Figure 9.5. Because TOP2 activity and DSBs are higher in S-phase, I wanted to confirm that the differences detected here between mock-depleted and wild type versus TDP2-depleted cells reflected transcription-induced DSBs, rather than differences in cell cycle progression. This was unlikely since I counted only cells that lacked CENP-F staining, and which were thus most likely in G1. However, to confirm this I measured EdU incorporation. I conducted two different experiments; one in which EdU was present for a short time before cell fixation and one in which it was present the entire time of the treatment with estradiol. The results are presented in Figures 9.4B and C respectively, and reveal that there was no difference in the number or fraction of S-phase cells in mock-depleted and TDP2-depleted MCF7 cells.
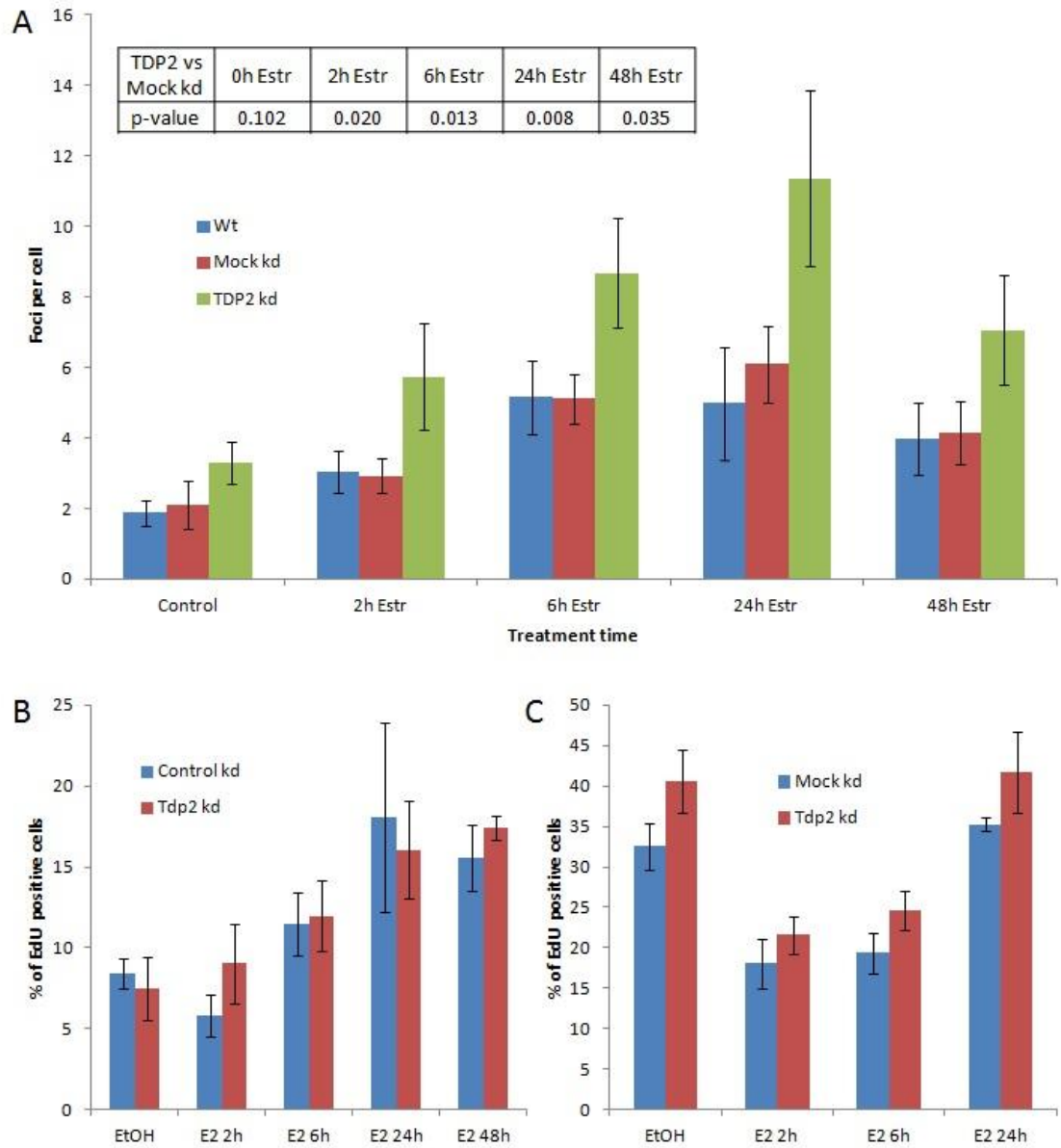
Figure 9.4: γH2AX foci per cell after treatment with estradiol in TDP2-depleted MCF7 breast cancer cells. Measured by immunostaining for γH2AX. Cell lines used were TDP2 depleted, mock depleted and wild type MCF7 breast cancer cells. A) γH2AX foci per cell were counted and plotted against treatment time with estradiol (E2). B) EdU incorporation. EdU was present for 15min before cell fixation. C) EdU was present the entire time of the treatment with estradiol.
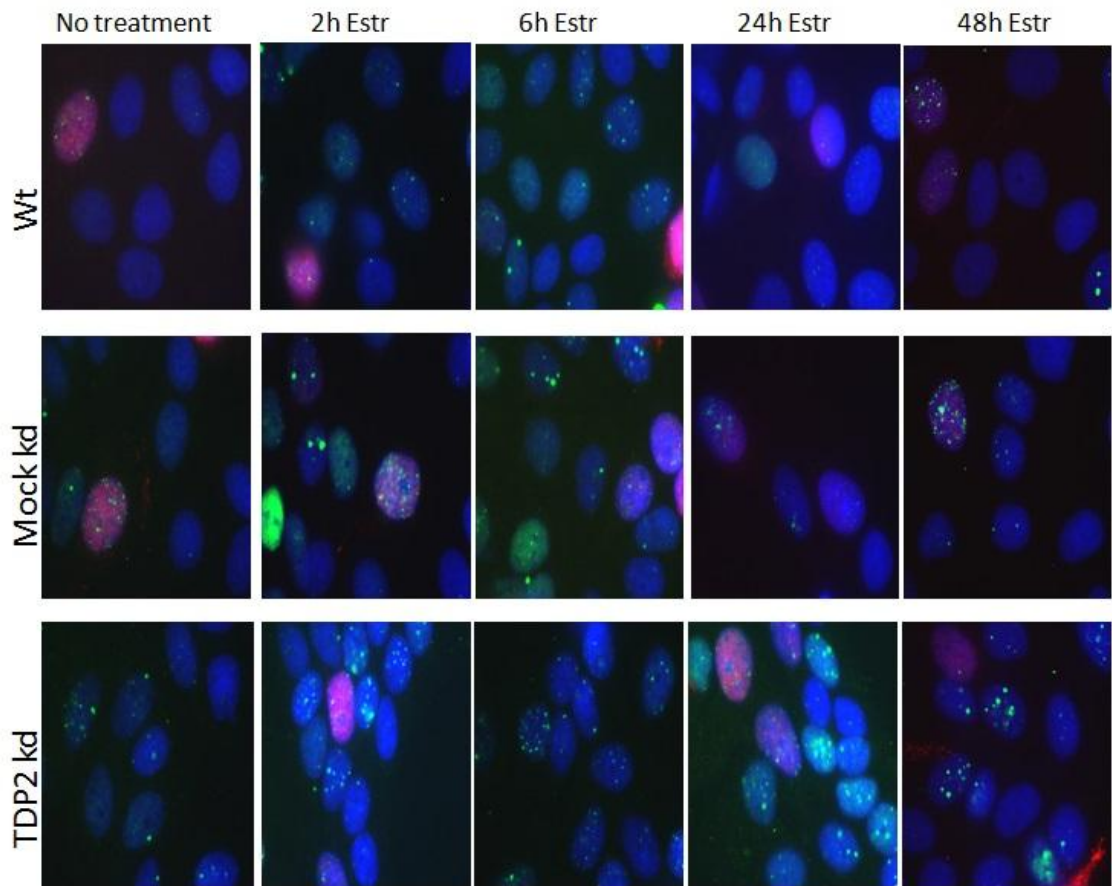
Figure 9.5: Representative images of γH2AX foci per cell after treatment with estradiol in TDP2-depleted MCF7 breast cancer cells. First row depicts wild type MCF7 cells, second row depicts mock-depleted MCF7 cells and third row depicts TDP2-depleted MCF cells.

Next, I explored the possible synergistic impact of combining estradiol treatment with etoposide. These experiments revealed that the survival of TDP2-depleted cells is significantly lower that of the mock-depleted cells, but this is not further enhanced by additional treatment with estradiol (Fig.9.6). Estradiol treatment alone did not affect cell survival.
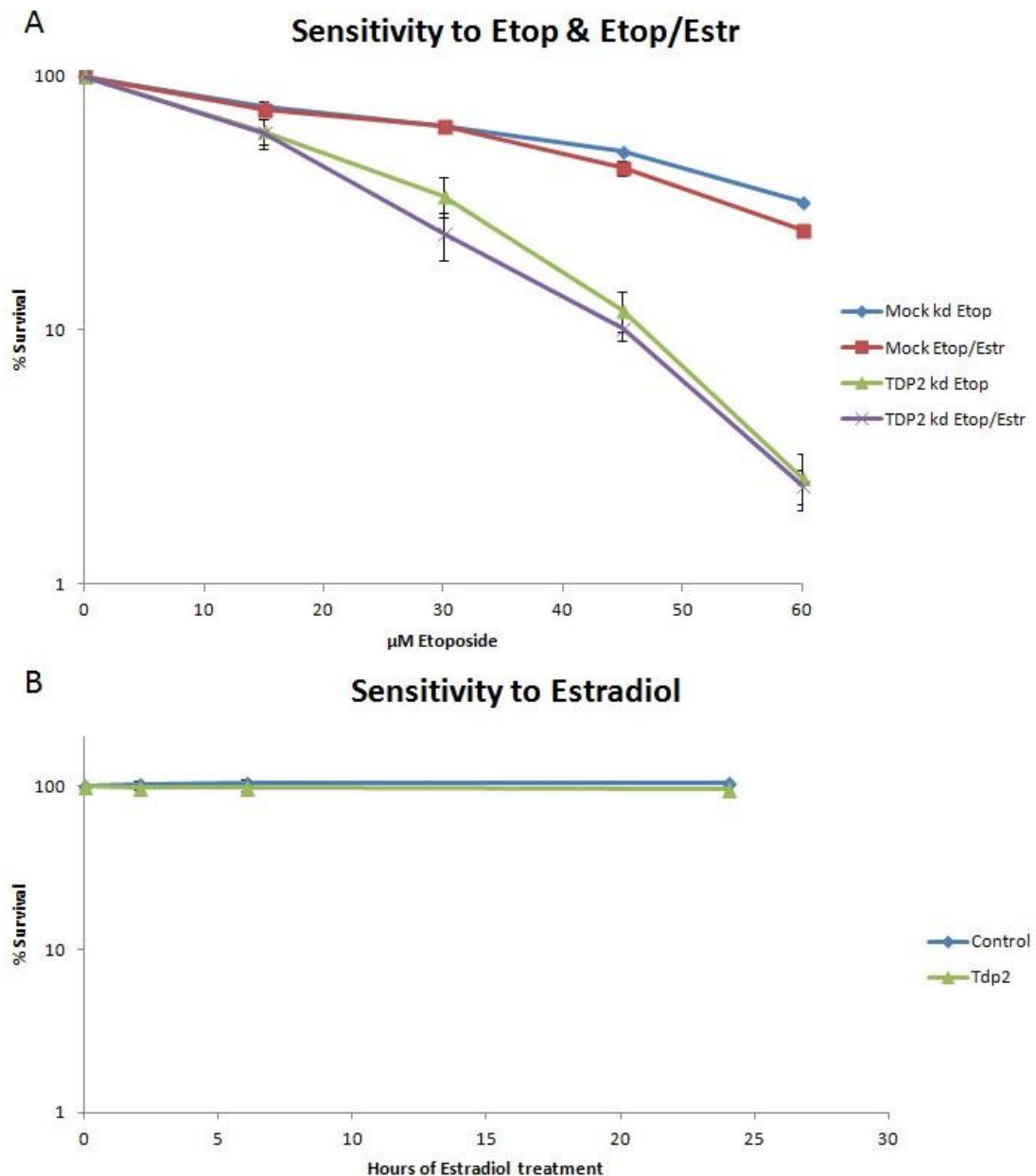
Figure 9.6: Comparison of sensitivity to etoposide, estradiol/etoposide or estradiol in TDP2-depleted or mock-depleted MCF7 breast cancer cell line. A) Etoposide sensitivity was measured using Clonogenic Cell Survival assay in the indicated cancer cell lines. Etoposide concentrations used were 0μM, 15μM, 30μM, 45μM and 60μM. Estradiol concentration used was 100nM. Cells were treated for 3h with estradiol or EtOH and 3h with etoposide or DMSO and left to grow in colonies for two weeks. The colonies were counted and the percentage of survival for each concentration was calculated from the untreated control. The percentages of survival were plotted against the corresponding etoposide concentrations. Data are the mean (+/- SEM) from three independent experiments. B) Estradiol sensitivity was measured using Clonogenic Cell Survival assay in the indicated cancer cell lines. Estradiol concentration used was 100nM. Cells were treated for 3h, 6h and 24h with estradiol or EtOH and left to grow in colonies for two weeks. The colonies were counted and the percentage of survival for each concentration was calculated from the untreated control. The percentages of survival were plotted against the corresponding treatment time. Data are the mean (+/- SEM) from three independent experiments.

As a final experiment of my project, I utilised the TDP2-depleted MCF7 cell line to explore DSB formation in the form of γH2AX foci after a combined treatment of estradiol and etoposide. As previously shown, treating these cell lines with estradiol caused a significant increase in γH2AX foci per cell in the TDP2-depleted cells (Fig.9.4A). This can also be seen in Figure 9.7. When both cell lines were treated with a high dose of etoposide only, they both responded with a high number of γH2AX foci per cell. Interestingly, when both cell lines were treated with a combination of estradiol and etoposide, the number of foci per cell increased significantly in the TDP2-depleted cells, but not in the mock-depleted cells.

	DMSO/EtOH	DMSO/Estr	50nM Etop/EtOH	50nM Etop/Estr
p-values	0.003	0.008	0.287	0.006
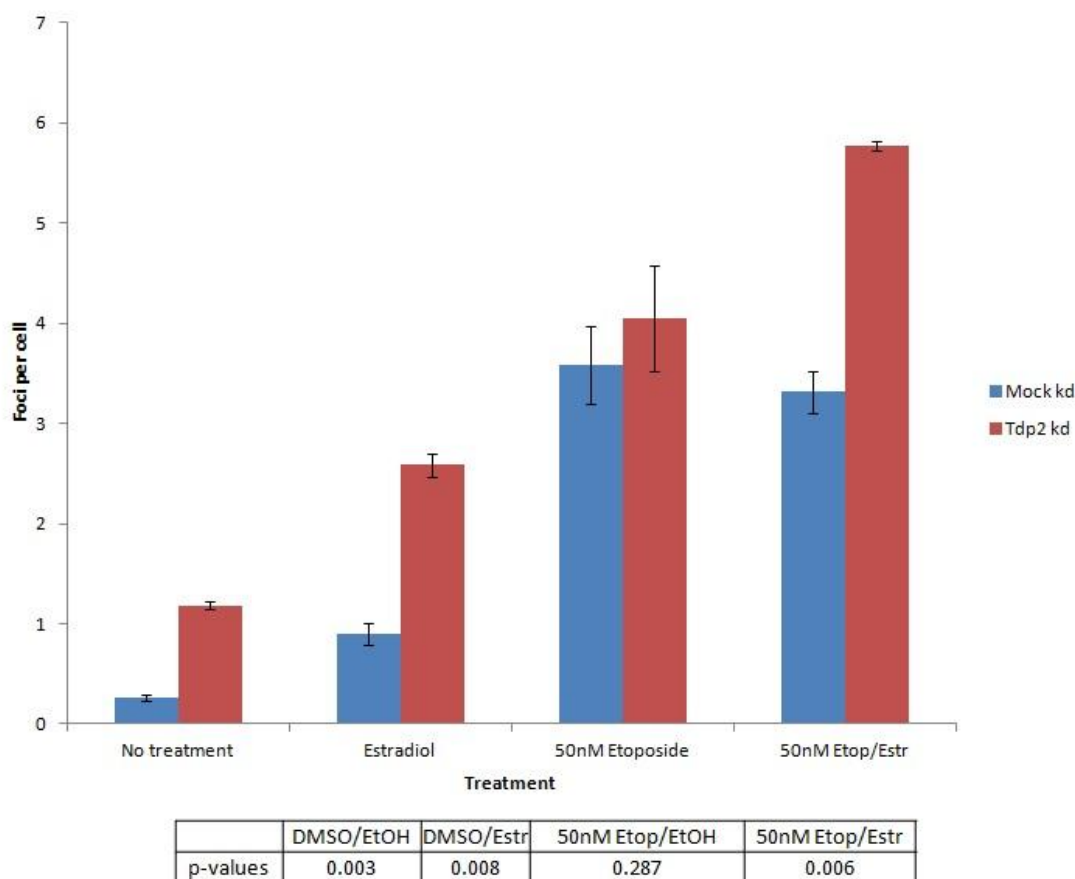
Figure 9.7: γH2AX foci induction by estradiol, etoposide or estradiol and etoposide in TDP2-depleted MCF7 breast cancer cells. Measured by immunostaining for γH2AX. Cell lines used were TDP2-depleted and mock-depleted MCF7 breast cancer cells. γH2AX foci per cell were counted and plotted against treatment type. Treatment time was 24h. P-values represent 1-tailed, paired t-test.

9.3 Discussion

This chapter of my project attempted to determine whether TDP2 can be exploited therapeutically in cancer treatment, in the context of hormone driven cancers such as breast cancer. The hypothesis was that inducing TOP2-dependent transcription programmes with estradiol in breast cancer cells lacking TDP2 might cause sufficient TOP2B-induced DSBs, before or after etoposide treatment, to overwhelm cells with DSBs.

To answer this question, the first step was to directly manipulate TDP2 protein levels by knocking down TDP2 in a cell line that is both relevant and accessible. For those reasons, MCF7 was chosen. After successfully knocking down TDP2, the TDP2-depleted cell line was characterised. It is more sensitive to etoposide treatment compared to both wild type and mock-depleted cells (established by clonogenic cell survival assays) and it demonstrated delayed repair kinetics. Cell viability was also lower, though these results were not as clear as the clonogenic cell survival assay results. The differences observed via clonogenic cell survival assay results are statistically significant, while the differences observed in the Alamar Blue result are not consistently statistically significant. Unlike in previous Alamar Blue experiments where small variations in TDP2 levels among different lung and breast cancer cell lines proved difficult to be recognised, depletion of TDP2 was enough to be detected by Alamar Blue assay after continuous treatment with increasing concentrations of etoposide. Furthermore, the small observed difference was significant for the concentrations of 500nM and 750nM of etoposide (Appendix 1). Additional Alamar Blue experiments demonstrated that very high concentrations of etoposide do not have a distinguishable effect on all three MCF7 cell lines (Appendices 3 & 4).

The observations that showed increased etoposide sensitivity in cell lacking TDP2 are in agreement with multiple publications that have shown similar results in different cell lines. For example, it has been shown that depletion of TDP2 in A549 and MEF cells, resulted in

increased sensitivity to etoposide treatment (Gomez-Herreros et al., 2013). The same has been shown for DT40 cells, where they exhibited higher sensitivity to etoposide after *TDP2* depletion (Zeng et al., 2011).

Transcription can be induced by estradiol treatment, peaking at 24h of treatment in wild type MCF7 cells (Appendix 5), mock-depleted MCF7 cells and TDP2-depleted MCF7 cells. Similar results were found when treatment with DHT (Dihydrotestosterone – an androgen hormone) induced transcription of AR (androgen receptor) responsive genes, but not control genes in both wt and TDP2-depleted LNCaP prostate cancer cells (Gómez-Herreros et al., 2014). Even though the differences I observed were not significant, there was a trend that showed lower induction levels in TDP2-depleted MCF7 cells compared to mock-depleted cells. This was seen at significant levels by (Gómez-Herreros et al., 2014), agreeing with their suggestion that TDP2 is required for induction of AR (and in my system perhaps ER too, although it will require targeted experiments) responsive genes, even at endogenous levels of TOP2-induced DSBs. This is a very important observation as TOP2B has already been shown to be essential in transcription regulation by introducing transient, site-specific DSBs, which probably occur frequently enough to cause an important demand of TDP2 (Ju et al., 2006, Pommier et al., 2016).

Estradiol alone also seems to cause increased number of DSBs in TDP2-depleted cells compared to wild type and mock-depleted cells, shown as γH2AX foci per cell. This supports the hypothesis that inducing TOP2B-dependent transcription programmes with estradiol generates TOP2B-dependent DSBs, visualised as γH2AX foci. This has been seen before by other groups (Williamson and Lees-Miller, 2011). At the same time, the impact of TDP2 depletion on the repair of TOP2B-mediated DSBs is evident by the near 2-fold increase of the number of DSBs after 24h of estradiol treatment.

Combining the two treatments (estradiol and etoposide) has a significant effect in cell survival in the TDP2-depleted cells, compared to the mock-depleted cells. When I tried to determine whether this was due to the induction of transcription and what that means for a cell, the results were not as expected. There seemed to be no significant difference in the percentage of survival between TDP2-depleted cells treated with etoposide only, compared with estradiol and etoposide treatment. The same lack of significant difference was observed in the mock-depleted cells. In the previous chapter, I showed that there was no significant difference between these treatments in wild type MCF7 cells.

A reason for this is that the induction of transcription and the subsequent DSBs are possibly not enough to have a significant impact on cell survival, although my results are in accordance with what other people have seen (Williamson and Lees-Miller, 2011). Another study has also reported lack of an increased effect on etoposide sensitivity after treatment with estradiol, even though that group also saw induction of transcription (Sebestyén et al., 1997). At the point of choosing a cell line, I could not have foreseen that the level of transcriptional induction might have been an issue. In hindsight I would have chosen a cell line with higher transcriptional induction. In addition, perhaps even the fractional levels of TDP2 remaining in the TDP2-depleted cells might have been enough to compensate for all the abortive DSBs.

Nevertheless, I was able to demonstrate by observation of γH2AX foci that treatment with estradiol alone caused a significantly higher number of DSBs in TDP2 depleted cells than in mock-depleted cells. In addition I showed that a combined treatment of etoposide and estradiol caused a significantly higher number of DSBs (depicted as γH2AX foci per cell) in the TDP2-depleted cells than the mock-depleted cells. Furthermore, I have shown that lack of TDP2 causes delayed DSB repair kinetics. Delayed repaired kinetics after etoposide treatment have also been shown in *TDP2*$^{Δ1-3}$ mice primary cortical astrocytes compared to *TDP2*$^{+/+}$, in

TDP2+/$^{\Delta1\text{-}3}$ mice cerebellar granule neurons compared to *TDP2*$^{\Delta1\text{-}3}$ and in human lymphoblastoid cells with mutated *TDP2* compared to normal (Gómez-Herreros et al., 2014).

Based on these two observations, the following hypothesis can be generated. Firstly, inhibiting TDP2, while inducing transcription with estradiol, will increase TOP2B-mediated DSBs (compared to estradiol-treated uninhibited cells). By treating these cells with etoposide, the DSBs will be made abortive. The lack of functional TDP2 will lead to delayed repair kinetics, meaning that abortive DSBs will persist long enough to cause cells to enter apoptosis and die. Therefore, I believe that under the right conditions, TDP2 can be therapeutically exploited. Any future studies would greatly benefit from a cell line with higher transcription induction levels and the development of a TDP2-targeting agent, in the same manner that one exists for TOP2.

10. CHAPTER TEN – Discussion

10.1 Overview

TDP2 has been identified as the enzyme that efficiently restores 5'-phosphate termini at DNA double-strand breaks in preparation for DNA ligation in human cells (Cortes Ledesma et al., 2009). It has been previously shown that cellular depletion of TDP2 leads to increased sensitivity to the commonly used TOP2-targeting agent etoposide (Zeng et al., 2011). In a recent small scale study that includes some lung and breast cancer cell lines, it was shown that TDP2 levels vary among the different cell lines, ranging from very high to very low (Li et al., 2011a).

Induction of transcriptional programs, for example by stimulating breast or prostate cancer cells with estrogens or androgens respectively, can involve the formation of TOP2B-mediated DSBs and the recruitment of DSB repair proteins, such as PARP1 (Ju et al., 2006). TOP2B is believed to be recruited with the estrogen receptor to regulatory sites on target genes for efficient activation of those genes (Haffner et al., 2011, Ju et al., 2006, Williamson and Lees-Miller, 2011). TDP2 has been shown to protect such genes from these naturally occurring DSBs, as cells with depleted TDP2 have demonstrated reduced induction of transcription, suggesting how important the resolution of those TOP2B cleavage complexes is (Gómez-Herreros et al., 2014).

Bringing all this information together, the aim of this thesis was to characterise the mRNA and protein levels of TDP2, TOP2A and TOP2B in a panel of non-small cell lung cancer cell lines and a panel of breast cancer cell lines, to explore whether these levels exhibit a large variation, as previously reported. The next step was to investigate the correlation between intrinsic TDP2, TOP2A and TOP2B protein levels and resistance to the TOP2-targeting agent etoposide, based on the hypothesis that cell lines with low levels of TDP2 and high levels of TOP2A and TOP2B

will be more sensitive to etoposide, compared to cell lines with high levels of TDP2 and low levels of TOP2A and TOP2B. The target of the first part of my project was to establish TDP2 as a biomarker for sensitivity to etoposide and a clinical predictor of treatment outcome.

The second part of my project was based on the model of TOP2B-mediated DSBs during estradiol-induced transcription. This aimed to explore the possibility of TDP2 as a novel therapeutic target for inhibition. The hypothesis was that inducing transcription in selected breast cancer cell lines with estradiol would lead to TOP2B-mediated DSBs and treating them with etoposide, thus making DSBs abortive, would overwhelm TDP2 and render it unable to cope with the high number of abortive DSBs, leading to apoptosis. It was also important to explore this model in a breast cancer cell line that lacked TDP2, which, if the results were promising, would steer future projects towards a compound that is able to inhibit TDP2 directly without the need of a knock-out.

10.2 TDP2, TOP2A and TOP2B mRNA and protein levels vary among a panel of lung and breast cancer cell lines.

I first investigated the mRNA and protein levels of TDP2, TOP2A and TOP2B in a panel of lung and a panel of breast cancer cell lines. The RT-qPCR and western blot data that emerged revealed a clear variation in mRNA and protein levels in TDP2, TOP2A and TOP2B among both cancer cell line panels. However, these results were not the same. It is very common for protein levels to not reflect mRNA levels, due to the abundance of mechanisms that act post-translation, which have an impact on protein production and stability (Vogel and Marcotte, 2012, Stark et al., 2006, Dickson et al., 2007, Taquet et al., 2009). Nonetheless, they both provide information in regard to how variant TDP2, TOP2A and TOP2B mRNA and protein levels are in different lung and breast cancer cell lines. These observations agree with a previous study that showed that in non-small cell lung cancer cell lines the levels of TDP2 varied, with H460, A459 and H226 demonstrating medium to high levels, and H292 and H1650

showing almost no TDP2 (Li et al., 2011a). I was also able to demonstrate that there was no correlation between the protein levels of TDP2, TOP2A and TOP2B. Just because a cell line has very high levels of TDP2, doesn't mean it will also have very high levels of TOP2A and TOP2B, although that doesn't mean that it can't, as shown by MOR and LCC9 which have high TDP2 and TOP2 levels. From these results, certain cell lines emerged as the ideal candidates to potentially be utilised to explore whether TDP2 could be a biomarker for sensitivity to the TOP2-targeting agent etoposide and therefore exploited therapeutically. The ideal candidate to further this idea would be a cell line with low levels of TDP2 and high levels of TOP2A and TOP2B. On the other hand, a cell line with high levels of TDP2 and low levels of TOP2A and TOP2B would be the least desirable candidate, which could serve as the baseline. High levels of TOP2A and TOP2B could mean more TOP2-mediated abortive DSBs within a cell and low levels of TDP2 could mean that TDP2 could be overwhelmed by these abortive DSBs, leading the cell to apoptosis. An abundance of TDP2 (high protein levels), on the other hand, would be able to resolve the fewer TOP2-mediated DSBs caused by low levels of TOP2A and TOP2B. Among the lung cancer cell lines, H226 and A549 emerged as good candidates and among the breast cancer cell lines ZR751 fulfilled the criteria.

In an attempt to minimise problems with the detection of the protein levels, three different concentrations of whole cell extract were used, that were determined by accurate cell counting and a strict sample preparation protocol. There had been attempts to utilise Bradford protein assay to measure protein concentrations experimentally, but significant protein degradation was observed and therefore this technique was abandoned. The repetition of the experiments a minimum of three times and the use of three volume based concentrations was implemented to maximise accuracy. Available equipment that could accommodate the number and volumes of the samples was used, but I believe that larger gels that could hold more samples at the same time might have improved the consistency of the results through the different concentrations. Perhaps a lower transfer voltage and longer transfer times could

have also helped with reducing discrepancies further, but the current protocol was assessed to be acceptable given the time available.

10.3 No direct correlation was detected between TDP2, TOP2A and TOP2B protein levels and sensitivity to etoposide in a panel of lung and breast cancer cell lines.

Once levels of TDP2, TOP2A and TOP2B were measured, the project progressed to explore the effects of etoposide on cell viability and survival of increasing treatment with etoposide. All the cell lines were treated with etoposide only. A549 was the lung cancer cell line least sensitive to etoposide and HCC95 the most sensitive. MM453 and BT474 were the breast cancer cell lines with highest resistance to etoposide and LCC9 and T47D the least resistant. A comparison of TDP2 protein levels and cell viability or survival failed to correlate TDP2 protein levels and sensitivity to etoposide in either breast or lung cancer cell lines. It appears that the protein levels do not necessarily translate to etoposide sensitivity or resistance. Cell viability and cell survival did not necessarily agree either, but it is important to remember that they offer different kinds of information. Alamar Blue is an assay that essentially demonstrates how many cells are still metabolically active after a certain treatment. This possesses limitations and makes this assay less sensitive at detecting small differences. It cannot differentiate between cells that have entered the path to apoptosis and healthy cells and this is a restriction not found in the clonogenic cell survival assay. The latter is able to show how many cells survived the treatment and were able to proliferate and produce colonies. Unfortunately, not all cells I utilised are able to grow in colonies and therefore are not suitable for that type of assay. In addition, there seemed to be no correlation between sensitivity to etoposide and TOP2A/TDP2 and TOP2B/TDP2 protein ratios either. All these results were verified with the calculation of correlation coefficients and their corresponding p-values.

I corroborated my results by using etoposide IC50 values from an appropriate database (CancerDr), for all of the tested cell lines that had this information available. I calculated the correlation coefficients and corresponding p-values against the TDP2 and TOP2 protein level results that I had obtained experimentally. None of these results were significant either.

The lack of correlation could be attributed to the levels of activity of TDP2 or perhaps even mutations that render it inactive. To explore this mechanistically, I would have utilised an activity assay to establish TDP2 catalytic activity for 5'-phosphotyrosyl linkage at the 5' ends of a ^{32}P-radiolabeled 20-bp oligonucleotide substrate analyzed by denaturing PAGE/phosphorimaging as described by Cortes Ledesma and co-workers; or I could have used a more recently developed assay that uses p-nitrophenyl-thymidine-5'-phosphate (T5PNP), a well-established substrate for snake venom phosphodiesterase, as a substrate for TDP2 (Cortes Ledesma et al., 2009, Adhikari et al., 2011). The biggest advantage of a T5PNP assay compared to the gel-based assay is that it can be used as continuous colorimetric assays in a 96-well format (Adhikari et al., 2011). In previous publications results obtained from both assays were comparable, but they also showed that TDP2 works more efficiently on the traditional 5'-phosphotyrosyl substrate than on T5PNP, while T5PNP is not a non-specific general phosphodiesterase substrate (Adhikari et al., 2011). Although this could have potentially explained the lack of direct correlation between TDP2 protein levels and resistance to etoposide, it would not have been particularly useful in a clinical setting. If the project ever moved outside of a molecular lab and into a diagnostics lab, running an activity assay would be very expensive and time consuming, and therefore unlikely to be used. Nevertheless, *TDP2* mutations have been identified in individuals with intellectual disability, seizures and ataxia, where mutant *TDP2* mRNA was found to be truncated and subject to nonsense-mediated decay (Gómez-Herreros et al., 2014).

A recently published study with a similar structure to mine investigated whether TDP1 or TOP1 levels alone could be a predictive biomarker of sensitivity to the TOP1 poison irinotecan, in colorectal cancer cell lines (Meisenberg et al., 2015). Their conclusion was that there was no correlation between intrinsic TDP1 or TOP1 levels and irinotecan, a system very similar to the TDP2-TOP2-etoposide system I was investigating. They only used 6 cell lines to study their colorectal cancer model, while I used two types of models (breast and lung cancer) with 8 cell lines each. Perhaps neither of those numbers is large enough to observe such correlation, or the opposite is true, as handling a large number of cell lines at the same time only serves to complicate the experimental study process rather than contributing to the observations made. In addition, the same group has shown in a previous study that TDP1 levels alone had failed to show a correlation to Topotecan (another water soluble analog of camptothecin) (Meisenberg et al., 2014). Instead they observed a correlation of Topotecan resistance with TDP1/TOP1 ratios only when certain data points were excluded, without a clear and logical reason other than being outliers. Furthermore, they used cell viability assays and not cell survival assays, the differences of which I have discussed earlier. Even though their aim was similar to mine, the differences in approach mean that it's not surprising that their results do not agree with the results detailed here.

There is always the question of whether immortalised cell lines accurately represent what is occurring *in vivo* – a more complex and variable model. They usually differ but there are methodologies that can be followed to bring those two models closer (Pan et al., 2009). Another variable that makes the system I use more complicated than the studies that just looked into effects of cells lacking TDP2, is the use of multiple cell lines. Even though every precaution was taken to maintain the cell lines at optimum conditions, by checking for mycoplasma and not over-culturing the cells, one fact still remains; some of the cell lines were obtained directly from a cell line repository and some were obtained from resources available within the University. Although the hypothesis that I was working under did not consider the

background of each cell line as much, and was based on experimental observations, this is a situation that should be considered carefully when using immortalised cell lines for research (Hughes et al., 2007).

Moving forward from cell line limitations in laboratory research, my results were not the only ones unable to show a correlation between TDP2 levels and response to etoposide. Although at a much smaller scale, earlier this year a group showed such lack of correlation in a much smaller number of lung cancer cell lines, in which even knocking down TDP2 didn't increase the number of DSBs after etoposide treatment, compared to wt (Bian et al., 2016).

There is also data that suggests a functional redundancy between the nuclease activity of Mre11 and TDP2 in terms of genome stability, which could be interpreted as another reason why no correlation was observed in my experiments (Hoa et al., 2016). The same paper demonstrated that heterozygous TDP2$^{+/-}$ cells were not more sensitive to etoposide than wt cells, which could imply that even a small presence of TDP2 is enough to respond to etoposide-induced abortive DSBs. This could explain why the intrinsic variability of TDP2 protein levels in the cell lines I studied is too subtle to have a significant impact on etoposide resistance

10.4 Database-derived mutational profiles of cancer cell lines are a primary useful tool in experimental result interpretation

In my attempt to investigate further why a correlation between TDP2 protein levels, TOP2A/TDP2 protein ratios or TOP2B/TDP2 protein ratios and etoposide sensitivity does not exist, I decided to collate mutational profiles of the cell lines I utilised with curated databases. This process differs from all the previous experimental methods that I used, as it relies on information provided by other people. Therefore, very strict guidelines needed to be followed. Chapter 6 has revealed some very interesting results and has demonstrated how useful such collections of information can be. This work also highlights the synergy in combining experimental work with database and literature mining.

The first interesting mutation is EGFR. In lung cancer cell lines it was shown to have a significant positive correlation with TDP2 protein levels, and could possibly affect the relationship between TDP2 protein levels and sensitivity to etoposide, although that could be through its effect on TDP2. PIK3CA mutation in breast cancer cell lines had a significant negative effect on TOP2A/TDP2 and TOP2B/TDP2 ratios, but not on TDP2 protein levels. This is in agreement with a group that has shown that the PIK3CA mutation decreased TOP2 expression levels (Fountzilas et al., 2012). Another interesting gene was PTEN, whose mutation in lung cancer cell lines had a significant negative effect on TDP2 protein levels and it was also shown to have a possibly significant effect on the relationship between TDP2 protein levels and sensitivity to etoposide, although similar to EGFR that could be due to its effect on TDP2 levels. A very frequent mutation in cancer, TP53 in breast cancer cell lines showed a significant positive correlation with TDP2 protein levels. SMARCA4 mutation in lung cancer cell lines appeared to have a significant effect on the relationship between TDP2 protein levels and etoposide sensitivity, when the mutation is taken into account.

To summarise, I propose that three mutations (EGFR, PTEN and TP53) could possibly have an effect on TDP2 protein levels, something that would be interesting to explore further, especially how this effect arises at a molecular level. Mutant TP53 has already been shown to target TDP2 transcription and mtp53 knock-down was shown to reduce both the mRNA and protein levels of TDP2, which agrees with the positive correlation I discovered (Do et al., 2012). I also identified a mutation (PIK3CA) that appears to negatively affect TOP2 protein levels, which also agrees with published work (Fountzilas et al., 2012). Finally, I identified a mutation (SMARCA4) that has a significant effect on the relationship between TDP2 protein levels and etoposide sensitivity. These 5 mutations could all be contributing to the lack of correlation between etoposide sensitivity and either TDP2 protein levels or TOP2/TDP2 protein ratios observed in Chapter 5. As certain cell lines are not used as extensively in scientific experiments, in some cases potentially interesting mutations may have been missed simply

because information was unavailable for the minimum required number of cell lines for statistical analysis.

Chapter 6 has resulted in a collection of mutation profiles of several breast and lung cancer cell lines that are available for use in the Caldecott lab and as an extension, the whole Genome Damage and Stability Centre, as well as other collaborators. This extensive list of thousands of mutations is available in Appendices 6 and 7. This chapter demonstrated the usefulness of mutation profiles in research, information that is consistently used in cancer pathology and the focus of many studies (McConechy et al., 2012, Chang et al., 2014, Bolli et al., 2014, Pereira et al., 2016).

10.5 Exploring the possibility of the use of TDP2 as a potential biomarker of sensitivity to other anti-cancer drugs

As evidence is pointing toward TDP2 not being a predictive biomarker of sensitivity to the TOP2 poison etoposide, I decided to further utilise my results from Chapters 3 and 4 and search for another anti-cancer drug for which TDP2 could potentially be a predictive biomarker. My decision was also based on the comparison of the statistical analysis of my results for etoposide IC50 and D37 values and the same analysis of IC50 values mined from appropriate databases, which were in agreement. Since there is a wealth of curated information on anti-cancer drug IC50 values in a broad range of cancer cell lines, I searched and found as many anti-cancer drugs as possible for which there was enough information to be able to use statistical methods for analysis.

Given how precious research funding is, a hypothesis is worth exploring when it is based on sound reasoning. My analysis looked at a large number of drugs with a variety of targets and thus, even though their IC50 values' correlation was statistically significant with the phenotypes studied, some of the results could be false positives. That means that the phenotypes might not actually correlate with the IC50 concentrations of these drugs if a

broader set of cell lines was examined. Taking this into consideration I supplemented correlation tests with further literature searches so as to focus only on drugs currently in use for cancer therapy, or presently in clinical trials for their potential use in cancer therapy, rather than commercially available inhibitors for laboratory use. I propose that the following anti-cancer drug results are significant enough to warrant further investigation. Vinblastine, ABT888, AG014688, Bleomycin, Camptothecin and Bicalutamide. Bicalutamide, AG014699, Camptothecin and Vinblastine all showed a direct correlation with TDP2 protein levels. Bicalutamide showed a positive correlation and the rest a negative. ABT888 and Bleomycin demonstrated a positive correlation with TDP2 via its TOP2A/TDP2 protein ratio. The background of these drugs offers a logical explanation as to why it is reasonable that TDP2 should be explored as a biomarker of sensitivity for these drugs.

The negative correlation between resistance to Vinblastine and TDP2 protein levels could be due to increased proliferation in NSCLC observed when TDP2 is overexpressed (Li et al., 2011a). With high levels of TDP2 there is increased proliferation, and thus more tubulin (Vinblastine's target). Additionally, Vinblastine along with etoposide induces the formation of apoptotic TOP1 cleavage complexes (Sordet et al., 2006), which TDP2 has been shown to promote the repair of in TDP1 knock-out cells (Zeng et al., 2012).

ABT888 and AG014688 are PARP1 inhibitors in which resistance was negatively correlated with TDP2 protein levels in lung cancer cell lines (AG014688) and positively correlated with TOP2A/TDP2 protein ratios in breast lines (ABT888). If the positive correlation between TOP2A/TDP2 protein ratios is due to lower levels of TDP2 rather than high TOP2A levels, then these results agree; even if the opposite is true, they are not mutually exclusive. Multiple lines of evidence allude to an interesting relationship between PARP1 and TOP2 (Sabisz et al., 2010, Ju et al., 2006, Mandraju et al., 2011), and since TOP2 is so closely related to TDP2, I believe

there is possibility for TDP2 or its ratio to TOP2 to be a candidate for sensitivity to PARP1 inhibitors.

Bleomycin works by introducing DSBs and SSBs, whose ratio and absolute values can determine which pathway will be followed for cell death: apoptosis, mitotic cell death, or pseudoapoptosis (Tounekti et al., 2001). Given that Bleomycin resistance is characterised by reduced DNA damage and the role of TOP2A and TDP2 in DSBR, this correlation is worth exploring further (Wang et al., 2013).

TDP2-depleted cells have shown lack of hypersensitivity to Camptothecin (compared to etoposide). However TDP2 has been shown to step in when TDP1 is absent and promote the repair of TOP1-mediated DNA damage in its place (Zeng et al., 2011, Zeng et al., 2012). Thus, it might be interesting to further expand on my results showing that high TDP2 protein levels had a negative impact on Camptothecin resistance.

Bicalutamide is a selective antagonist of the androgen receptor, used primarily for the treatment of prostate cancer (Singh et al., 2000). Androgen-deprived prostate cancer cells display transient TOP2B-dependent DSBs, when exposed to androgen; the same group also showed that exposing androgen-deprived prostate cancer cells to Bicalutamide lead to increased DSBs via γH2AX foci formation (Coulter et al., 2015). In Chapter 7, I showed that there is a positive correlation between resistance to Bicalutamide and TDP2 protein levels. Given the involvement of TDP2 in resolving TOP2-mediated DSBs, I believe that it merits further exploration *in vivo*.

It is important to note that similar to my project, such research projects entail a relatively high risk, as they are not mechanistic and therefore additional preliminary checks should be carried out. I would propose an investment in cell lines that are relevant to the types of cancer that these drugs are used for. But even before proceeding to obtain such cell lines for experimental data, they should be scrutinised for how extensive the available information on their profile is,

for example are IC50 values for these drugs available and do they vary, something that could be influenced by how frequently these cell lines are used in similar models. These are important questions that need to be answered to ensure that risk in minimised.

It is not easy to identify biomarkers. Cancer therapy is becoming more targeted, but the progress of developing tools to determine which patients will benefit from such therapies is slower. A biomarker should be highly sensitive, specific, and reliable with high prognostic value, organ specificity and it should correlate with tumour stages, according to (Daniel and Lalitha, 2016). That is of course not always the case. There are several research groups or even centres dedicated to the sole purpose of identifying and characterising new cancer biomarkers (e.g. the Biomarkers and Systems Medicine (BSM) group at the University of Glasgow, the Centre for Cancer Biomarkers (CCBIO) at the University of Bergen), suggesting how challenging and complicated this type of research can be. New proteomic and genomic technologies are utilised, but advance bioinformatics tools are equally indispensible as they allow the simultaneous analysis of thousands of biological molecules (Mäbert et al., 2014).

10.6 PARP1 inhibition and estradiol treatment failed to further sensitise LCC9, MLET5 and MCF7 breast cancer cells to etoposide treatment

Chapter 8 examined whether triggering TOP2-dependent transcription by estradiol treatment would result in cell sensitivity to PARP1 inhibitor (to inhibit SSBR and/or replication fork maintenance), either in the presence or absence of etoposide (to convert TOP2 cleavage complexes into abortive DSBs). However, a difference between etoposide treatment and the combination treatment was not observed.

There have been other studies in the past, in which treatment of etoposide and a PARP1 inhibitor only showed weak or no combination effect. One study used Olaparib, a PARP inhibitor currently used in cancer therapy, and etoposide, but the results were disappointing in terms of a synergistic effect, where an insignificant difference was found between single

treatment and combined treatment (Murai et al., 2014). An older study showed that PARP inhibitor NU1025 did not increase DSBs or cytotoxic effect caused by etoposide alone and, in addition, PARP was not activated after treatment with etoposide, even after the concentrations caused significant levels of apoptosis (Bowman et al., 2001). Finally, a different study had shown that the PARP inhibitor 3-Aminobenzamide, when combined with etoposide had no significant effect in the human lymphoma cell line HT58 (Sebestyén et al., 1997).

In 2010, a group investigated the effects of estradiol on the cytotoxicity of etoposide in A549 lung cancer cells and did not see an increase in etoposide cytotoxicity, even though estradiol treatment showed a tendency towards increased proliferation rates (Sebestyén et al., 1997). These observations agree with the lack of a synergistic effect in my results.

10.7 TDP2-depletion in MCF7 breast cancer cells increases sensitivity to etoposide and formation of TOP2B-mediated DSBs due to estradiol treatment.

I attempted to determine whether TDP2 can be exploited therapeutically in cancer treatment, in the context of slow proliferating hormone-dependent cancers such as breast tumours, where resistance has developed or other types of treatment have failed. The hypothesis was that inducing TOP2-dependent transcription programmes with estradiol in breast cancer cells lacking TDP2 might cause sufficient TOP2B-induced DSBs, before or after etoposide treatment, to overwhelm cells with DSBs. I successfully knocked down TDP2 in MCF7 breast cancer cell and I characterised the TDP2-depleted cell line.

The first important result was that TDP2 knock-down cells demonstrated increased sensitivity to etoposide treatment in clonogenic cell survival assays, compared to wild type and mock-depleted cells. This observation is in agreement with multiple publications that have shown similar results in different cell lines. It has been shown that depletion of TDP2 in A549 and MEF cells, resulted in increased sensitivity to etoposide treatment (Gomez-Herreros et al., 2013). The same has been shown for DT40 cells, which exhibited higher sensitivity to etoposide after

TDP2 depletion (Zeng et al., 2011). The increased sensitivity was also observed with cell viability assays. More specifically, unlike in previous cell viability experiments in this thesis where small variations in TDP2 levels among different lung and breast cancer cell lines proved difficult to be recognised, depletion of TDP2 was enough to be detected after continuous treatment with increasing concentrations of etoposide.

The second important result was that the TDP2 knock-down demonstrated delayed repair kinetics following etoposide treatment. Delayed repaired kinetics after etoposide treatment have been shown in *TDP2$^{\Delta1-3}$* mice primary cortical astrocytes compared to *TDP2$^{+/+}$*, in *TDP2^{+}/$^{\Delta1-3}$* mice cerebellar granule neurons compared to *TDP2$^{\Delta1-3}$* and in human lymphoblastoid cells with mutated *TDP2* compared to normal (Gómez-Herreros et al., 2014).

The third important result was that transcription can be induced by estradiol treatment, peaking at 24h of treatment in wild type MCF7 cells, mock-depleted MCF7 cells and TDP2-depleted cells. Similar results were observed when treatment with DHT induced transcription of AR (androgen receptor) responsive genes, but not control genes, in both wt and TDP2-depleted LNCaP prostate cancer cells (Gómez-Herreros et al., 2014). Even though the differences I observed were not significant, there was a trend that showed lower induction levels in TDP2-depleted MCF7 cells compared to mock-depleted cells. This was seen at significant levels by (Gómez-Herreros et al., 2014), agreeing with their suggestion that TDP2 is required for induction of AR responsive genes (and in my system perhaps ER too, although it will require targeted experiments), even at endogenous levels of TOP2-induced DSBs. This is a very important observation as TOP2B has already been show to be essential in transcription regulation by introducing transient, site-specific DSBs, which probably occur frequently enough to cause an important demand of TDP2 (Ju et al., 2006, Pommier et al., 2016).

Furthermore, estradiol alone also seems to cause a significantly increased number of DSBs in TDP2-depleted cells compared to mock-depleted cells, shown as γH2AX foci per cell. This

supports the hypothesis that inducing TOP2-dependent transcription programmes with estradiol generates TOP2-dependent DSBs. This agrees with a previously published study (Williamson and Lees-Miller, 2011). At the same time, the impact of TDP2 depletion on the repair of TOP2B-mediated DSBs is evident by the near 2-fold increase of the number of DSBs after 24h of estradiol treatment.

Additionally, I was able to demonstrate that combining the two treatments (estradiol and etoposide) has a significant effect on cell survival in the TDP2-depleted cells, compared to the mock-depleted cells. I was not able to demonstrate though, that induction of transcription was the reason behind this difference. There was no observed significant difference in the percentage of survival between TDP2-depleted cells treated with etoposide only, compared with estradiol and etoposide treatment. The same lack of significant difference was observed in the mock-depleted cells. This agrees with my results in Chapter 8, where I observed no significant difference between these treatments in wild type MCF7 cells. A reason for this could be that the level of induction of transcription in MCF7 cells and the subsequent DSBs are possibly not enough to have a significant impact on cell survival even in the absence of TDP2, although my results are in accordance with what other people have seen (Williamson and Lees-Miller, 2011). Another study has also reported lack of an increased effect on etoposide sensitivity after treatment with estradiol, even though that group also saw induction of transcription (Sebestyén et al., 1997). In hindsight, perhaps there would have been a benefit if I had chosen a cell line with higher levels of induction of transcription. Or perhaps the explanation lies with the fractional levels of TDP2 remaining in the TDP2-depleted cells, meaning that they might have been enough to compensate for all the abortive DSBs.

To summarise, I demonstrated that treatment with estradiol induces a significantly higher number of DSBs (depicted as γH2AX foci per cell) in the TDP2-depleted cells compared with the mock-depleted cells; a combined treatment of etoposide and estradiol causes a significantly

higher number of DSBs (depicted as γH2AX foci per cell) in the TDP2-depleted than the mock-depleted cells; and lack of TDP2 causes delayed DSB repair kinetics. Based on these observations the following hypothesis can be generated. Firstly, inhibiting TDP2, while inducing transcription with estradiol, will increase TOP2B-mediated DSBs (compared to estradiol-treated uninhibited cells). By treating these cells with etoposide, the DSBs will be made abortive. The lack of functional TDP2 will lead to delayed repair kinetics, meaning that abortive DSBs will persist long enough to overwhelm the cell and result in permanent cell cycle arrest and cell death. Therefore, I believe that under the right conditions, TDP2 can be therapeutically exploited.

10.8 Overview and further work

In my thesis I showed that inherent TDP2 protein levels or inherent TOP2A/TDP2 and TOP2B/TDP2 protein level ratios do not have the potential for use as predictive biomarkers of etoposide resistance in lung and breast cancer cells. I investigated whether this could be due to the mutational profile of the cell lines I used and I proposed that EGFR, PIK3CA, PTEN, TP53 and SMARCA4 mutations could have contributed to the lack of correlation that I observed. I also investigated the potential of using TDP2 as a biomarker of sensitivity for other anti-cancer drugs. I proposed that a correlation between TDP2 and drug resistance should be explored experimentally for Vinblastine, ABT888, AG014688, Bleomycin, Camptothecin and Bicalutamide. I was not able to show a synergistic effect of treatment with estradiol, PARP1 inhibitor and etoposide when comparing clonogenicity after the combination treatment and treatment with etoposide. Finally, I showed that treatment with estradiol induces a significantly higher number of DSBs (depicted as γH2AX foci per cell) in TDP2-depleted cells; a combined treatment of etoposide and estradiol causes a significantly higher number of DSBs (depicted as γH2AX foci per cell) in the TDP2-depleted cells; and lack of TDP2 causes delayed DSB repair kinetics. Based on these observations I proposed a hypothesis. Treating breast cancer cells lacking TDP2 with estradiol will induce a higher number of transcription-related,

TOP2B-mediated DSBs compared to cells with TDP2. At that stage, treating these cells with etoposide will render this large number of breaks abortive. Finally the lack of functional TDP2 will lead to delayed repair kinetics, meaning that abortive DSBs will persist long enough to overwhelm the cell and result in permanent cell cycle arrest and cell death. Therefore, I believe that under the right conditions, TDP2 can be therapeutically exploited and targeted for inhibition, especially in slow proliferating hormone-dependent cancers such as breast tumours, where resistance has developed or other types of treatment have failed.

I believe the future of therapeutic exploitation of TDP2 lies with the development of a TDP2 targeting agent, in the same manner that one exists for TOP2. Such work would allow successful inhibition of TDP2 in different types of cell lines, which will have an effect on resistance to etoposide; something that I believe has been proven conclusively. The acquisition of such a useful tool could open up the possibility of targeting TDP2 for pharmacological inhibition to potentially increase tumour sensitivity to TOP2 poisons, particularly those that develop resistance during the course of treatment. This of course is the opinion of several research groups too, as there are many encouraging published papers about the future of a successful TDP2 inhibitor.

Toxoflavin derivatives and deazaflavins were identified as sub-micromolar and selective inhibitors of TDP2 via a high-throughput screening; however, even though they both demonstrated a structure-activity relationship, both posed challenges (Raoof et al., 2013). Isoquinoline-1,3-dione compounds were also reported as selective TDP2 inhibitors (Kankanala et al., 2016). Deazaflavin was identified by another study as a good candidate (Marchand et al., 2016). Another paper identified a selective inhibitor via a high-throughput screening and subsequent validation studies, NSC111041 (Kont et al., 2016). A different group reported another collective of TDP2 inhibitors that emerged from experimentally verifying the selective inhibitory activity of relevant molecules from the NCI diversity small molecule database

(Kossmann et al., 2016). Finally, another group confirmed that deazaflavin inhibits TDP2 using a fluorescence-based assay suitable for high-throughput screening, but they went one step further and determined the crystal structures of these compounds bound to a 'humanized' form of murine TDP2. Interestingly, these structures demonstrated that the compounds act as competitive ligands for the binding site of an incoming DNA substrate, something that could potentially point the way towards generating new and effective TDP2 inhibitors (Hornyak et al., 2016).

Administering such an inhibitor to patients might have an effect on cell viability on its own, although additional experiments would need to be designed to specifically explore this. Even though the MCF7 TDP2 knock-down I used was not utilised for comparison of physiological differences compared to the MCF7 mock knock-down in the absence of any treatment, I observed a difference during cell culture of those two cell lines. TDP2-depleted cells' growth rate was slower than the mock-depleted, which could imply that near-absence of TDP2 had an effect on proliferation rates. Furthermore, it is interesting that treatment with estradiol induced a higher number of DSBs (visualised as γH2AX foci) in the TDP2-depleted cells, while induction of transcription of ER responsive genes was not significantly different for the TDP2-depleted and mock-depleted cells, although there is evidence that suggest that it is lower for AR-responsive genes in *TDP2*-depleted cells (Gómez-Herreros et al., 2014). Even though treatment with estradiol alone did not yield different cell survival results between these two cell lines, it is still possible that the introduction of a significantly higher number of DSBs could potentially have therapeutic applications. As previously discussed, perhaps a different TDP2-inhibited cell line in which induction of transcription could be achieved at higher levels could be a good candidate, so that treatment with estradiol would introduce a high enough number of TOP2B-mediated DSBs, which in turn would lead to significantly different results in cell survival.

High-throughput RNA interference (RNAi) screens could be used to identify synthetic lethal interactions with TDP2. The approach that my project had, required considerable infrastructure and a substantial investment for each cell line that I screened. Screens that are designed to be specific to the context of the inactivated contender, in this case TDP2, and specific to treatments previously characterised have been shown to be successful in generating interpretable data and exposing genotype-specific sensitivities (Silva et al., 2008, Ngo et al., 2006, Schlabach et al., 2008).

My project was solely focused on lung and breast cancer cells, as initially the objective was to compare levels of TDP2 in various cancer cells and if these levels were variable enough, to investigate a correlation between TDP2 levels and sensitivity to etoposide, as a clinical predictor for treatment outcome. Using cells derived from normal tissue was not considered as it would not have an impact on the outcome and the conclusions of the findings. One of the hallmarks of cancer is increased proliferation and it was hypothesised that cancer cells would rely heavily on TDP2 for repairing TOP2-mediated DSBs, unlike normal tissue.

As with any kind of cancer treatment though, etoposide causes side effects due to the inability of the treatment to distinguish between healthy cells and cancer cells when targeting TOP2. Targeted therapies though, still strike cancer cells more effectively as the small molecule compounds usually used, target specific proteins that play a key role in cancer cell growth and survival. Inhibiting TDP2 with a drug would potentially have similar side effects, as it would probably target all cells indiscriminately. On the other hand, there is a way of administering such treatments in a more specific manner, by using monoclonal antibodies. Monoclonal antibodies can recognise specific areas unique to cancer cells, bind to them and deliver the small molecule compound directly to the cancer cell without harming the healthy cells of the patient (Scott et al., 2012). Therefore even once an effective TDP2 inhibitor is developed, the compound may need to be optimised with targeted delivery to ensure minimal side effects on

healthy tissue. On the other hand strong synergy with other drugs that are specific to cancer

cells may reduce the detrimental effect on healthy tissue.

Bibliography

Abaan, O. D., Polley, E. C., Davis, S. R., Zhu, Y. J., Bilke, S., Walker, R. L., Pineda, M., Gindin, Y., Jiang, Y., Reinhold, W. C., Holbeck, S. L., Simon, R. M., Doroshow, J. H., Pommier, Y. and Meltzer, P. S. (2013) 'The exomes of the NCI-60 panel: a genomic resource for cancer biology and systems pharmacology', *Cancer Res,* 73(14), pp. 4372-82.

Adachi, N., Suzuki, H., Iiizumi, S. and Koyama, H. (2003) 'Hypersensitivity of nonhomologous DNA end-joining mutants to VP-16 and ICRF-193: implications for the repair of topoisomerase II-mediated DNA damage', *J Biol Chem,* 278(38), pp. 35897-902.

Adhikari, S., Karmahapatra, S. K., Elias, H., Dhopeshwarkar, P., Williams, R. S., Byers, S., Uren, A. and Roy, R. (2011) 'Development of a novel assay for human tyrosyl DNA phosphodiesterase 2', *Analytical Biochemistry,* 416(1), pp. 112-116.

Ahnesorg, P., Smith, P. and Jackson, S. P. (2006) 'XLF interacts with the XRCC4-DNA ligase IV complex to promote DNA nonhomologous end-joining', *Cell,* 124(2), pp. 301-13.

Alpsoy, A., Yasa, S. and Gündüz, U. (2014) 'Etoposide resistance in MCF-7 breast cancer cell line is marked by multiple mechanisms', *Biomed Pharmacother,* 68(3), pp. 351-5.

Asikainen, T. M., Ahmad, A., Schneider, B. K., Ho, W. B., Arend, M., Brenner, M., Günzler, V. and White, C. W. (2005) 'Stimulation of HIF-1alpha, HIF-2alpha, and VEGF by prolyl 4-hydroxylase inhibition in human lung endothelial and epithelial cells', *Free Radic Biol Med,* 38(8), pp. 1002-13.

Audeh, M. W., Carmichael, J., Penson, R. T., Friedlander, M., Powell, B., Bell-McGuinn, K. M., Scott, C., Weitzel, J. N., Oaknin, A., Loman, N., Lu, K., Schmutzler, R. K., Matulonis, U., Wickens, M. and Tutt, A. (2010) 'Oral poly(ADP-ribose) polymerase inhibitor olaparib in patients with BRCA1 or BRCA2 mutations and recurrent ovarian cancer: a proof-of-concept trial', *Lancet,* 376(9737), pp. 245-51.

Azzam, E. I., Jay-Gerin, J. P. and Pain, D. (2012) 'Ionizing radiation-induced metabolic oxidative stress and prolonged cell injury', *Cancer Lett,* 327(1-2), pp. 48-60.

Baird, C. L., Harkins, T. T., Morris, S. K. and Lindsley, J. E. (1999) 'Topoisomerase II drives DNA transport by hydrolyzing one ATP', *Proc Natl Acad Sci U S A,* 96(24), pp. 13685-90.

Banks-Schlegel, S. P., Gazdar, A. F. and Harris, C. C. (1985) 'Intermediate filament and cross-linked envelope expression in human lung tumor cell lines', *Cancer Res,* 45(3), pp. 1187-97.

Baranello, L., Levens, D., Gupta, A. and Kouzine, F. (2012) 'The importance of being supercoiled: how DNA mechanics regulate dynamic processes', *Biochim Biophys Acta,* 1819(7), pp. 632-8.

Bareschino, M. A., Schettino, C., Rossi, A., Maione, P., Sacco, P. C., Zeppa, R. and Gridelli, C. (2011) 'Treatment of advanced non small cell lung cancer', *J Thorac Dis,* 3(2), pp. 122-33.

Barretina, J., Caponigro, G., Stransky, N., Venkatesan, K., Margolin, A. A., Kim, S., Wilson, C. J., Lehár, J., Kryukov, G. V., Sonkin, D., Reddy, A., Liu, M., Murray, L., Berger, M. F., Monahan, J. E., Morais, P., Meltzer, J., Korejwa, A., Jané-Valbuena, J., Mapa, F. A., Thibault, J., Bric-Furlong, E., Raman, P., Shipway, A., Engels, I. H., Cheng, J., Yu, G. K., Yu, J., Aspesi, P., de Silva, M., Jagtap, K., Jones, M. D., Wang, L., Hatton, C., Palescandolo, E., Gupta, S., Mahan, S., Sougnez, C., Onofrio, R. C., Liefeld, T., MacConaill, L., Winckler, W., Reich, M., Li, N., Mesirov, J. P., Gabriel, S. B., Getz, G., Ardlie, K., Chan, V., Myer, V. E., Weber, B. L., Porter, J., Warmuth, M., Finan, P., Harris, J. L., Meyerson, M., Golub, T. R., Morrissey, M. P., Sellers, W. R., Schlegel, R. and Garraway, L. A. (2012) 'The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity', *Nature,* 483(7391), pp. 603-7.

Beckman, K. B. and Ames, B. N. (1997) 'Oxidative decay of DNA', *J Biol Chem,* 272(32), pp. 19633-6.

Benjamin, R. C. and Gill, D. M. (1980) 'Poly(ADP-ribose) synthesis in vitro programmed by damaged DNA. A comparison of DNA molecules containing different types of strand breaks', *J Biol Chem,* 255(21), pp. 10502-8.

Beral, V. (2003) 'Breast cancer and hormone-replacement therapy in the Million Women Study', *Lancet,* 362(9382), pp. 419-27.

Beral, V., Bull, D., Green, J. and Reeves, G. (2007) 'Ovarian cancer and hormone replacement therapy in the Million Women Study', *Lancet,* 369(9574), pp. 1703-10.

Bian, K., Muppani, N. R., Elkhadragy, L., Wang, W., Zhang, C., Chen, T., Jung, S., Seternes, O. M. and Long, W. (2016) 'ERK3 regulates TDP2-mediated DNA damage response and chemoresistance in lung cancer cells', *Oncotarget,* 7(6), pp. 6665-75.

Bolli, N., Avet-Loiseau, H., Wedge, D. C., Van Loo, P., Alexandrov, L. B., Martincorena, I., Dawson, K. J., Iorio, F., Nik-Zainal, S., Bignell, G. R., Hinton, J. W., Li, Y., Tubio, J. M., McLaren, S., O' Meara, S., Butler, A. P., Teague, J. W., Mudie, L., Anderson, E., Rashid, N., Tai, Y. T., Shammas, M. A., Sperling, A. S., Fulciniti, M., Richardson, P. G., Parmigiani, G., Magrangeas, F., Minvielle, S., Moreau, P., Attal, M., Facon, T., Futreal, P. A., Anderson, K. C., Campbell, P. J. and Munshi, N. C. (2014) 'Heterogeneity of genomic evolution and mutational profiles in multiple myeloma', *Nat Commun,* 5, pp. 2997.

Bolton, E. E., Wang, Y., Thiessen, P. A. and Bryant, S. H. (2008) 'Chapter 12 - PubChem: Integrated Platform of Small Molecules and Biological Activities', *Annual Reports in Computational Chemistry,* 4, pp. 217-241.

Bouaoun, L., Sonkin, D., Ardin, M., Hollstein, M., Byrnes, G., Zavadil, J. and Olivier, M. (2016) 'TP53 Variations in Human Cancers: New Lessons from the IARC TP53 Database and Genomics Data', *Hum Mutat,* 37(9), pp. 865-76.

Bowman, K. J., Newell, D. R., Calvert, A. H. and Curtin, N. J. (2001) 'Differential effects of the poly (ADP-ribose) polymerase (PARP) inhibitor NU1025 on topoisomerase I and II inhibitor cytotoxicity in L1210 cells in vitro', *Br J Cancer,* 84(1), pp. 106-12.

Bramson, J., Prévost, J., Malapetsa, A., Noë, A. J., Poirier, G. G., DesNoyers, S., Alaoui-Jamali, M. and Panasci, L. (1993) 'Poly(ADP-ribose) polymerase can bind melphalan damaged DNA', *Cancer Res,* 53(22), pp. 5370-3.

Breathnach, O. S., Freidlin, B., Conley, B., Green, M. R., Johnson, D. H., Gandara, D. R., O'Connell, M., Shepherd, F. A. and Johnson, B. E. (2001) 'Twenty-two years of phase III trials for patients with advanced non-small-cell lung cancer: sobering results', *J Clin Oncol,* 19(6), pp. 1734-42.

Bromberg, K. D., Burgin, A. B. and Osheroff, N. (2003) 'A two-drug model for etoposide action against human topoisomerase IIalpha', *J Biol Chem,* 278(9), pp. 7406-12.

Brummelkamp, T. R., Bernards, R. and Agami, R. (2002) 'A system for stable expression of short interfering RNAs in mammalian cells', *Science,* 296(5567), pp. 550-3.

Brunner, N., Boulay, V., Fojo, A., Freter, C. E., Lippman, M. E. and Clarke, R. (1993a) 'Acquisition of hormone-independent growth in MCF-7 cells is accompanied by increased expression of estrogen-regulated genes but without detectable DNA amplifications', *Cancer Res,* 53(2), pp. 283-90.

Brunner, N., Frandsen, T. L., Holst-Hansen, C., Bei, M., Thompson, E. W., Wakeling, A. E., Lippman, M. E. and Clarke, R. (1993b) 'MCF7/LCC2: a 4-hydroxytamoxifen resistant human breast cancer variant that retains sensitivity to the steroidal antiestrogen ICI 182,780', *Cancer Res,* 53(14), pp. 3229-32.

Bryant, H. E., Schultz, N., Thomas, H. D., Parker, K. M., Flower, D., Lopez, E., Kyle, S., Meuth, M., Curtin, N. J. and Helleday, T. (2005) 'Specific killing of BRCA2-deficient tumours with inhibitors of poly(ADP-ribose) polymerase', *Nature,* 434(7035), pp. 913-7.

Brünner, N., Boysen, B., Jirus, S., Skaar, T. C., Holst-Hansen, C., Lippman, J., Frandsen, T., Spang-Thomsen, M., Fuqua, S. A. and Clarke, R. (1997) 'MCF7/LCC9: an antiestrogen-resistant MCF-7 variant in which acquired resistance to the steroidal antiestrogen ICI 182,780 confers an early cross-resistance to the nonsteroidal antiestrogen tamoxifen', *Cancer Res,* 57(16), pp. 3486-93.

Cadet, J., Douki, T., Gasparutto, D. and Ravanat, J. L. (2003) 'Oxidative damage to DNA: formation, measurement and biochemical features', *Mutat Res,* 531(1-2), pp. 5-23.

Cailleau, R., Olive, M. and Cruciger, Q. V. (1978) 'Long-term human breast carcinoma cell lines of metastatic origin: preliminary characterization', *In Vitro,* 14(11), pp. 911-5.

Caldecott, K. W. (2001) 'Mammalian DNA single-strand break repair: an X-ra(y)ted affair', *Bioessays,* 23(5), pp. 447-55.

Caldecott, K. W. (2008) 'Single-strand break repair and genetic disease', *Nat Rev Genet,* 9(8), pp. 619-31.

Caldecott, K. W., Aoufouchi, S., Johnson, P. and Shall, S. (1996) 'XRCC1 polypeptide interacts with DNA polymerase beta and possibly poly (ADP-ribose) polymerase, and DNA ligase III is a novel molecular 'nick-sensor' in vitro', *Nucleic Acids Res,* 24(22), pp. 4387-94.

Cannan, W. J. and Pederson, D. S. (2016) 'Mechanisms and Consequences of Double-Strand DNA Break Formation in Chromatin', *J Cell Physiol,* 231(1), pp. 3-14.

Carboni, S., Hiver, A., Szyndralewiez, C., Gaillard, P., Gotteland, J. P. and Vitte, P. A. (2004) 'AS601245 (1,3-benzothiazol-2-yl (2-[[2-(3-pyridinyl) ethyl] amino]-4 pyrimidinyl) acetonitrile): a c-Jun NH2-terminal protein kinase inhibitor with neuroprotective properties', *J Pharmacol Exp Ther,* 310(1), pp. 25-32.

Carney, D. N., Gazdar, A. F., Bepler, G., Guccion, J. G., Marangos, P. J., Moody, T. W., Zweig, M. H. and Minna, J. D. (1985) 'Establishment and identification of small cell lung cancer cell lines having classic and variant features', *Cancer Res,* 45(6), pp. 2913-23.

Cascón, A. and Robledo, M. (2012) 'MAX and MYC: a heritable breakup', *Cancer Res,* 72(13), pp. 3119-24.

Champoux, J. J. (2001) 'DNA topoisomerases: structure, function, and mechanism', *Annu Rev Biochem,* 70, pp. 369-413.

Chan, N., Pires, I. M., Bencokova, Z., Coackley, C., Luoto, K. R., Bhogal, N., Lakshman, M., Gottipati, P., Oliver, F. J., Helleday, T., Hammond, E. M. and Bristow, R. G. (2010) 'Contextual synthetic lethality of cancer cell kill based on the tumor microenvironment', *Cancer Res,* 70(20), pp. 8045-54.

Chang, H. H., Watanabe, G. and Lieber, M. R. (2015) 'Unifying the DNA end-processing roles of the artemis nuclease: Ku-dependent artemis resection at blunt DNA ends', *J Biol Chem,* 290(40), pp. 24036-50.

Chang, L. C., Chiu, H. M., Shun, C. T., Liang, J. T., Lin, J. T., Chen, C. C., Lee, Y. C. and Wu, M. S. (2014) 'Mutational profiles of different macroscopic subtypes of colorectal adenoma reveal distinct pathogenetic roles for KRAS, BRAF and PIK3CA', *BMC Gastroenterol,* 14, pp. 221.

Chen, L., Trujillo, K., Sung, P. and Tomkinson, A. E. (2000) 'Interactions of the DNA ligase IV-XRCC4 complex with DNA ends and the DNA-dependent protein kinase', *J Biol Chem,* 275(34), pp. 26196-205.

Cheng, K. C., Cahill, D. S., Kasai, H., Nishimura, S. and Loeb, L. A. (1992) '8-Hydroxyguanine, an abundant form of oxidative DNA damage, causes G----T and A----C substitutions', *J Biol Chem,* 267(1), pp. 166-72.

Clancy, S. (2008) 'DNA damage & repair: mechanisms for maintaining DNA integrity.', *Nature Education,* 1(1).

Clark, K., Plater, L., Peggie, M. and Cohen, P. (2009) 'Use of the pharmacological inhibitor BX795 to study the regulation and physiological roles of TBK1 and IkappaB kinase epsilon: a distinct upstream kinase mediates Ser-172 phosphorylation and activation', *J Biol Chem,* 284(21), pp. 14136-46.

Comsa, S., Cimpean, A. M. and Raica, M. (2015) 'The Story of MCF-7 Breast Cancer Cell Line: 40 years of Experience in Research', *Anticancer Res,* 35(6), pp. 3147-54.

Cooke, M. S., Evans, M. D., Dizdaroglu, M. and Lunec, J. (2003) 'Oxidative DNA damage: mechanisms, mutation, and disease', *Faseb j,* 17(10), pp. 1195-214.

Cortes Ledesma, F., El Khamisy, S. F., Zuma, M. C., Osborn, K. and Caldecott, K. W. (2009) 'A human 5'-tyrosyl DNA phosphodiesterase that repairs topoisomerase-mediated DNA damage', *Nature,* 461(7264), pp. 674-8.

Cortes, U., Tong, W. M., Coyle, D. L., Meyer-Ficca, M. L., Meyer, R. G., Petrilli, V., Herceg, Z., Jacobson, E. L., Jacobson, M. K. and Wang, Z. Q. (2004) 'Depletion of the 110-kilodalton isoform of poly(ADP-ribose) glycohydrolase increases sensitivity to genotoxic and endotoxic stress in mice', *Mol Cell Biol,* 24(16), pp. 7163-78.

Coulter, J. B., M.C., H., Mian, O. Y., Hedayati, M., Zhang, Y., Zhou, H., Nelso, W., Yegnasubramanian, S. and DeWeese, T. L. (2015) 'Generation of DNA Double Strand Breaks in Prostate Cancer Cells by Androgen Receptor Antagonists, Hydroxyflutamide and Bicalutamide', *International Journal of Radiation Oncology*Biology*Physics,* 93(3), pp. S185.

Cowell, I. G., Sondka, Z., Smith, K., Lee, K. C., Manville, C. M., Sidorczuk-Lesthuruge, M., Rance, H. A., Padget, K., Jackson, G. H., Adachi, N. and Austin, C. A. (2012) 'Model for MLL translocations in therapy-related leukemia involving topoisomerase IIbeta-mediated DNA strand breaks and gene proximity', *Proc Natl Acad Sci U S A,* 109(23), pp. 8989-94.

Critchlow, S. E., Bowater, R. P. and Jackson, S. P. (1997) 'Mammalian DNA double-strand break repair protein XRCC4 interacts with DNA ligase IV', *Curr Biol,* 7(8), pp. 588-98.

Curtin, N. J. (2005) 'PARP inhibitors for cancer therapy', *Expert Rev Mol Med,* 7(4), pp. 1-20.

D'Amours, D., Desnoyers, S., D'Silva, I. and Poirier, G. G. (1999) 'Poly(ADP-ribosyl)ation reactions in the regulation of nuclear functions', *Biochem J,* 342 (Pt 2), pp. 249-68.

Daniel, D. and Lalitha, R. M. (2016) 'Tumor markers - A bird's eye view', *Journal of Oral and Maxillofacial Surgery, Medicine, and Pathology,* 28(6), pp. 475-480.

Davis, A. J., Chen, B. P. and Chen, D. J. (2014) 'DNA-PK: a dynamic enzyme in a versatile DSB repair pathway', *DNA Repair (Amst),* 17, pp. 21-9.

de Boer, H. R., Guerrero Llobet, S. and van Vugt, M. A. (2016) 'Controlling the response to DNA damage by the APC/C-Cdh1', *Cell Mol Life Sci,* 73(5), pp. 949-60.

De Ioannes, P., Malu, S., Cortes, P. and Aggarwal, A. K. (2012) 'Structural basis of DNA ligase IV-Artemis interaction in nonhomologous end-joining', *Cell Rep,* 2(6), pp. 1505-12.

Delacôte, F. and Lopez, B. S. (2008) 'Importance of the cell cycle phase for the choice of the appropriate DSB repair pathway, for genome stability maintenance: the trans-S double-strand break repair model', *Cell Cycle,* 7(1), pp. 33-8.

Demple, B. and DeMott, M. S. (2002) 'Dynamics and diversions in base excision DNA repair of oxidized abasic lesions', *Oncogene,* 21(58), pp. 8926-34.

Deweese, J. E. and Osheroff, N. (2009) 'The DNA cleavage reaction of topoisomerase II: wolf in sheep's clothing', *Nucleic Acids Res,* 37(3), pp. 738-48.

Deweese, J. E. and Osheroff, N. (2010) 'The use of divalent metal ions by type II topoisomerases', *Metallomics,* 2(7), pp. 450-9.

Dianov, G. L. and Hübscher, U. (2013) 'Mammalian base excision repair: the forgotten archangel', *Nucleic Acids Res,* 41(6), pp. 3483-90.

Dickson, B. C., Mulligan, A. M., Zhang, H., Lockwood, G., O'Malley, F. P., Egan, S. E. and Reedijk, M. (2007) 'High-level JAG1 mRNA and protein predict poor outcome in breast cancer', *Mod Pathol,* 20(6), pp. 685-93.

Ditchfield, C., Johnson, V. L., Tighe, A., Ellston, R., Haworth, C., Johnson, T., Mortlock, A., Keen, N. and Taylor, S. S. (2003) 'Aurora B couples chromosome alignment with anaphase by targeting BubR1, Mad2, and Cenp-E to kinetochores', *J Cell Biol,* 161(2), pp. 267-80.

Do, P. M., Varanasi, L., Fan, S., Li, C., Kubacka, I., Newman, V., Chauhan, K., Daniels, S. R., Boccetta, M., Garrett, M. R., Li, R. and Martinez, L. A. (2012) 'Mutant p53 cooperates with ETS2 to promote etoposide resistance', *Genes Dev,* 26(8), pp. 830-45.

Donawho, C. K., Luo, Y., Penning, T. D., Bauch, J. L., Bouska, J. J., Bontcheva-Diaz, V. D., Cox, B. F., DeWeese, T. L., Dillehay, L. E., Ferguson, D. C., Ghoreishi-Haack, N. S., Grimm, D. R., Guan, R., Han, E. K., Holley-Shanks, R. R., Hristov, B., Idler, K. B., Jarvis, K., Johnson, E. F., Kleinberg, L. R., Klinghofer, V., Lasko, L. M., Liu, X., Marsh, K. C., McGonigal, T. P., Meulbroek, J. A., Olson, A. M., Palma, J. P., Rodriguez, L. E., Shi, Y., Stavropoulos, J. A., Tsurutani, A. C., Zhu, G. D., Rosenberg, S. H., Giranda, V. L. and Frost, D. J. (2007) 'ABT-888, an orally active poly(ADP-ribose) polymerase inhibitor that potentiates DNA-damaging agents in preclinical tumor models', *Clin Cancer Res,* 13(9), pp. 2728-37.

Durkacz, B. W., Omidiji, O., Gray, D. A. and Shall, S. (1980) '(ADP-ribose)n participates in DNA excision repair', *Nature,* 283(5747), pp. 593-6.

Eguren, M., Álvarez-Fernández, M., García, F., López-Contreras, A. J., Fujimitsu, K., Yaguchi, H., Luque-García, J. L., Fernández-Capetillo, O., Muñoz, J., Yamano, H. and Malumbres, M. (2014) 'A synthetic lethal interaction between APC/C and topoisomerase poisons uncovered by proteomic screens', *Cell Rep,* 6(4), pp. 670-83.

El-Khamisy, S. F., Masutani, M., Suzuki, H. and Caldecott, K. W. (2003) 'A requirement for PARP-1 for the assembly or stability of XRCC1 nuclear foci at sites of oxidative DNA damage', *Nucleic Acids Res,* 31(19), pp. 5526-33.

El-Khamisy, S. F., Saifi, G. M., Weinfeld, M., Johansson, F., Helleday, T., Lupski, J. R. and Caldecott, K. W. (2005) 'Defective DNA single-strand break repair in spinocerebellar ataxia with axonal neuropathy-1', *Nature,* 434(7029), pp. 108-13.

End, D. W., Smets, G., Todd, A. V., Applegate, T. L., Fuery, C. J., Angibaud, P., Venet, M., Sanz, G., Poignet, H., Skrzat, S., Devine, A., Wouters, W. and Bowden, C. (2001) 'Characterization of the antitumor effects of the selective farnesyl protein transferase inhibitor R115777 in vivo and in vitro', *Cancer Res,* 61(1), pp. 131-7.

Engel, L. W., Young, N. A., Tralka, T. S., Lippman, M. E., O'Brien, S. J. and Joyce, M. J. (1978) 'Establishment and characterization of three new continuous cell lines derived from human breast carcinomas', *Cancer Res,* 38(10), pp. 3352-64.

Farmer, H., McCabe, N., Lord, C. J., Tutt, A. N., Johnson, D. A., Richardson, T. B., Santarosa, M., Dillon, K. J., Hickson, I., Knights, C., Martin, N. M., Jackson, S. P., Smith, G. C. and Ashworth, A. (2005) 'Targeting the DNA repair defect in BRCA mutant cells as a therapeutic strategy', *Nature,* 434(7035), pp. 917-21.

Ferro, A. M. and Olivera, B. M. (1982) 'Poly(ADP-ribosylation) in vitro. Reaction parameters and enzyme mechanism', *J Biol Chem,* 257(13), pp. 7808-13.

Fisher, A. E., Hochegger, H., Takeda, S. and Caldecott, K. W. (2007) 'Poly(ADP-ribose) polymerase 1 accelerates single-strand break repair in concert with poly(ADP-ribose) glycohydrolase', *Mol Cell Biol,* 27(15), pp. 5597-605.

Fleming, R. A., Miller, A. A. and Stewart, C. F. (1989) 'Etoposide: an update', *Clin Pharm,* 8(4), pp. 274-93.

Fong, P. C., Boss, D. S., Yap, T. A., Tutt, A., Wu, P., Mergui-Roelvink, M., Mortimer, P., Swaisland, H., Lau, A., O'Connor, M. J., Ashworth, A., Carmichael, J., Kaye, S. B., Schellens, J. H. and de Bono, J. S. (2009) 'Inhibition of poly(ADP-ribose) polymerase in tumors from BRCA mutation carriers', *N Engl J Med,* 361(2), pp. 123-34.

Forbes, S. A., Beare, D., Gunasekaran, P., Leung, K., Bindal, N., Boutselakis, H., Ding, M., Bamford, S., Cole, C., Ward, S., Kok, C. Y., Jia, M., De, T., Teague, J. W., Stratton, M. R., McDermott, U. and Campbell, P. J. (2015) 'COSMIC: exploring the world's knowledge of somatic mutations in human cancer', *Nucleic Acids Res,* 43(Database issue), pp. D805-11.

Forget, A. L. and Kowalczykowski, S. C. (2010) 'Single-molecule imaging brings Rad51 nucleoprotein filaments into focus', *Trends Cell Biol,* 20(5), pp. 269-76.

Fountzilas, G., Christodoulou, C., Bobos, M., Kotoula, V., Eleftheraki, A. G., Xanthakis, I., Batistatou, A., Pentheroudakis, G., Xiros, N., Papaspirou, I., Koumarianou, A., Papakostas, P., Bafaloukos, D., Skarlos, D. V. and Kalogeras, K. T. (2012) 'Topoisomerase II alpha gene amplification is a favorable prognostic factor in patients with HER2-positive metastatic breast cancer treated with trastuzumab', *J Transl Med,* 10, pp. 212.

Frampton, J. E. (2015) 'Olaparib: a review of its use as maintenance therapy in patients with ovarian cancer', *BioDrugs,* 29(2), pp. 143-50.

Fry, D. W., Harvey, P. J., Keller, P. R., Elliott, W. L., Meade, M., Trachet, E., Albassam, M., Zheng, X., Leopold, W. R., Pryer, N. K. and Toogood, P. L. (2004) 'Specific inhibition of cyclin-dependent kinase 4/6 by PD 0332991 and associated antitumor activity in human tumor xenografts', *Mol Cancer Ther,* 3(11), pp. 1427-38.

Furr, B. J., Valcaccia, B., Curry, B., Woodburn, J. R., Chesterson, G. and Tucker, H. (1987) 'ICI 176,334: a novel non-steroidal, peripherally selective antiandrogen', *J Endocrinol,* 113(3), pp. R7-9.

Futreal, P. A., Coin, L., Marshall, M., Down, T., Hubbard, T., Wooster, R., Rahman, N. and Stratton, M. R. (2004) 'A census of human cancer genes', *Nat Rev Cancer,* 4(3), pp. 177-83.

Gao, R., Huang, S. Y., Marchand, C. and Pommier, Y. (2012) 'Biochemical characterization of human tyrosyl-DNA phosphodiesterase 2 (TDP2/TTRAP): a Mg(2+)/Mn(2+)-dependent phosphodiesterase specific for the repair of topoisomerase cleavage complexes', *J Biol Chem,* 287(36), pp. 30842-52.

García-Echeverría, C., Pearson, M. A., Marti, A., Meyer, T., Mestan, J., Zimmermann, J., Gao, J., Brueggen, J., Capraro, H. G., Cozens, R., Evans, D. B., Fabbro, D., Furet, P., Porta, D. G., Liebetanz, J., Martiny-Baron, G., Ruetz, S. and Hofmann, F. (2004) 'In vivo antitumor activity of NVP-AEW541-A novel, potent, and selective inhibitor of the IGF-IR kinase', *Cancer Cell,* 5(3), pp. 231-9.

Garnett, M. J., Edelman, E. J., Heidorn, S. J., Greenman, C. D., Dastur, A., Lau, K. W., Greninger, P., Thompson, I. R., Luo, X., Soares, J., Liu, Q., Iorio, F., Surdez, D., Chen, L., Milano, R. J., Bignell, G. R., Tam, A. T., Davies, H., Stevenson, J. A., Barthorpe, S., Lutz, S. R., Kogera, F., Lawrence, K., McLaren-Douglas, A., Mitropoulos, X., Mironenko, T., Thi, H., Richardson, L., Zhou, W., Jewitt, F., Zhang, T., O'Brien, P., Boisvert, J. L., Price, S., Hur, W., Yang, W., Deng, X., Butler, A., Choi, H. G., Chang, J. W., Baselga, J., Stamenkovic, I., Engelman, J. A., Sharma, S. V., Delattre, O., Saez-Rodriguez, J., Gray, N. S., Settleman, J., Futreal, P. A., Haber, D. A., Stratton, M. R., Ramaswamy, S., McDermott, U. and Benes, C. H. (2012) 'Systematic identification of genomic markers of drug sensitivity in cancer cells', *Nature,* 483(7391), pp. 570-5.

Gazdar, A. e. a. (1996) 'NCI-Navy Medical Oncology Branch cell line supplement', *J Cell Biochem Suppl,* 24, pp. 1-291.

Germain, M., Affar, E. B., D'Amours, D., Dixit, V. M., Salvesen, G. S. and Poirier, G. G. (1999) 'Cleavage of automodified poly(ADP-ribose) polymerase during apoptosis. Evidence for involvement of caspase-7', *J Biol Chem,* 274(40), pp. 28379-84.

Germe, T., Miller, K. and Cooper, J. P. (2009) 'A non-canonical function of topoisomerase II in disentangling dysfunctional telomeres', *Embo j,* 28(18), pp. 2803-11.

Ghosh, M. G., Thompson, D. A. and Weigel, R. J. (2000) 'PDZK1 and GREB1 are estrogen-regulated genes expressed in hormone-responsive breast cancer', *Cancer Res,* 60(22), pp. 6367-75.

Giard, D. J., Aaronson, S. A., Todaro, G. J., Arnstein, P., Kersey, J. H., Dosik, H. and Parks, W. P. (1973) 'In vitro cultivation of human tumors: establishment of cell lines derived from a series of solid tumors', *J Natl Cancer Inst,* 51(5), pp. 1417-23.

Gillman, K. W., Starrett, J. E., Parker, M. F., Xie, K., Bronson, J. J., Marcin, L. R., McElhone, K. E., Bergstrom, C. P., Mate, R. A., Williams, R., Meredith, J. E., Burton, C. R., Barten, D. M.,

Toyn, J. H., Roberts, S. B., Lentz, K. A., Houston, J. G., Zaczek, R., Albright, C. F., Decicco, C. P., Macor, J. E. and Olson, R. E. (2010) 'Discovery and Evaluation of BMS-708163, a Potent, Selective and Orally Bioavailable γ-Secretase Inhibitor', *ACS Med Chem Lett,* 1(3), pp. 120-4.

Godon, C., Cordelières, F. P., Biard, D., Giocanti, N., Mégnin-Chanet, F., Hall, J. and Favaudon, V. (2008) 'PARP inhibition versus PARP-1 silencing: different outcomes in terms of single-strand break repair and radiation susceptibility', *Nucleic Acids Res,* 36(13), pp. 4454-64.

Goh, H. S., Yao, J. and Smith, D. R. (1995) 'p53 point mutation and survival in colorectal cancer patients', *Cancer Res,* 55(22), pp. 5217-21.

Gomez-Herreros, F., Romero-Granados, R., Zeng, Z., Alvarez-Quilon, A., Quintero, C., Ju, L., Umans, L., Vermeire, L., Huylebroeck, D., Caldecott, K. W. and Cortes-Ledesma, F. (2013) 'TDP2-dependent non-homologous end-joining protects against topoisomerase II-induced DNA breaks and genome instability in cells and in vivo', *PLoS Genet,* 9(3), pp. e1003226.

Govindan, R. and Weber, J. (2014) 'TP53 mutations and lung cancer: not all mutations are created equal', *Clin Cancer Res,* 20(17), pp. 4419-21.

Grawunder, U., Wilm, M., Wu, X., Kulesza, P., Wilson, T. E., Mann, M. and Lieber, M. R. (1997) 'Activity of DNA ligase IV stimulated by complex formation with XRCC4 protein in mammalian cells', *Nature,* 388(6641), pp. 492-5.

Grawunder, U., Zimmer, D. and Lieber, M. R. (1998) 'DNA ligase IV binds to XRCC4 via a motif located between rather than within its BRCT domains', *Curr Biol,* 8(15), pp. 873-6.

Grue, P., Grasser, A., Sehested, M., Jensen, P. B., Uhse, A., Straub, T., Ness, W. and Boege, F. (1998) 'Essential mitotic functions of DNA topoisomerase IIalpha are not adopted by topoisomerase IIbeta in human H69 cells', *J Biol Chem,* 273(50), pp. 33660-6.

Grundy, G. J., Rulten, S. L., Arribas-Bosacoma, R., Davidson, K., Kozik, Z., Oliver, A. W., Pearl, L. H. and Caldecott, K. W. (2016) 'The Ku-binding motif is a conserved module for recruitment and stimulation of non-homologous end-joining proteins', *Nat Commun,* 7, pp. 11242.

Gu, J., Lu, H., Tsai, A. G., Schwarz, K. and Lieber, M. R. (2007) 'Single-stranded DNA ligation and XLF-stimulated incompatible DNA end ligation by the XRCC4-DNA ligase IV complex: influence of terminal DNA sequence', *Nucleic Acids Res,* 35(17), pp. 5755-62.

Gupta, A., Patnaik, M. M. and Naina, H. V. (2014) 'MYST3/CREBBP Rearranged Acute Myeloid Leukemia after Adjuvant Chemotherapy for Breast Cancer', *Case Rep Oncol Med,* 2014, pp. 361748.

Gómez-Herreros, F., Schuurs-Hoeijmakers, J. H., McCormack, M., Greally, M. T., Rulten, S., Romero-Granados, R., Counihan, T. J., Chaila, E., Conroy, J., Ennis, S., Delanty, N., Cortés-Ledesma, F., de Brouwer, A. P., Cavalleri, G. L., El-Khamisy, S. F., de Vries, B. B. and Caldecott, K. W. (2014) 'TDP2 protects transcription from abortive topoisomerase activity and is required for normal neural function', *Nat Genet,* 46(5), pp. 516-21.

Haffner, M. C., Aryee, M. J., Toubaji, A., Esopi, D. M., Albadine, R., Gurel, B., Isaacs, W. B., Bova, G. S., Liu, W., Xu, J., Meeker, A. K., Netto, G., De Marzo, A. M., Nelson, W. G. and Yegnasubramanian, S. (2010) 'Androgen-induced TOP2B-mediated double-strand breaks and prostate cancer gene rearrangements', *Nat Genet,* 42(8), pp. 668-75.

Haffner, M. C., De Marzo, A. M., Meeker, A. K., Nelson, W. G. and Yegnasubramanian, S. (2011) 'Transcription-induced DNA double strand breaks: both oncogenic force and potential therapeutic target?', *Clin Cancer Res,* 17(12), pp. 3858-64.

Hahne, J. C., Okuducu, A. F., Sahin, A., Fafeur, V., Kiriakidis, S. and Wernert, N. (2008) 'The transcription factor ETS-1: its role in tumour development and strategies for its inhibition', *Mini Rev Med Chem,* 8(11), pp. 1095-105.

Hakii, H., Fujiki, H., Suganuma, M., Nakayasu, M., Tahira, T., Sugimura, T., Scheuer, P. J. and Christensen, S. B. (1986) 'Thapsigargin, a histamine secretagogue, is a non-12-O-

tetradecanoylphorbol-13-acetate (TPA) type tumor promoter in two-stage mouse skin carcinogenesis', *J Cancer Res Clin Oncol,* 111(3), pp. 177-81.

Hammel, M., Rey, M., Yu, Y., Mani, R. S., Classen, S., Liu, M., Pique, M. E., Fang, S., Mahaney, B. L., Weinfeld, M., Schriemer, D. C., Lees-Miller, S. P. and Tainer, J. A. (2011) 'XRCC4 protein interactions with XRCC4-like factor (XLF) create an extended grooved scaffold for DNA ligation and double strand break repair', *J Biol Chem,* 286(37), pp. 32638-50.

Hande, K. R. (1998) 'Etoposide: four decades of development of a topoisomerase II inhibitor', *Eur J Cancer,* 34(10), pp. 1514-21.

Hartlerode, A. J. and Scully, R. (2009) 'Mechanisms of double-strand break repair in somatic mammalian cells', *Biochem J,* 423(2), pp. 157-68.

Hasinoff, B. B., Wu, X., Nitiss, J. L., Kanagasabai, R. and Yalowich, J. C. (2012) 'The anticancer multi-kinase inhibitor dovitinib also targets topoisomerase I and topoisomerase II', *Biochem Pharmacol,* 84(12), pp. 1617-26.

Heck, M. M. and Earnshaw, W. C. (1986) 'Topoisomerase II: A specific marker for cell proliferation', *J Cell Biol,* 103(6 Pt 2), pp. 2569-81.

Heck, M. M., Hittelman, W. N. and Earnshaw, W. C. (1988) 'Differential expression of DNA topoisomerases I and II during the eukaryotic cell cycle', *Proc Natl Acad Sci U S A,* 85(4), pp. 1086-90.

Heeres, J. T. and Hergenrother, P. J. (2007) 'Poly(ADP-ribose) makes a date with death', *Curr Opin Chem Biol,* 11(6), pp. 644-53.

Hegde, M. L., Hazra, T. K. and Mitra, S. (2008) 'Early steps in the DNA base excision/single-strand interruption repair pathway in mammalian cells', *Cell Res,* 18(1), pp. 27-47.

Hoa, N. N., Shimizu, T., Zhou, Z. W., Wang, Z. Q., Deshpande, R. A., Paull, T. T., Akter, S., Tsuda, M., Furuta, R., Tsusui, K., Takeda, S. and Sasanuma, H. (2016) 'Mre11 Is Essential for the Removal of Lethal Topoisomerase 2 Covalent Cleavage Complexes', *Mol Cell,* 64(3), pp. 580-592.

Hoeijmakers, J. H. (2001) 'Genome maintenance mechanisms for preventing cancer', *Nature,* 411(6835), pp. 366-74.

Hornyak, P., Askwith, T., Walker, S., Komulainen, E., Paradowski, M., Pennicott, L. E., Bartlett, E. J., Brissett, N. C., Raoof, A., Watson, M., Jordan, A. M., Ogilvie, D. J., Ward, S. E., Atack, J. R., Pearl, L. H., Caldecott, K. W. and Oliver, A. W. (2016) 'Mode of action of DNA-competitive small molecule inhibitors of tyrosyl DNA phosphodiesterase 2', *Biochem J,* 473(13), pp. 1869-79.

Hosoya, N. and Miyagawa, K. (2014) 'Targeting DNA damage response in cancer therapy', *Cancer Sci,* 105(4), pp. 370-88.

Hsiang, Y. H., Wu, H. Y. and Liu, L. F. (1988) 'Proliferation-dependent regulation of DNA topoisomerase II in cultured human cells', *Cancer Res,* 48(11), pp. 3230-5.

Hsieh, P. and Yamane, K. (2008) 'DNA mismatch repair: molecular mechanism, cancer, and ageing', *Mech Ageing Dev,* 129(7-8), pp. 391-407.

Huertas, P. and Jackson, S. P. (2009) 'Human CtIP mediates cell cycle control of DNA end resection and double strand break repair', *J Biol Chem,* 284(14), pp. 9558-65.

Hughes, P., Marshall, D., Reid, Y., Parkes, H. and Gelber, C. (2007) 'The costs of using unauthenticated, over-passaged cell lines: how much more data do we need?', *Biotechniques,* 43(5), pp. 575, 577-8, 581-2 passim.

Iles, N., Rulten, S., El-Khamisy, S. F. and Caldecott, K. W. (2007) 'APLF (C2orf13) is a novel human protein involved in the cellular response to chromosomal DNA strand breaks', *Mol Cell Biol,* 27(10), pp. 3793-803.

Ingle, J. N. (2002) 'Estrogen as therapy for breast cancer', *Breast Cancer Res,* 4(4), pp. 133-6.

Iverson, C., Larson, G., Lai, C., Yeh, L. T., Dadson, C., Weingarten, P., Appleby, T., Vo, T., Maderna, A., Vernier, J. M., Hamatake, R., Miner, J. N. and Quart, B. (2009) 'RDEA119/BAY 869766: a potent, selective, allosteric inhibitor of MEK1/2 for the treatment of cancer', *Cancer Res,* 69(17), pp. 6839-47.

Jackson, S. P. and Bartek, J. (2009) 'The DNA-damage response in human biology and disease', *Nature,* 461(7267), pp. 1071-8.

Jordan, M. A. and Wilson, L. (2004) 'Microtubules as a target for anticancer drugs', *Nat Rev Cancer,* 4(4), pp. 253-65.

Ju, B. G., Lunyak, V. V., Perissi, V., Garcia-Bassets, I., Rose, D. W., Glass, C. K. and Rosenfeld, M. G. (2006) 'A topoisomerase IIbeta-mediated dsDNA break required for regulated transcription', *Science,* 312(5781), pp. 1798-802.

Ju, B. G. and Rosenfeld, M. G. (2006) 'A breaking strategy for topoisomerase IIbeta/PARP-1-dependent regulated transcription', *Cell Cycle,* 5(22), pp. 2557-60.

Kamileri, I., Karakasilioti, I. and Garinis, G. A. (2012) 'Nucleotide excision repair: new tricks with old bricks', *Trends Genet,* 28(11), pp. 566-73.

Kankanala, J., Marchand, C., Abdelmalak, M., Aihara, H., Pommier, Y. and Wang, Z. (2016) 'Isoquinoline-1,3-diones as Selective Inhibitors of Tyrosyl DNA Phosphodiesterase II (TDP2)', *J Med Chem,* 59(6), pp. 2734-46.

Kastan, M. B. and Bartek, J. (2004) 'Cell-cycle checkpoints and cancer', *Nature,* 432(7015), pp. 316-23.

Kechagioglou, P., Papi, R. M., Provatopoulou, X., Kalogera, E., Papadimitriou, E., Grigoropoulos, P., Nonni, A., Zografos, G., Kyriakidis, D. A. and Gounaris, A. (2014) 'Tumor suppressor PTEN in breast cancer: heterozygosity, mutations and protein expression', *Anticancer Res,* 34(3), pp. 1387-400.

Keydar, I., Chen, L., Karby, S., Weiss, F. R., Delarea, J., Radu, M., Chaitcik, S. and Brenner, H. J. (1979) 'Establishment and characterization of a cell line of human breast carcinoma origin', *Eur J Cancer,* 15(5), pp. 659-70.

Khanna, K. K. and Jackson, S. P. (2001) 'DNA double-strand breaks: signaling, repair and the cancer connection', *Nat Genet,* 27(3), pp. 247-54.

Kim, J., Petz, L. N., Ziegler, Y. S., Wood, J. R., Potthoff, S. J. and Nardulli, A. M. (2000) 'Regulation of the estrogen-responsive pS2 gene in MCF-7 human breast cancer cells', *J Steroid Biochem Mol Biol,* 74(4), pp. 157-68.

Kim, M. Y., Zhang, T. and Kraus, W. L. (2005) 'Poly(ADP-ribosyl)ation by PARP-1: 'PAR-laying' NAD+ into a nuclear signal', *Genes Dev,* 19(17), pp. 1951-67.

Kimura, K., Saijo, M., Ui, M. and Enomoto, T. (1994) 'Growth state- and cell cycle-dependent fluctuation in the expression of two forms of DNA topoisomerase II and possible specific modification of the higher molecular weight form in the M phase', *J Biol Chem,* 269(2), pp. 1173-6.

Kont, Y. S., Dutta, A., Mallisetty, A., Mathew, J., Minas, T., Kraus, C., Dhopeshwarkar, P., Kallakury, B., Mitra, S., Üren, A. and Adhikari, S. (2016) 'Depletion of tyrosyl DNA phosphodiesterase 2 activity enhances etoposide-mediated double-strand break formation and cell killing', *DNA Repair (Amst),* 43, pp. 38-47.

Korwek, Z., Sewastianik, T., Bielak-Zmijewska, A., Mosieniak, G., Alster, O., Moreno-Villanueva, M., Moreno-Villaneuva, M., Burkle, A. and Sikora, E. (2012) 'Inhibition of ATM blocks the etoposide-induced DNA damage response and apoptosis of resting human T cells', *DNA Repair (Amst),* 11(11), pp. 864-73.

Kossmann, B. R., Abdelmalak, M., Lopez, S., Tender, G., Yan, C., Pommier, Y., Marchand, C. and Ivanov, I. (2016) 'Discovery of selective inhibitors of tyrosyl-DNA phosphodiesterase 2 by targeting the enzyme DNA-binding cleft', *Bioorg Med Chem Lett,* 26(14), pp. 3232-6.

Kouzminova, E. A. and Kuzminov, A. (2006) 'Fragmentation of replicating chromosomes triggered by uracil in DNA', *J Mol Biol,* 355(1), pp. 20-33.

Kuliopulos, A., Mohanlal, R. and Covic, L. (2004) 'Effect of selective inhibition of the p38 MAP kinase pathway on platelet aggregation', *Thromb Haemost,* 92(6), pp. 1387-93.

Kumar, R., Chaudhary, K., Gupta, S., Singh, H., Kumar, S., Gautam, A., Kapoor, P. and Raghava, G. P. (2013) 'CancerDR: cancer drug resistance database', *Sci Rep,* 3, pp. 1445.

Kuzminov, A. (2001) 'Single-strand interruptions in replicating chromosomes cause double-strand breaks', *Proc Natl Acad Sci U S A,* 98(15), pp. 8241-6.

Lage, H., Helmbach, H., Dietel, M. and Schadendorf, D. (2000) 'Modulation of DNA topoisomerase II activity and expression in melanoma cells with acquired drug resistance', *Br J Cancer,* 82(2), pp. 488-91.

Lasfargues, E. Y., Coutinho, W. G. and Redfield, E. S. (1978) 'Isolation of two human tumor epithelial cell lines from solid breast carcinomas', *J Natl Cancer Inst,* 61(4), pp. 967-78.

Launonen, V. (2005) 'Mutations in the human LKB1/STK11 gene', *Hum Mutat,* 26(4), pp. 291-7.

Le Page, F., Schreiber, V., Dherin, C., De Murcia, G. and Boiteux, S. (2003) 'Poly(ADP-ribose) polymerase-1 (PARP-1) is required in murine cell lines for base excision repair of oxidative DNA damage in the absence of DNA polymerase beta', *J Biol Chem,* 278(20), pp. 18471-7.

Leahy, J. J., Golding, B. T., Griffin, R. J., Hardcastle, I. R., Richardson, C., Rigoreau, L. and Smith, G. C. (2004) 'Identification of a highly potent and selective DNA-dependent protein kinase (DNA-PK) inhibitor (NU7441) by screening of chromenone libraries', *Bioorg Med Chem Lett,* 14(24), pp. 6083-7.

Lechner, J. F., McClendon, I. A., LaVeck, M. A., Shamsuddin, A. M. and Harris, C. C. (1983) 'Differential control by platelet factors of squamous differentiation in normal and malignant human bronchial epithelial cells', *Cancer Res,* 43(12 Pt 1), pp. 5915-21.

Lee, C. K., Lee, M. E., Lee, W. S., Kim, J. M., Park, K. H., Kim, T. S., Lee, K. Y., Ahn, J. B., Chung, H. C. and Rha, S. Y. (2015a) 'Dovitinib (TKI258), a multi-target angiokinase inhibitor, is effective regardless of KRAS or BRAF mutation status in colorectal cancer', *Am J Cancer Res,* 5(1), pp. 72-86.

Lee, H. J., Seo, A. N., Kim, E. J., Jang, M. H., Kim, Y. J., Kim, J. H., Kim, S. W., Ryu, H. S., Park, I. A., Im, S. A., Gong, G., Jung, K. H., Kim, H. J. and Park, S. Y. (2015b) 'Prognostic and predictive values of EGFR overexpression and EGFR copy number alteration in HER2-positive breast cancer', *Br J Cancer,* 112(1), pp. 103-11.

Lerner, E. C., Qian, Y., Blaskovich, M. A., Fossum, R. D., Vogt, A., Sun, J., Cox, A. D., Der, C. J., Hamilton, A. D. and Sebti, S. M. (1995) 'Ras CAAX peptidomimetic FTI-277 selectively blocks oncogenic Ras signaling by inducing cytoplasmic accumulation of inactive Ras-Raf complexes', *J Biol Chem,* 270(45), pp. 26802-6.

Li, C., Fan, S., Owonikoko, T. K., Khuri, F. R., Sun, S. Y. and Li, R. (2011a) 'Oncogenic role of EAPII in lung cancer development and its activation of the MAPK-ERK pathway', *Oncogene,* 30(35), pp. 3802-3812.

Li, C., Sun, S. Y., Khuri, F. R. and Li, R. (2011b) 'Pleiotropic functions of EAPII/TTRAP/TDP2: cancer development, chemoresistance and beyond', *Cell Cycle,* 10(19), pp. 3274-83.

Li, T. K. and Liu, L. F. (2001) 'Tumor cell death induced by topoisomerase-targeting drugs', *Annu Rev Pharmacol Toxicol,* 41, pp. 53-77.

Lieber, M. R. (2010) 'The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway', *Annu Rev Biochem,* 79, pp. 181-211.

Lieber, M. R. and Karanjawala, Z. E. (2004) 'Ageing, repetitive genomes and DNA damage', *Nat Rev Mol Cell Biol,* 5(1), pp. 69-75.

Lieber, M. R., Ma, Y., Pannicke, U. and Schwarz, K. (2003) 'Mechanism and regulation of human non-homologous DNA end-joining', *Nat Rev Mol Cell Biol,* 4(9), pp. 712-20.

Lindahl, T. (1993) 'Instability and decay of the primary structure of DNA', *Nature,* 362(6422), pp. 709-15.

Lindahl, T. and Nyberg, B. (1972) 'Rate of depurination of native deoxyribonucleic acid', *Biochemistry,* 11(19), pp. 3610-8.

Lindsley, C. W., Zhao, Z., Leister, W. H., Robinson, R. G., Barnett, S. F., Defeo-Jones, D., Jones, R. E., Hartman, G. D., Huff, J. R., Huber, H. E. and Duggan, M. E. (2005) 'Allosteric Akt (PKB) inhibitors: discovery and SAR of isozyme selective inhibitors', *Bioorg Med Chem Lett,* 15(3), pp. 761-4.

Liu, D. W., Chen, S. T. and Liu, H. P. (2005) 'Choice of endogenous control for gene expression in nonsmall cell lung cancer', *Eur Respir J,* 26(6), pp. 1002-8.

Liu, L. L., Zhao, H., Ma, T. F., Ge, F., Chen, C. S. and Zhang, Y. P. (2015) 'Identification of valid reference genes for the normalization of RT-qPCR expression studies in human breast cancer cell lines treated with and without transient transfection', *PLoS One,* 10(1), pp. e0117058.

Liu, Y., Prasad, R., Beard, W. A., Kedar, P. S., Hou, E. W., Shock, D. D. and Wilson, S. H. (2007) 'Coordination of steps in single-nucleotide base excision repair mediated by apurinic/apyrimidinic endonuclease 1 and DNA polymerase beta', *J Biol Chem,* 282(18), pp. 13532-41.

Lyng, M. B., Laenkholm, A. V., Pallisgaard, N. and Ditzel, H. J. (2008) 'Identification of genes for normalization of real-time RT-PCR data in breast carcinomas', *BMC Cancer,* 8, pp. 20.

Ma, Y., Pannicke, U., Schwarz, K. and Lieber, M. R. (2002) 'Hairpin opening and overhang processing by an Artemis/DNA-dependent protein kinase complex in nonhomologous end joining and V(D)J recombination', *Cell,* 108(6), pp. 781-94.

Macrae, C. J., McCulloch, R. D., Ylanko, J., Durocher, D. and Koch, C. A. (2008) 'APLF (C2orf13) facilitates nonhomologous end-joining and undergoes ATM-dependent hyperphosphorylation following ionizing radiation', *DNA Repair (Amst),* 7(2), pp. 292-302.

Mahaney, B. L., Meek, K. and Lees-Miller, S. P. (2009) 'Repair of ionizing radiation-induced DNA double-strand breaks by non-homologous end-joining', *Biochem J,* 417(3), pp. 639-50.

Majewski, I. J., Nuciforo, P., Mittempergher, L., Bosma, A. J., Eidtmann, H., Holmes, E., Sotiriou, C., Fumagalli, D., Jimenez, J., Aura, C., Prudkin, L., Díaz-Delgado, M. C., de la Peña, L., Loi, S., Ellis, C., Schultz, N., de Azambuja, E., Harbeck, N., Piccart-Gebhart, M., Bernards, R. and Baselga, J. (2015) 'PIK3CA mutations are associated with decreased benefit to neoadjuvant human epidermal growth factor receptor 2-targeted therapies in breast cancer', *J Clin Oncol,* 33(12), pp. 1334-9.

Malu, S., De Ioannes, P., Kozlov, M., Greene, M., Francis, D., Hanna, M., Pena, J., Escalante, C. R., Kurosawa, A., Erdjument-Bromage, H., Tempst, P., Adachi, N., Vezzoni, P., Villa, A., Aggarwal, A. K. and Cortes, P. (2012) 'Artemis C-terminal region facilitates V(D)J recombination through its interactions with DNA Ligase IV and DNA-PKcs', *J Exp Med,* 209(5), pp. 955-63.

Mandraju, R., Chekuri, A., Bhaskar, C., Duning, K., Kremerskothen, J. and Kondapi, A. K. (2011) 'Topoisomerase IIβ associates with Ku70 and PARP-1 during double strand break repair of DNA in neurons', *Arch Biochem Biophys,* 516(2), pp. 128-37.

Marchand, C., Abdelmalak, M., Kankanala, J., Huang, S. Y., Kiselev, E., Fesen, K., Kurahashi, K., Sasanuma, H., Takeda, S., Aihara, H., Wang, Z. and Pommier, Y. (2016) 'Deazaflavin Inhibitors of Tyrosyl-DNA Phosphodiesterase 2 (TDP2) Specific for the Human Enzyme and Active against Cellular TDP2', *ACS Chem Biol,* 11(7), pp. 1925-33.

Mari, P. O., Florea, B. I., Persengiev, S. P., Verkaik, N. S., Brüggenwirth, H. T., Modesti, M., Giglia-Mari, G., Bezstarosti, K., Demmers, J. A., Luider, T. M., Houtsmuller, A. B. and van Gent, D. C. (2006) 'Dynamic assembly of end-joining complexes requires interaction between Ku70/80 and XRCC4', *Proc Natl Acad Sci U S A,* 103(49), pp. 18597-602.

Marks, P. A. and Breslow, R. (2007) 'Dimethyl sulfoxide to vorinostat: development of this histone deacetylase inhibitor as an anticancer drug', *Nat Biotechnol,* 25(1), pp. 84-90.

Martin, G. M., Smith, A. C., Ketterer, D. J., Ogburn, C. E. and Disteche, C. M. (1985) 'Increased chromosomal aberrations in first metaphases of cells isolated from the kidneys of aged mice', *Isr J Med Sci,* 21(3), pp. 296-301.

Masson, M., Niedergang, C., Schreiber, V., Muller, S., Menissier-de Murcia, J. and de Murcia, G. (1998) 'XRCC1 is specifically associated with poly(ADP-ribose) polymerase and

negatively regulates its activity following DNA damage', *Mol Cell Biol,* 18(6), pp. 3563-71.

Maxwell, K. N., Wubbenhorst, B., D'Andrea, K., Garman, B., Long, J. M., Powers, J., Rathbun, K., Stopfer, J. E., Zhu, J., Bradbury, A. R., Simon, M. S., DeMichele, A., Domchek, S. M. and Nathanson, K. L. (2015) 'Prevalence of mutations in a panel of breast cancer susceptibility genes in BRCA1/2-negative patients with early-onset breast cancer', *Genet Med,* 17(8), pp. 630-8.

Maynard, S., Schurman, S. H., Harboe, C., de Souza-Pinto, N. C. and Bohr, V. A. (2009) 'Base excision repair of oxidative DNA damage and association with cancer and aging', *Carcinogenesis,* 30(1), pp. 2-10.

McClendon, A. K. and Osheroff, N. (2007) 'DNA topoisomerase II, genotoxicity, and cancer', *Mutat Res,* 623(1-2), pp. 83-97.

McConechy, M. K., Ding, J., Cheang, M. C., Wiegand, K. C., Senz, J., Tone, A. A., Yang, W., Prentice, L. M., Tse, K., Zeng, T., McDonald, H., Schmidt, A. P., Mutch, D. G., McAlpine, J. N., Hirst, M., Shah, S. P., Lee, C. H., Goodfellow, P. J., Gilks, C. B. and Huntsman, D. G. (2012) 'Use of mutation profiles to refine the classification of endometrial carcinomas', *J Pathol,* 228(1), pp. 20-30.

McIlwraith, M. J., McIlwraith, M. J., Vaisman, A., Liu, Y., Fanning, E., Woodgate, R. and West, S. C. (2005) 'Human DNA polymerase eta promotes DNA synthesis from strand invasion intermediates of homologous recombination', *Mol Cell,* 20(5), pp. 783-92.

McVeigh, T. P., Jung, S. Y., Kerin, M. J., Salzman, D. W., Nallur, S., Nemec, A. A., Dookwah, M., Sadofsky, J., Paranjape, T., Kelly, O., Chan, E., Miller, N., Sweeney, K. J., Zelterman, D., Sweasy, J., Pilarski, R., Telesca, D., Slack, F. J. and Weidhaas, J. B. (2015) 'Estrogen withdrawal, increased breast cancer risk and the KRAS-variant', *Cell Cycle,* 14(13), pp. 2091-9.

Medina, P. P., Romero, O. A., Kohno, T., Montuenga, L. M., Pio, R., Yokota, J. and Sanchez-Cespedes, M. (2008) 'Frequent BRG1/SMARCA4-inactivating mutations in human lung cancer cell lines', *Hum Mutat,* 29(5), pp. 617-22.

Mehrotra, P. V., Ahel, D., Ryan, D. P., Weston, R., Wiechens, N., Kraehenbuehl, R., Owen-Hughes, T. and Ahel, I. (2011) 'DNA repair factor APLF is a histone chaperone', *Mol Cell,* 41(1), pp. 46-55.

Mehta, A. and Haber, J. E. (2014) 'Sources of DNA double-strand breaks and models of recombinational DNA repair', *Cold Spring Harb Perspect Biol,* 6(9), pp. a016428.

Meisenberg, C., Gilbert, D. C., Chalmers, A., Haley, V., Gollins, S., Ward, S. E. and El-Khamisy, S. F. (2015) 'Clinical and cellular roles for TDP1 and TOP1 in modulating colorectal cancer response to irinotecan', *Mol Cancer Ther,* 14(2), pp. 575-85.

Meisenberg, C., Ward, S. E., Schmid, P. and El-Khamisy, S. F. (2014) 'TDP1/TOP1 Ratio as a Promising Indicator for the Response of Small Cell Lung Cancer to Topotecan', *J Cancer Sci Ther,* 6(7), pp. 258-267.

Menon, V. and Povirk, L. F. (2016) 'End-processing nucleases and phosphodiesterases: An elite supporting cast for the non-homologous end joining pathway of DNA double-strand break repair', *DNA Repair (Amst),* 43, pp. 57-68.

Miknyoczki, S. J., Dionne, C. A., Klein-Szanto, A. J. and Ruggeri, B. A. (1999) 'The novel Trk receptor tyrosine kinase inhibitor CEP-701 (KT-5555) exhibits antitumor efficacy against human pancreatic carcinoma (Panc1) xenograft growth and in vivo invasiveness', *Ann N Y Acad Sci,* 880, pp. 252-62.

Mimori, T. and Hardin, J. A. (1986) 'Mechanism of interaction between Ku protein and DNA', *J Biol Chem,* 261(22), pp. 10375-9.

Montecucco, A., Zanetta, F. and Biamonti, G. (2015) 'Molecular mechanisms of etoposide', *EXCLI J,* 14, pp. 95-108.

Moroni, F. (2008) 'Poly(ADP-ribose)polymerase 1 (PARP-1) and postischemic brain damage', *Curr Opin Pharmacol,* 8(1), pp. 96-103.

Morse, D. L., Carroll, D., Weberg, L., Borgstrom, M. C., Ranger-Moore, J. and Gillies, R. J. (2005) 'Determining suitable internal standards for mRNA quantification of increasing cancer progression in human breast cells by real-time reverse transcriptase polymerase chain reaction', *Anal Biochem,* 342(1), pp. 69-77.

Mortazavi, F., Lu, J., Phan, R., Lewis, M., Trinidad, K., Aljilani, A., Pezeshkpour, G. and Tamanoi, F. (2015) 'Significance of KRAS/PAK1/Crk pathway in non-small cell lung cancer oncogenesis', *BMC Cancer,* 15, pp. 381.

Murai, J., Huang, S. Y., Das, B. B., Renaud, A., Zhang, Y., Doroshow, J. H., Ji, J., Takeda, S. and Pommier, Y. (2012) 'Trapping of PARP1 and PARP2 by Clinical PARP Inhibitors', *Cancer Res,* 72(21), pp. 5588-99.

Murai, J., Zhang, Y., Morris, J., Ji, J., Takeda, S., Doroshow, J. H. and Pommier, Y. (2014) 'Rationale for poly(ADP-ribose) polymerase (PARP) inhibitors in combination therapy with camptothecins or temozolomide based on PARP trapping versus catalytic inhibition', *J Pharmacol Exp Ther,* 349(3), pp. 408-16.

Mäbert, K., Cojoc, M., Peitzsch, C., Kurth, I., Souchelnytskyi, S. and Dubrovska, A. (2014) 'Cancer biomarker discovery: current status and future perspectives', *Int J Radiat Biol,* 90(8), pp. 659-77.

Ngo, V. N., Davis, R. E., Lamy, L., Yu, X., Zhao, H., Lenz, G., Lam, L. T., Dave, S., Yang, L., Powell, J. and Staudt, L. M. (2006) 'A loss-of-function RNA interference screen for molecular targets in cancer', *Nature,* 441(7089), pp. 106-10.

Nguyen, M., Marcellus, R. C., Roulston, A., Watson, M., Serfass, L., Murthy Madiraju, S. R., Goulet, D., Viallet, J., Bélec, L., Billot, X., Acoca, S., Purisima, E., Wiegmans, A., Cluse, L., Johnstone, R. W., Beauparlant, P. and Shore, G. C. (2007) 'Small molecule obatoclax (GX15-070) antagonizes MCL-1 and overcomes MCL-1-mediated resistance to apoptosis', *Proc Natl Acad Sci U S A,* 104(49), pp. 19512-7.

Nitiss, J. L. (2009a) 'DNA topoisomerase II and its growing repertoire of biological functions', *Nat Rev Cancer,* 9(5), pp. 327-37.

Nitiss, J. L. (2009b) 'Targeting DNA topoisomerase II in cancer chemotherapy', *Nat Rev Cancer,* 9(5), pp. 338-50.

Nitiss, J. L. and Nitiss, K. C. (2013) 'Tdp2: a means to fixing the ends', *PLoS Genet,* 9(3), pp. e1003370.

Nitiss, K. C., Malik, M., He, X., White, S. W. and Nitiss, J. L. (2006) 'Tyrosyl-DNA phosphodiesterase (Tdp1) participates in the repair of Top2-mediated DNA damage', *Proc Natl Acad Sci U S A,* 103(24), pp. 8953-8.

Nylander, S., Kull, B., Bjorkman, J. A., Ulvinge, J. C., Oakes, N., Emanuelsson, B. M., Andersson, M., Skarby, T., Inghardt, T., Fjellstrom, O. and Gustafsson, D. (2012) 'Human target validation of phosphoinositide 3-kinase (PI3K)beta: effects on platelets and insulin sensitivity, using AZD6482 a novel PI3Kbeta inhibitor', *J Thromb Haemost,* 10(10), pp. 2127-36.

O'Hare, T., Shakespeare, W. C., Zhu, X., Eide, C. A., Rivera, V. M., Wang, F., Adrian, L. T., Zhou, T., Huang, W. S., Xu, Q., Metcalf, C. A., Tyner, J. W., Loriaux, M. M., Corbin, A. S., Wardwell, S., Ning, Y., Keats, J. A., Wang, Y., Sundaramoorthi, R., Thomas, M., Zhou, D., Snodgrass, J., Commodore, L., Sawyer, T. K., Dalgarno, D. C., Deininger, M. W., Druker, B. J. and Clackson, T. (2009) 'AP24534, a pan-BCR-ABL inhibitor for chronic myeloid leukemia, potently inhibits the T315I mutant and overcomes mutation-based resistance', *Cancer Cell,* 16(5), pp. 401-12.

O'Shaughnessy, J. (2005) 'Extending survival with chemotherapy in metastatic breast cancer', *Oncologist,* 10 Suppl 3, pp. 20-9.

Ochi, T., Blackford, A. N., Coates, J., Jhujh, S., Mehmood, S., Tamura, N., Travers, J., Wu, Q., Draviam, V. M., Robinson, C. V., Blundell, T. L. and Jackson, S. P. (2015) 'DNA repair. PAXX, a paralog of XRCC4 and XLF, interacts with Ku to promote DNA double-strand break repair', *Science,* 347(6218), pp. 185-8.

Ostendorf, B. N., le Coutre, P., Kim, T. D. and Quintás-Cardama, A. (2014) 'Nilotinib', *Recent Results Cancer Res,* 201, pp. 67-80.

Pan, C., Kumar, C., Bohl, S., Klingmueller, U. and Mann, M. (2009) 'Comparative proteomic phenotyping of cell lines and primary cells to assess preservation of cell type-specific functions', *Mol Cell Proteomics,* 8(3), pp. 443-50.

Pavlopoulou, A., Spandidos, D. A. and Michalopoulos, I. (2015) 'Human cancer databases (review)', *Oncol Rep,* 33(1), pp. 3-18.

Pei, H., Li, C., Adereth, Y., Hsu, T., Watson, D. K. and Li, R. (2005) 'Caspase-1 is a direct target gene of ETS1 and plays a role in ETS1-induced apoptosis', *Cancer Res,* 65(16), pp. 7205-13.

Pei, H. P., Yordy, J. S., Leng, Q. X., Zhao, Q. H., Watson, D. K. and Li, R. Z. (2003) 'EAPII interacts with ETS1 and modulates its transcriptional function', *Oncogene,* 22(18), pp. 2699-2709.

Pelden, S., Insawang, T., Thuwajit, C. and Thuwajit, P. (2013) 'The trefoil factor 1 (TFF1) protein involved in doxorubicininduced apoptosis resistance is upregulated by estrogen in breast cancer cells', *Oncol Rep,* 30(3), pp. 1518-26.

Pellegrini, L., Yu, D. S., Lo, T., Anand, S., Lee, M., Blundell, T. L. and Venkitaraman, A. R. (2002) 'Insights into DNA recombination from the structure of a RAD51-BRCA2 complex', *Nature,* 420(6913), pp. 287-93.

Pendleton, M., Lindsey, R. H., Jr., Felix, C. A., Grimwade, D. and Osheroff, N. (2014) 'Topoisomerase II and leukemia', *Ann N Y Acad Sci,* 1310, pp. 98-110.

Pereira, B., Chin, S. F., Rueda, O. M., Vollan, H. K., Provenzano, E., Bardwell, H. A., Pugh, M., Jones, L., Russell, R., Sammut, S. J., Tsui, D. W., Liu, B., Dawson, S. J., Abraham, J., Northen, H., Peden, J. F., Mukherjee, A., Turashvili, G., Green, A. R., McKinney, S., Oloumi, A., Shah, S., Rosenfeld, N., Murphy, L., Bentley, D. R., Ellis, I. O., Purushotham, A., Pinder, S. E., Børresen-Dale, A. L., Earl, H. M., Pharoah, P. D., Ross, M. T., Aparicio, S. and Caldas, C. (2016) 'The somatic mutation profiles of 2,433 breast cancers refines their genomic and transcriptomic landscapes', *Nat Commun,* 7, pp. 11479.

Plummer, R., Jones, C., Middleton, M., Wilson, R., Evans, J., Olsen, A., Curtin, N., Boddy, A., McHugh, P., Newell, D., Harris, A., Johnson, P., Steinfeldt, H., Dewji, R., Wang, D., Robson, L. and Calvert, H. (2008) 'Phase I study of the poly(ADP-ribose) polymerase inhibitor, AG014699, in combination with temozolomide in patients with advanced solid tumors', *Clin Cancer Res,* 14(23), pp. 7917-23.

Pogozelski, W. K. and Tullius, T. D. (1998) 'Oxidative Strand Scission of Nucleic Acids: Routes Initiated by Hydrogen Abstraction from the Sugar Moiety', *Chem Rev,* 98(3), pp. 1089-1108.

Pommier, Y. (2006) 'Topoisomerase I inhibitors: camptothecins and beyond', *Nat Rev Cancer,* 6(10), pp. 789-802.

Pommier, Y. (2013) 'Drugging topoisomerases: lessons and challenges', *ACS Chem Biol,* 8(1), pp. 82-95.

Pommier, Y., Huang, S. Y., Gao, R., Das, B. B., Murai, J. and Marchand, C. (2014) 'Tyrosyl-DNA-phosphodiesterases (TDP1 and TDP2)', *DNA Repair (Amst),* 19, pp. 114-29.

Pommier, Y., Leo, E., Zhang, H. and Marchand, C. (2010) 'DNA topoisomerases and their poisoning by anticancer and antibacterial drugs', *Chem Biol,* 17(5), pp. 421-33.

Pommier, Y., Sun, Y., Huang, S. N. and Nitiss, J. L. (2016) 'Roles of eukaryotic topoisomerases in transcription, replication and genomic stability', *Nat Rev Mol Cell Biol,* 17(11), pp. 703-721.

Pourquier, P. and Pommier, Y. (2001) 'Topoisomerase I-mediated DNA damage', *Adv Cancer Res,* 80, pp. 189-216.

Pray, L. (2008) 'DNA Replication and Causes of Mutation.', *Nature Education,* 1(1).

Program, S. E. a. E. R. S. (2015) *SEER Cancer Statistics Factsheets: Breast Cancer*. National Cancer Institute, Bethesda, MD. Available at: http://seer.cancer.gov/statfacts/html/breast.html (Accessed: 11 September 2015.

Purnell, M. R. and Whish, W. J. (1980) 'Novel inhibitors of poly(ADP-ribose) synthetase', *Biochem J,* 185(3), pp. 775-7.

Pype, S., Declercq, W., Ibrahimi, A., Michiels, C., Van Rietschoten, J. G., Dewulf, N., de Boer, M., Vandenabeele, P., Huylebroeck, D. and Remacle, J. E. (2000) 'TTRAP, a novel protein that associates with CD40, tumor necrosis factor (TNF) receptor-75 and TNF receptor-associated factors (TRAFs), and that inhibits nuclear factor-kappa B activation', *J Biol Chem,* 275(24), pp. 18586-93.

Pâques, F. and Haber, J. E. (1999) 'Multiple pathways of recombination induced by double-strand breaks in Saccharomyces cerevisiae', *Microbiol Mol Biol Rev,* 63(2), pp. 349-404.

Pérez-Ramírez, C., Cañadas-Garre, M., Molina, M., Faus-Dáder, M. J. and Calleja-Hernández, M. (2015) 'PTEN and PI3K/AKT in non-small-cell lung cancer', *Pharmacogenomics,* 16(16), pp. 1843-62.

Qi, X. C., Xie, D. J., Yan, Q. F., Wang, Y. R., Zhu, Y. X., Qian, C. and Yang, S. X. (2013) 'LRIG1 dictates the chemo-sensitivity of temozolomide (TMZ) in U251 glioblastoma cells via down-regulation of EGFR/topoisomerase-2/Bcl-2', *Biochem Biophys Res Commun,* 437(4), pp. 565-72.

Raoof, A., Depledge, P., Hamilton, N. M., Hamilton, N. S., Hitchin, J. R., Hopkins, G. V., Jordan, A. M., Maguire, L. A., McGonagle, A. E., Mould, D. P., Rushbrooke, M., Small, H. F., Smith, K. M., Thomson, G. J., Turlais, F., Waddell, I. D., Waszkowycz, B., Watson, A. J. and Ogilvie, D. J. (2013) 'Toxoflavins and deazaflavins as the first reported selective small molecule inhibitors of tyrosyl-DNA phosphodiesterase II', *J Med Chem,* 56(16), pp. 6352-70.

Reardon, J. T. and Sancar, A. (2006) 'Purification and characterization of Escherichia coli and human nucleotide excision repair enzyme systems', *Methods Enzymol,* 408, pp. 189-213.

Reid, A., Vidal, L., Shaw, H. and de Bono, J. (2007) 'Dual inhibition of ErbB1 (EGFR/HER1) and ErbB2 (HER2/neu)', *Eur J Cancer,* 43(3), pp. 481-9.

Renodon-Cornière Axelle, W. P., Le Breton Magali, Fleury Fabrice (2013) 'New Potential Therapeutic Approaches by Targeting Rad51- Dependent Homologous Recombination', *New Research Directions in DNA Repair*: InTech.

Reynolds, J. J., Walker, A. K., Gilmore, E. C., Walsh, C. A. and Caldecott, K. W. (2012) 'Impact of PNKP mutations associated with microcephaly, seizures and developmental delay on enzyme activity and DNA strand break repair', *Nucleic Acids Res,* 40(14), pp. 6608-19.

Riballo, E., Kühne, M., Rief, N., Doherty, A., Smith, G. C., Recio, M. J., Reis, C., Dahm, K., Fricke, A., Krempler, A., Parker, A. R., Jackson, S. P., Gennery, A., Jeggo, P. A. and Löbrich, M. (2004) 'A pathway of double-strand break rejoining dependent upon ATM, Artemis, and proteins locating to gamma-H2AX foci', *Mol Cell,* 16(5), pp. 715-24.

Rich, T., Allen, R. L. and Wyllie, A. H. (2000) 'Defying death after DNA damage', *Nature,* 407(6805), pp. 777-83.

Riedl, S. J. and Shi, Y. (2004) 'Molecular mechanisms of caspase regulation during apoptosis', *Nat Rev Mol Cell Biol,* 5(11), pp. 897-907.

Risbridger, G. P., Davis, I. D., Birrell, S. N. and Tilley, W. D. (2010) 'Breast and prostate cancer: more similar than different', *Nat Rev Cancer,* 10(3), pp. 205-12.

Roberts, W. G., Ung, E., Whalen, P., Cooper, B., Hulford, C., Autry, C., Richter, D., Emerson, E., Lin, J., Kath, J., Coleman, K., Yao, L., Martinez-Alsina, L., Lorenzen, M., Berliner, M., Luzzio, M., Patel, N., Schmitt, E., LaGreca, S., Jani, J., Wessel, M., Marr, E., Griffor, M. and Vajdos, F. (2008) 'Antitumor activity and pharmacology of a selective focal adhesion kinase inhibitor, PF-562,271', *Cancer Res,* 68(6), pp. 1935-44.

Rocha, J. C., Busatto, F. F., Guecheva, T. N. and Saffi, J. (2016) 'Role of nucleotide excision repair proteins in response to DNA damage induced by topoisomerase II inhibitors', *Mutat Res Rev Mutat Res,* 768, pp. 68-77.

Rodriguez-Nieto, S. and Sanchez-Cespedes, M. (2009) 'BRG1 and LKB1: tales of two tumor suppressor genes on chromosome 19p and lung cancer', *Carcinogenesis,* 30(4), pp. 547-54.

Rouleau, M., Patel, A., Hendzel, M. J., Kaufmann, S. H. and Poirier, G. G. (2010) 'PARP inhibition: PARP1 and beyond', *Nat Rev Cancer,* 10(4), pp. 293-301.

Rugo, H. S., Herbst, R. S., Liu, G., Park, J. W., Kies, M. S., Steinfeldt, H. M., Pithavala, Y. K., Reich, S. D., Freddo, J. L. and Wilding, G. (2005) 'Phase I trial of the oral antiangiogenesis agent AG-013736 in patients with advanced solid tumors: pharmacokinetic and clinical results', *J Clin Oncol,* 23(24), pp. 5474-83.

Rulten, S. L., Cortes-Ledesma, F., Guo, L., Iles, N. J. and Caldecott, K. W. (2008) 'APLF (C2orf13) is a novel component of poly(ADP-ribose) signaling in mammalian cells', *Mol Cell Biol,* 28(14), pp. 4620-8.

Rulten, S. L., Fisher, A. E., Robert, I., Zuma, M. C., Rouleau, M., Ju, L., Poirier, G., Reina-San-Martin, B. and Caldecott, K. W. (2011) 'PARP-3 and APLF function together to accelerate nonhomologous end-joining', *Mol Cell,* 41(1), pp. 33-45.

Saberi, A., Hochegger, H., Szuts, D., Lan, L., Yasui, A., Sale, J. E., Taniguchi, Y., Murakawa, Y., Zeng, W., Yokomori, K., Helleday, T., Teraoka, H., Arakawa, H., Buerstedde, J. M. and Takeda, S. (2007) 'RAD18 and poly(ADP-ribose) polymerase independently suppress the access of nonhomologous end joining to double-strand breaks and facilitate homologous recombination-mediated repair', *Mol Cell Biol,* 27(7), pp. 2562-71.

Sabisz, M., Wesierska-Gadek, J. and Skladanowski, A. (2010) 'Increased cytotoxicity of an unusual DNA topoisomerase II inhibitor compound C-1305 toward HeLa cells with downregulated PARP-1 activity results from re-activation of the p53 pathway and modulation of mitotic checkpoints', *Biochem Pharmacol,* 79(10), pp. 1387-97.

Salerno, S., Da Settimo, F., Taliani, S., Simorini, F., La Motta, C., Fornaciari, G. and Marini, A. M. (2010) 'Recent advances in the development of dual topoisomerase I and II inhibitors as anticancer drugs', *Curr Med Chem,* 17(35), pp. 4270-90.

Sambrook, J. and Russell, D. W. (2006) 'SDS-Polyacrylamide Gel Electrophoresis of Proteins', *CSH Protoc,* 2006(4).

San Filippo, J., Chi, P., Sehorn, M. G., Etchin, J., Krejci, L. and Sung, P. (2006) 'Recombination mediator and Rad51 targeting activities of a human BRCA2 polypeptide', *J Biol Chem,* 281(17), pp. 11649-57.

Sander, M. and Hsieh, T. (1983) 'Double strand DNA cleavage by type II DNA topoisomerase from Drosophila melanogaster', *J Biol Chem,* 258(13), pp. 8421-8.

Sathya, S., Sudhagar, S., Sarathkumar, B. and Lakshmi, B. S. (2014) 'EGFR inhibition by pentacyclic triterpenes exhibit cell cycle and growth arrest in breast cancer cells', *Life Sci,* 95(1), pp. 53-62.

Schellenberg, M. J., Appel, C. D., Adhikari, S., Robertson, P. D., Ramsden, D. A. and Williams, R. S. (2012) 'Mechanism of repair of 5'-topoisomerase II-DNA adducts by mammalian tyrosyl-DNA phosphodiesterase 2', *Nat Struct Mol Biol,* 19(12), pp. 1363-71.

Schlabach, M. R., Luo, J., Solimini, N. L., Hu, G., Xu, Q., Li, M. Z., Zhao, Z., Smogorzewska, A., Sowa, M. E., Ang, X. L., Westbrook, T. F., Liang, A. C., Chang, K., Hackett, J. A., Harper, J. W., Hannon, G. J. and Elledge, S. J. (2008) 'Cancer proliferation gene discovery through functional genomics', *Science,* 319(5863), pp. 620-4.

Schvartzman, J. B. and Stasiak, A. (2004) 'A topological view of the replicon', *EMBO Rep,* 5(3), pp. 256-61.

Scott, A. M., Wolchok, J. D. and Old, L. J. (2012) 'Antibody therapy of cancer', *Nat Rev Cancer,* 12(4), pp. 278-87.

Sebestyén, A., Mihalik, R., Peták, I. and Kopper, L. (1997) 'Modulation of apoptosis signaling in etoposide-treated lymphoma cells', *Anticancer Res,* 17(4A), pp. 2609-14.

Sebolt-Leopold, J. S., Dudley, D. T., Herrera, R., Van Becelaere, K., Wiland, A., Gowan, R. C., Tecle, H., Barrett, S. D., Bridges, A., Przybranowski, S., Leopold, W. R. and Saltiel, A. R. (1999) 'Blockade of the MAP kinase pathway suppresses growth of colon tumors in vivo', *Nat Med,* 5(7), pp. 810-6.

Seth, A. and Watson, D. K. (2005) 'ETS transcription factors and their emerging roles in human cancer', *Eur J Cancer,* 41(16), pp. 2462-78.

Shi, K., Kurahashi, K., Gao, R., Tsutakawa, S. E., Tainer, J. A., Pommier, Y. and Aihara, H. (2012) 'Structural basis for recognition of 5'-phosphotyrosine adducts by Tdp2', *Nat Struct Mol Biol,* 19(12), pp. 1372-7.

Shiovitz, S. and Korde, L. A. (2015) 'Genetics of breast cancer: a topic in evolution', *Ann Oncol,* 26(7), pp. 1291-9.

Shoemaker, R. H. (2006) 'The NCI60 human tumour cell line anticancer drug screen', *Nat Rev Cancer,* 6(10), pp. 813-23.

Shrivastav, M., De Haro, L. P. and Nickoloff, J. A. (2008) 'Regulation of DNA double-strand break repair pathway choice', *Cell Res,* 18(1), pp. 134-47.

Siegel, R. L., Miller, K. D. and Jemal, A. (2015) 'Cancer statistics, 2015', *CA Cancer J Clin,* 65(1), pp. 5-29.

Silber, R., Potmesil, M. and Bank, B. B. (1989) 'Studies on drug resistance in chronic lymphocytic leukemia', *Adv Enzyme Regul,* 29, pp. 267-76.

Silva, J. M., Marran, K., Parker, J. S., Silva, J., Golding, M., Schlabach, M. R., Elledge, S. J., Hannon, G. J. and Chang, K. (2008) 'Profiling essential genes in human mammary cells by multiplex RNAi screening', *Science,* 319(5863), pp. 617-20.

Silwal-Pandit, L., Vollan, H. K., Chin, S. F., Rueda, O. M., McKinney, S., Osako, T., Quigley, D. A., Kristensen, V. N., Aparicio, S., Børresen-Dale, A. L., Caldas, C. and Langerød, A. (2014) 'TP53 mutation spectrum in breast cancer is subtype specific and has distinct prognostic relevance', *Clin Cancer Res,* 20(13), pp. 3569-80.

Singh, S. M., Gauthier, S. and Labrie, F. (2000) 'Androgen receptor antagonists (antiandrogens): structure-activity relationships', *Curr Med Chem,* 7(2), pp. 211-47.

Smith, J. A., Poteet-Smith, C. E., Xu, Y., Errington, T. M., Hecht, S. M. and Lannigan, D. A. (2005) 'Identification of the first specific inhibitor of p90 ribosomal S6 kinase (RSK) reveals an unexpected role for RSK in cancer cell proliferation', *Cancer Res,* 65(3), pp. 1027-34.

Sonoda, E., Hochegger, H., Saberi, A., Taniguchi, Y. and Takeda, S. (2006) 'Differential usage of non-homologous end-joining and homologous recombination in double strand break repair', *DNA Repair (Amst),* 5(9-10), pp. 1021-9.

Sordet, O., Goldman, A. and Pommier, Y. (2006) 'Topoisomerase II and tubulin inhibitors both induce the formation of apoptotic topoisomerase I cleavage complexes', *Mol Cancer Ther,* 5(12), pp. 3139-44.

Soubeyrand, S., Pope, L., De Chasseval, R., Gosselin, D., Dong, F., de Villartay, J. P. and Haché, R. J. (2006) 'Artemis phosphorylated by DNA-dependent protein kinase associates preferentially with discrete regions of chromatin', *J Mol Biol,* 358(5), pp. 1200-11.

Soule, H. D., Vazquez, J., Long, A., Albert, S. and Brennan, M. (1973) 'A human cell line from a pleural effusion derived from a breast carcinoma', *J Natl Cancer Inst,* 51(5), pp. 1409-16.

Stack, G., Kumar, V., Green, S., Ponglikitmongkol, M., Berry, M., Rio, M. C., Nunez, A. M., Roberts, M., Koehl, C., Bellocq, P. and et al. (1988) 'Structure and function of the pS2 gene and estrogen receptor in human breast cancer cells', *Cancer Treat Res,* 40, pp. 185-206.

Stark, A. M., Pfannenschmidt, S., Tscheslog, H., Maass, N., Rosel, F., Mehdorn, H. M. and Held-Feindt, J. (2006) 'Reduced mRNA and protein expression of BCL-2 versus decreased

mRNA and increased protein expression of BAX in breast cancer brain metastases: a real-time PCR and immunohistochemical evaluation', *Neurol Res,* 28(8), pp. 787-93.

Stratton, M. R., Campbell, P. J. and Futreal, P. A. (2009) 'The cancer genome', *Nature,* 458(7239), pp. 719-24.

Sugimura, K., Takebayashi, S., Taguchi, H., Takeda, S. and Okumura, K. (2008) 'PARP-1 ensures regulation of replication fork progression by homologous recombination on damaged DNA', *J Cell Biol,* 183(7), pp. 1203-12.

Sy, S. M., Huen, M. S. and Chen, J. (2009) 'PALB2 is an integral component of the BRCA complex required for homologous recombination repair', *Proc Natl Acad Sci U S A,* 106(17), pp. 7155-60.

Takashima, H., Boerkoel, C. F., John, J., Saifi, G. M., Salih, M. A., Armstrong, D., Mao, Y., Quiocho, F. A., Roa, B. B., Nakagawa, M., Stockton, D. W. and Lupski, J. R. (2002) 'Mutation of TDP1, encoding a topoisomerase I-dependent DNA damage repair enzyme, in spinocerebellar ataxia with axonal neuropathy', *Nat Genet,* 32(2), pp. 267-72.

Taquet, N., Dumont, S., Vonesch, J. L., Hentsch, D., Reimund, J. M. and Muller, C. D. (2009) 'Differential between protein and mRNA expression of CCR7 and SSTR5 receptors in Crohn's disease patients', *Mediators Inflamm,* 2009, pp. 285812.

Team, R. C. 2015. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.

Thomas, H. D., Calabrese, C. R., Batey, M. A., Canan, S., Hostomsky, Z., Kyle, S., Maegley, K. A., Newell, D. R., Skalitzky, D., Wang, L. Z., Webber, S. E. and Curtin, N. J. (2007) 'Preclinical selection of a novel poly(ADP-ribose) polymerase inhibitor for clinical trial', *Mol Cancer Ther,* 6(3), pp. 945-56.

Thompson, E. W., Katz, D., Shima, T. B., Wakeling, A. E., Lippman, M. E. and Dickson, R. B. (1989) 'ICI 164,384, a pure antagonist of estrogen-stimulated MCF-7 cell proliferation and invasiveness', *Cancer Res,* 49(24 Pt 1), pp. 6929-34.

Thomson, G., Watson, A., Caldecott, K., Denneny, O., Depledge, P., Hamilton, N., Hopkins, G., Jordan, A., Morrow, C., Raoof, A., Waddell, I. and Ogilvie, D. (2013) 'Generation of assays and antibodies to facilitate the study of human 5'-tyrosyl DNA phosphodiesterase', *Anal Biochem,* 436(2), pp. 145-50.

Tighe, A., Ray-Sinha, A., Staples, O. D. and Taylor, S. S. (2007) 'GSK-3 inhibitors induce chromosome instability', *BMC Cell Biol,* 8, pp. 34.

Tolhurst, R. S., Thomas, R. S., Kyle, F. J., Patel, H., Periyasamy, M., Photiou, A., Thiruchelvam, P. T., Lai, C. F., Al-Sabbagh, M., Fisher, R. A., Barry, S., Crnogorac-Jurcevic, T., Martin, L. A., Dowsett, M., Charles Coombes, R., Kamalati, T., Ali, S. and Buluwela, L. (2011) 'Transient over-expression of estrogen receptor-alpha in breast cancer cells promotes cell survival and estrogen-independent growth', *Breast Cancer Res Treat,* 128(2), pp. 357-68.

Tounekti, O., Kenani, A., Foray, N., Orlowski, S. and Mir, L. M. (2001) 'The ratio of single- to double-strand DNA breaks and their absolute values determine cell death pathway', *Br J Cancer,* 84(9), pp. 1272-9.

Trudel, S., Li, Z. H., Wei, E., Wiesmann, M., Chang, H., Chen, C., Reece, D., Heise, C. and Stewart, A. K. (2005) 'CHIR-258, a novel, multitargeted tyrosine kinase inhibitor for the potential treatment of t(4;14) multiple myeloma', *Blood,* 105(7), pp. 2941-8.

Twentyman, P. R., Fox, N. E., Wright, K. A. and Bleehen, N. M. (1986) 'Derivation and preliminary characterisation of adriamycin resistant lines of human lung cancer cells', *Br J Cancer,* 53(4), pp. 529-37.

Umezawa, H. (1965) 'Bleomycin and other antitumor antibiotics of high molecular weight', *Antimicrob Agents Chemother (Bethesda),* 5, pp. 1079-85.

Uziel, T., Lerenthal, Y., Moyal, L., Andegeko, Y., Mittelman, L. and Shiloh, Y. (2003) 'Requirement of the MRN complex for ATM activation by DNA damage', *EMBO J,* 22(20), pp. 5612-21.

Valabrega, G., Berrino, G., Milani, A., Aglietta, M. and Montemurro, F. (2015) 'A retrospective analysis of the activity and safety of oral Etoposide in heavily pretreated metastatic breast cancer patients', *Breast J,* 21(3), pp. 241-5.

Valkov, N. I. and Sullivan, D. M. (1997) 'Drug resistance to DNA topoisomerase I and II inhibitors in human leukemia, lymphoma, and multiple myeloma', *Semin Hematol,* 34(4 Suppl 5), pp. 48-62.

van Attikum, H. and Gasser, S. M. (2005) 'The histone code at DNA breaks: a guide to repair?', *Nat Rev Mol Cell Biol,* 6(10), pp. 757-65.

van Gent, D. C., Hoeijmakers, J. H. and Kanaar, R. (2001) 'Chromosomal stability and the DNA double-stranded break connection', *Nat Rev Genet,* 2(3), pp. 196-206.

Vassilev, L. T., Vu, B. T., Graves, B., Carvajal, D., Podlaski, F., Filipovic, Z., Kong, N., Kammlott, U., Lukacs, C., Klein, C., Fotouhi, N. and Liu, E. A. (2004) 'In vivo activation of the p53 pathway by small-molecule antagonists of MDM2', *Science,* 303(5659), pp. 844-8.

Veglia, F., Matullo, G. and Vineis, P. (2003) 'Bulky DNA adducts and risk of cancer: a meta-analysis', *Cancer Epidemiol Biomarkers Prev,* 12(2), pp. 157-60.

Vogel, C. and Marcotte, E. M. (2012) 'Insights into the regulation of protein abundance from proteomic and transcriptomic analyses', *Nat Rev Genet,* 13(4), pp. 227-32.

Wakeling, A. E., Dukes, M. and Bowler, J. (1991) 'A potent specific pure antiestrogen with clinical potential', *Cancer Res,* 51(15), pp. 3867-73.

Walker, J. R., Corpina, R. A. and Goldberg, J. (2001) 'Structure of the Ku heterodimer bound to DNA and its implications for double-strand break repair', *Nature,* 412(6847), pp. 607-14.

Wall, M. E., Wani, M. C., Cook, C. E., Palmer, K. H., McPhail, A. T. and Sim, G. A. (1966) 'Plant Antitumor Agents. I. The Isolation and Structure of Camptothecin, a Novel Alkaloidal Leukemia and Tumor Inhibitor from Camptotheca acuminata', *J. Am. Chem. Soc.,* 88(16), pp. 3888-3890.

Wang, J. C. (1996) 'DNA topoisomerases', *Annu Rev Biochem,* 65, pp. 635-92.

Wang, J. C. (2002) 'Cellular roles of DNA topoisomerases: a molecular perspective', *Nat Rev Mol Cell Biol,* 3(6), pp. 430-40.

Wang, L., Hu, H., Pan, Y., Wang, R., Li, Y., Shen, L., Yu, Y., Li, H., Cai, D., Sun, Y. and Chen, H. (2014) 'PIK3CA mutations frequently coexist with EGFR/KRAS mutations in non-small cell lung cancer and suggest poor prognosis in EGFR/KRAS wildtype subgroup', *PLoS One,* 9(2), pp. e88291.

Wang, Q., Cui, K., Espin-Garcia, O., Cheng, D., Qiu, X., Chen, Z., Moore, M., Bristow, R. G., Xu, W., Der, S. and Liu, G. (2013) 'Resistance to bleomycin in cancer cell lines is characterized by prolonged doubling time, reduced DNA damage and evasion of G2/M arrest and apoptosis', *PLoS One,* 8(12), pp. e82363.

Ward, J. F. (1975) 'Radiation-induced strand breakage in DNA', *Basic Life Sci,* 5B, pp. 471-2.

Ward, J. F., Evans, J. W., Limoli, C. L. and Calabro-Jones, P. M. (1987) 'Radiation and hydrogen peroxide induced free radical damage to DNA', *Br J Cancer Suppl,* 8, pp. 105-12.

Weil, M. K. and Chen, A. P. (2011) 'PARP inhibitor treatment in ovarian and breast cancer', *Curr Probl Cancer,* 35(1), pp. 7-50.

Weinfeld, M., Chaudhry, M. A., D'Amours, D., Pelletier, J. D., Poirier, G. G., Povirk, L. F. and Lees-Miller, S. P. (1997) 'Interaction of DNA-dependent protein kinase and poly(ADP-ribose) polymerase with radiation-induced DNA strand breaks', *Radiat Res,* 148(1), pp. 22-8.

Weisberg, E., Manley, P. W., Breitenstein, W., Brüggen, J., Cowan-Jacob, S. W., Ray, A., Huntly, B., Fabbro, D., Fendrich, G., Hall-Meyers, E., Kung, A. L., Mestan, J., Daley, G. Q., Callahan, L., Catley, L., Cavazza, C., Azam, M., Mohammed, A., Neuberg, D., Wright, R.

D., Gilliland, D. G. and Griffin, J. D. (2005) 'Characterization of AMN107, a selective inhibitor of native and mutant Bcr-Abl', *Cancer Cell,* 7(2), pp. 129-41.

White, A. W., Almassy, R., Calvert, A. H., Curtin, N. J., Griffin, R. J., Hostomsky, Z., Maegley, K., Newell, D. R., Srinivasan, S. and Golding, B. T. (2000) 'Resistance-modifying agents. 9. Synthesis and biological properties of benzimidazole inhibitors of the DNA repair enzyme poly(ADP-ribose) polymerase', *J Med Chem,* 43(22), pp. 4084-97.

Williamson, L. M. and Lees-Miller, S. P. (2011) 'Estrogen receptor alpha-mediated transcription induces cell cycle-dependent DNA double-strand breaks', *Carcinogenesis,* 32(3), pp. 279-85.

Wishart, D. S., Knox, C., Guo, A. C., Shrivastava, S., Hassanali, M., Stothard, P., Chang, Z. and Woolsey, J. (2006) 'DrugBank: a comprehensive resource for in silico drug discovery and exploration', *Nucleic Acids Res,* 34(Database issue), pp. D668-72.

Wistuba, II, Bryant, D., Behrens, C., Milchgrub, S., Virmani, A. K., Ashfaq, R., Minna, J. D. and Gazdar, A. F. (1999) 'Comparison of features of human lung cancer cell lines and their corresponding tumors', *Clin Cancer Res,* 5(5), pp. 991-1000.

Wittman, M. D., Carboni, J. M., Yang, Z., Lee, F. Y., Antman, M., Attar, R., Balimane, P., Chang, C., Chen, C., Discenza, L., Frennesson, D., Gottardis, M. M., Greer, A., Hurlburt, W., Johnson, W., Langley, D. R., Li, A., Li, J., Liu, P., Mastalerz, H., Mathur, A., Menard, K., Patel, K., Sack, J., Sang, X., Saulnier, M., Smith, D., Stefanski, K., Trainor, G., Velaparthi, U., Zhang, G., Zimmermann, K. and Vyas, D. M. (2009) 'Discovery of a 2,4-disubstituted pyrrolo[1,2-f][1,2,4]triazine inhibitor (BMS-754807) of insulin-like growth factor receptor (IGF-1R) kinase in clinical development', *J Med Chem,* 52(23), pp. 7360-3.

Wong, A. K., Pero, R., Ormonde, P. A., Tavtigian, S. V. and Bartel, P. L. (1997) 'RAD51 interacts with the evolutionarily conserved BRC motifs in the human breast cancer susceptibility gene brca2', *J Biol Chem,* 272(51), pp. 31941-4.

Xu, G. L., Pan, Y. K., Wang, B. Y., Huang, L., Tian, L., Xue, J. L., Chen, J. Z. and Jia, W. (2008) 'TTRAP is a novel PML nuclear bodies-associated protein', *Biochemical and Biophysical Research Communications,* 375(3), pp. 395-398.

Yang, J. C., Sequist, L. V., Geater, S. L., Tsai, C. M., Mok, T. S., Schuler, M., Yamamoto, N., Yu, C. J., Ou, S. H., Zhou, C., Massey, D., Zazulina, V. and Wu, Y. L. (2015a) 'Clinical activity of afatinib in patients with advanced non-small-cell lung cancer harbouring uncommon EGFR mutations: a combined post-hoc analysis of LUX-Lung 2, LUX-Lung 3, and LUX-Lung 6', *Lancet Oncol,* 16(7), pp. 830-8.

Yang, K., Guo, R. and Xu, D. (2016) 'Non-homologous end joining: advances and frontiers', *Acta Biochim Biophys Sin (Shanghai),* 48(7), pp. 632-40.

Yang, S. W., Burgin, A. B., Jr., Huizenga, B. N., Robertson, C. A., Yao, K. C. and Nash, H. A. (1996) 'A eukaryotic enzyme that can disjoin dead-end covalent complexes between DNA and type I topoisomerases', *Proc Natl Acad Sci U S A,* 93(21), pp. 11534-9.

Yang, Y., Dong, X., Xie, B., Ding, N., Chen, J., Li, Y., Zhang, Q., Qu, H. and Fang, X. (2015b) 'Databases and web tools for cancer genomics study', *Genomics Proteomics Bioinformatics,* 13(1), pp. 46-50.

Yoo, S. and Dynan, W. S. (1999) 'Geometry of a complex formed by double strand break repair proteins at a single DNA end: recruitment of DNA-PKcs induces inward translocation of Ku protein', *Nucleic Acids Res,* 27(24), pp. 4679-86.

Yuan, P., Di, L., Zhang, X., Yan, M., Wan, D., Li, L., Zhang, Y., Cai, J., Dai, H., Zhu, Q., Hong, R. and Xu, B. (2015) 'Efficacy of oral Etoposide in pretreated metastatic breast cancer: a multicenter phase 2 study', *Medicine (Baltimore),* 94(17), pp. e774.

Zahradka, P. and Ebisuzaki, K. (1982) 'A shuttle mechanism for DNA-protein interactions. The regulation of poly(ADP-ribose) polymerase', *Eur J Biochem,* 127(3), pp. 579-85.

Zarin, D. A., Tse, T., Williams, R. J. and Carr, S. (2016) 'Trial Reporting in ClinicalTrials.gov - The Final Rule', *N Engl J Med.*

Zechiedrich, E. L., Christiansen, K., Andersen, A. H., Westergaard, O. and Osheroff, N. (1989) 'Double-stranded DNA cleavage/religation reaction of eukaryotic topoisomerase II: evidence for a nicked DNA intermediate', *Biochemistry,* 28(15), pp. 6229-36.

Zeitlin, B. D., Joo, E., Dong, Z., Warner, K., Wang, G., Nikolovska-Coleska, Z., Wang, S. and Nör, J. E. (2006) 'Antiangiogenic effect of TW37, a small-molecule inhibitor of Bcl-2', *Cancer Res,* 66(17), pp. 8698-706.

Zeng, Z., Cortés-Ledesma, F., El Khamisy, S. F. and Caldecott, K. W. (2011) 'TDP2/TTRAP is the major 5'-tyrosyl DNA phosphodiesterase activity in vertebrate cells and is critical for cellular resistance to topoisomerase II-induced DNA damage.', *J Biol Chem,* 286(1), pp. 403-9.

Zeng, Z., Sharma, A., Ju, L., Murai, J., Umans, L., Vermeire, L., Pommier, Y., Takeda, S., Huylebroeck, D., Caldecott, K. W. and El-Khamisy, S. F. (2012) 'TDP2 promotes repair of topoisomerase I-mediated DNA damage in the absence of TDP1', *Nucleic Acids Res,* 40(17), pp. 8371-80.

Zhang, F., Ma, J., Wu, J., Ye, L., Cai, H., Xia, B. and Yu, X. (2009) 'PALB2 links BRCA1 and BRCA2 in the DNA-damage response', *Curr Biol,* 19(6), pp. 524-9.

Zhang, F., Wang, Y., Wang, L., Luo, X., Huang, K., Wang, C., Du, M., Liu, F., Luo, T. and Huang, D. (2013) 'Poly(ADP-ribose) polymerase 1 is a key regulator of estrogen receptor α-dependent gene transcription', *J Biol Chem,* 288(16), pp. 11348-57.

Zhang, Q., Major, M. B., Takanashi, S., Camp, N. D., Nishiya, N., Peters, E. C., Ginsberg, M. H., Jian, X., Randazzo, P. A., Schultz, P. G., Moon, R. T. and Ding, S. (2007) 'Small-molecule synergist of the Wnt/beta-catenin signaling pathway', *Proc Natl Acad Sci U S A,* 104(18), pp. 7444-8.

Zhen, D. B., Rabe, K. G., Gallinger, S., Syngal, S., Schwartz, A. G., Goggins, M. G., Hruban, R. H., Cote, M. L., McWilliams, R. R., Roberts, N. J., Cannon-Albright, L. A., Li, D., Moyes, K., Wenstrup, R. J., Hartman, A. R., Seminara, D., Klein, A. P. and Petersen, G. M. (2015) 'BRCA1, BRCA2, PALB2, and CDKN2A mutations in familial pancreatic cancer: a PACGENE study', *Genet Med,* 17(7), pp. 569-77.

Appendices

DMSO		Mock kd				TDP2 kd			
		Day 1	Day 3	Day 5	Day 7	Day 1	Day 3	Day 5	Day 7
WT	Day 1	0.295				0.065			
	Day 3		0.010				0.001		
	Day 5			0.041				0.038	
	Day 7				0.037				0.005
Mock kd	Day 1					0.053			
	Day 3						0.028		
	Day 5							0.232	
	Day 7								0.258

250 nM Etoposide		Mock kd				TDP2 kd			
		Day 1	Day 3	Day 5	Day 7	Day 1	Day 3	Day 5	Day 7
WT	Day 1	0.293				0.019			
	Day 3		0.065				0.002		
	Day 5			0.002				0.008	
	Day 7				0.023				0.002
Mock kd	Day 1					0.102			
	Day 3						0.054		
	Day 5							0.127	
	Day 7								0.087

500 nM Etoposide		Mock kd				TDP2 kd			
		Day 1	Day 3	Day 5	Day 7	Day 1	Day 3	Day 5	Day 7
WT	Day 1	0.179				0.047			
	Day 3		0.063				0.018		
	Day 5			0.004				0.000	
	Day 7				0.014				0.001
Mock kd	Day 1					0.065			
	Day 3						0.011		
	Day 5							0.014	
	Day 7								0.029

750 nM Etoposide		Mock kd				TDP2 kd			
		Day 1	Day 3	Day 5	Day 7	Day 1	Day 3	Day 5	Day 7
WT	Day 1	0.074				0.007			
	Day 3		0.011				0.020		
	Day 5			0.010				0.001	
	Day 7				0.044				0.001
Mock kd	Day 1					0.019			
	Day 3						0.036		
	Day 5							0.034	
	Day 7								0.097

1 µM Etoposide		Mock kd				TDP2 kd			
		Day 1	Day 3	Day 5	Day 7	Day 1	Day 3	Day 5	Day 7
WT	Day 1	0.200				0.066			
	Day 3		0.157				0.032		
	Day 5			0.009				0.008	
	Day 7				0.082				0.027
Mock kd	Day 1					0.050			
	Day 3						0.036		
	Day 5							0.150	
	Day 7								0.060

Appendix 1: P-values represent paired, 1-tailed t-test of Alamar Blue results of Fig.9.1B, C and D.

p-values	WT vs Mock kd	WT vs TDP2 kd	Mock kd vs TDP2 kd
15μM Etop	0.231	0.044	0.011
30μM Etop	0.221	0.014	0.011
45μM Etop	0.025	0.027	0.042
60μM Etop	0.457	0.067	0.047

Appendix 2: P-values represent paired, 1-tailed t-test of clonogenic cell survival assay of Fig.9.1E.
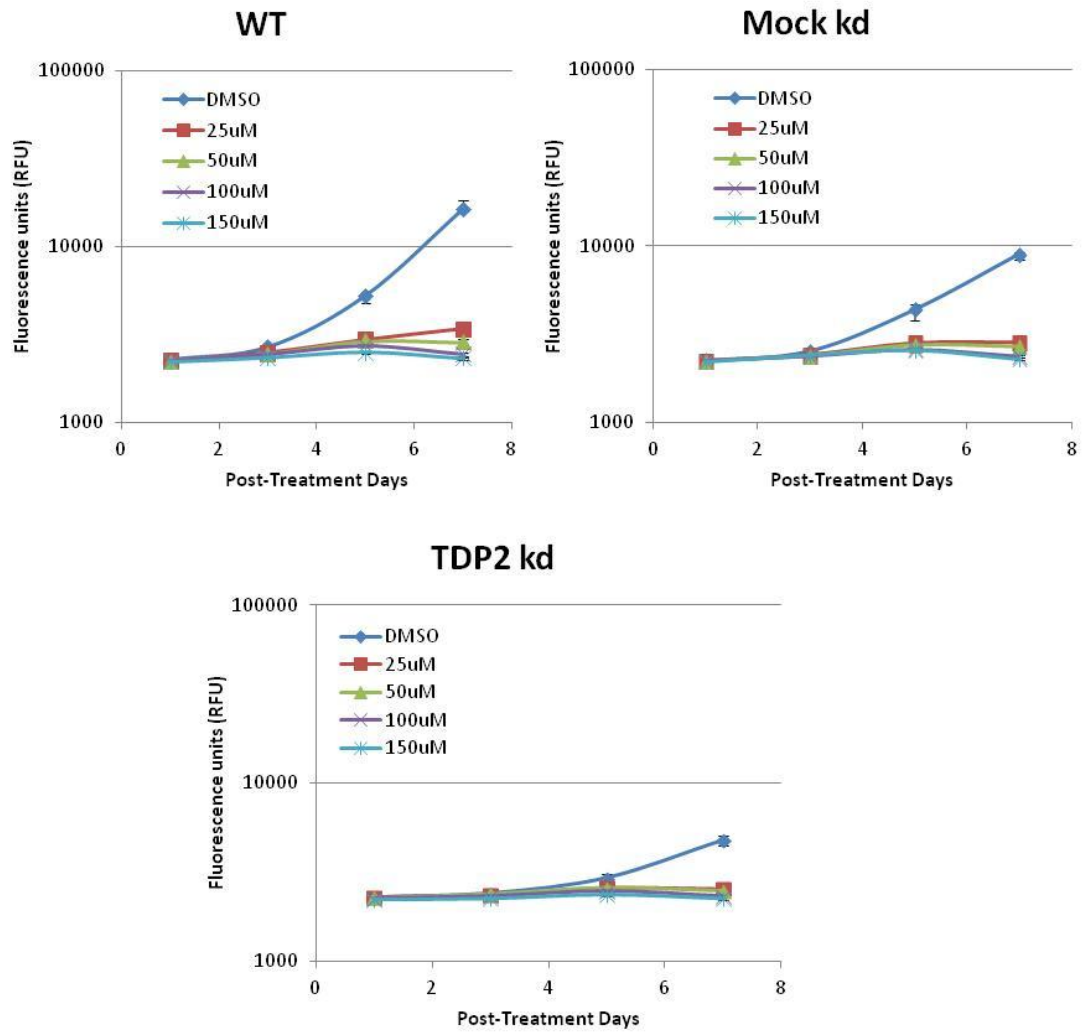
Appendix 3: Sensitivity to very high concentrations of etoposide in TDP2-depleted MCF7 breast cancer cells: Sensitivity to increasing concentrations of etoposide was measured using Alamar Blue assay in the indicated breast cancer cell lines. Etoposide concentrations used were 0µM, 25µM, 50µM, 100µM and 150µM. Cells were treated for 3h with etoposide or DMSO. Fluorescence measurements were taken at days 1, 3, 5 and 7 after treatment. These measurements were plotted against the corresponding day. Data are the mean (+/- SEM) from three independent experiments.

WT

Mock kd

TDP2 kd
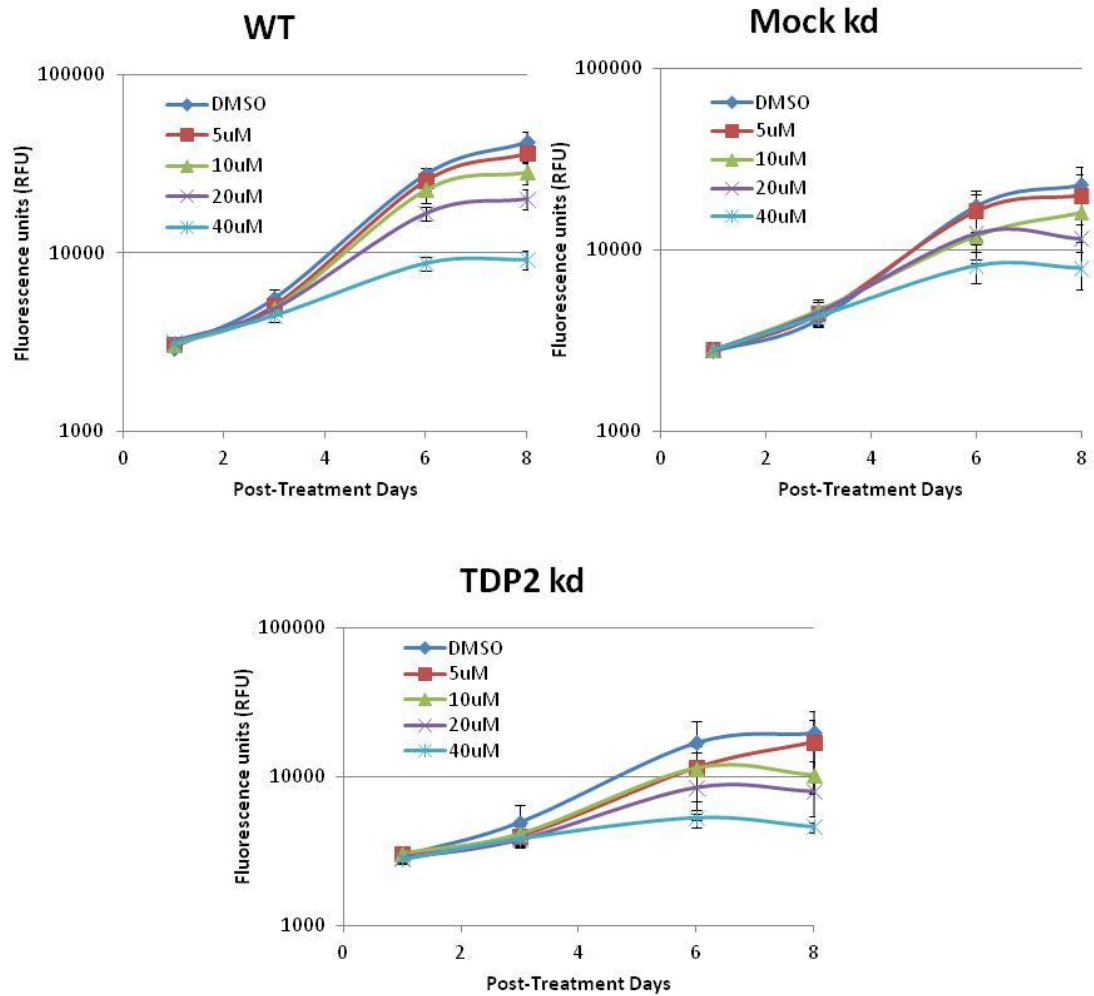
Appendix 4: Sensitivity to high concentrations of etoposide in TDP2-depleted MCF7 breast cancer cells: Sensitivity to increasing concentrations of etoposide was measured using Alamar Blue assay in the indicated breast cancer cell lines. Etoposide concentrations used were 0μM, 5μM, 10μM, 20μM and 40μM. Cells were treated for 3h with etoposide or DMSO. Fluorescence measurements were taken at days 1, 3, 6 and 8 after treatment. These measurements were plotted against the corresponding day. Data are the mean (+/- SEM) from three independent experiments.
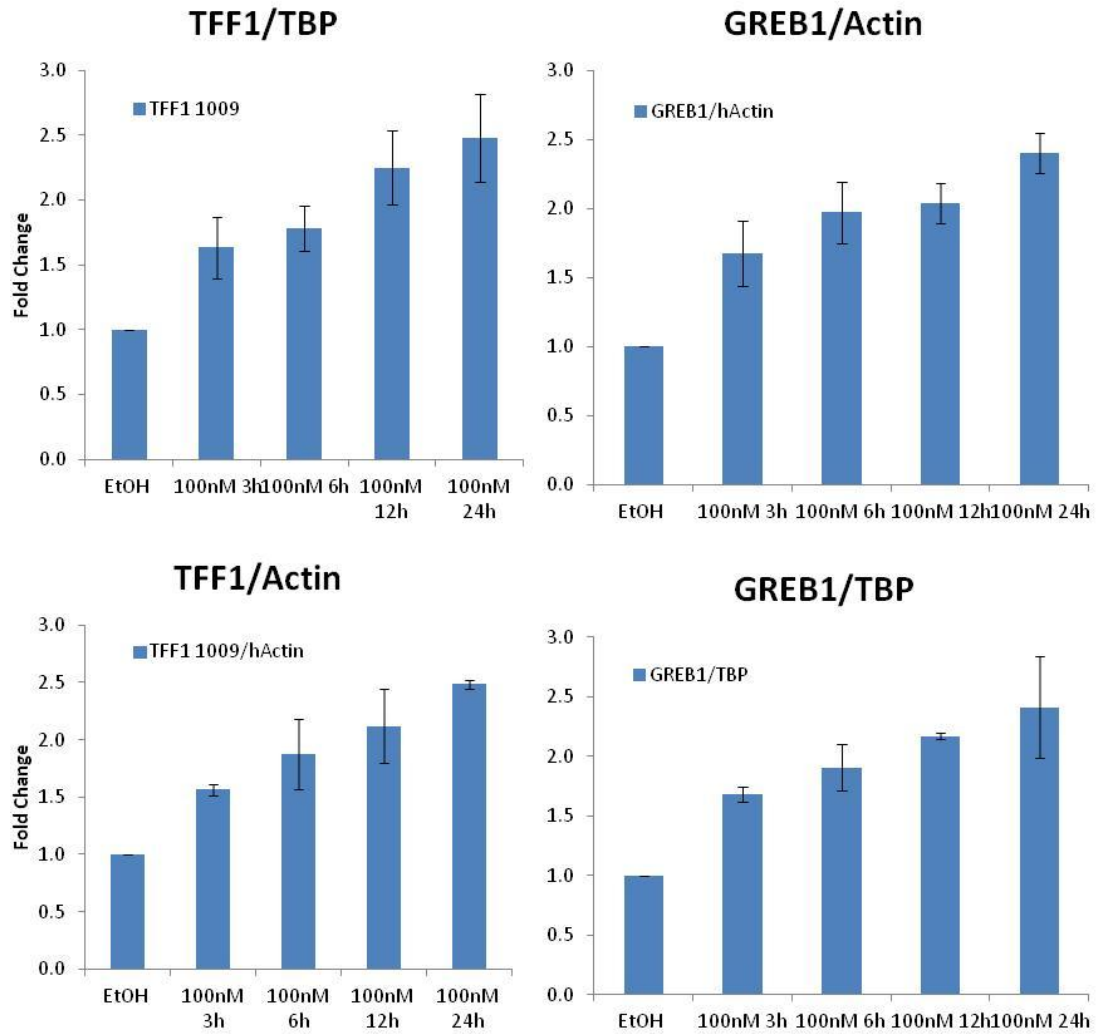
Appendix 5: Induction of transcription following estradiol treatment. TFF1 and GREB1 mRNA expression levels were measured using RT-qPCR in the MCF7 breast cancer cells after treatment with estradiol for 3h, 6h, 12h and 24h, and are expressed relative to the expression level of the TBP or Actin housekeeping gene. Relative expression (fold-change) was calculated from the equation relative expression=$2^{-\Delta Ct}$, where ΔCt is calculated by subtracting the Ct value of the housekeeping gene from the Ct value of the gene of interest. Data are the mean (+/- SEM) from three independent experiments.

Appendix 6: A collection of gene mutation status in a panel of breast cancer cell lines. No indicates wild type gene status and Yes indicates mutant gene status.

Gene Name	BT474	MCF7	MM361	MM453	T47D	ZR751
A2ML1			Yes			
AAK1	Yes	Yes		Yes		
AARS	Yes	Yes				
AARS2				Yes		
AASS	Yes					
ABCA1	Yes		Yes			
ABCA10					Yes	
ABCA12	Yes		Yes			
ABCA13			Yes	Yes		
ABCA2	Yes			Yes		
ABCA3			Yes			
ABCA4			Yes			
ABCA6				Yes		
ABCA7				Yes		
ABCA8		Yes			Yes	
ABCA9					Yes	
ABCB1			Yes			
ABCB11			Yes			
ABCB4		Yes				
ABCB6		Yes				
ABCC1			Yes			
ABCC12			Yes	Yes		
ABCC3	Yes					
ABCC5	Yes			Yes	Yes	
ABCC8	Yes					
ABCE1				Yes		
ABHD1			Yes			
ABHD16A	Yes					
ABHD2	Yes		Yes			
ABHD3			Yes			
ABI3BP		Yes				
ABL2			Yes			
ABLIM1			Yes			
ABLIM3	Yes					
AC007731_16	Yes					
AC010872_2		Yes				
AC092031.1					Yes	
AC112491_4			Yes			
AC142381_2			Yes			
ACACA					Yes	
ACACB			Yes	Yes		
ACAD10				Yes		
ACAN				Yes		
ACAT1				Yes		
ACAT2				Yes		
ACCN1			Yes	Yes		
ACCN3		Yes				
ACE		Yes				
ACO1			Yes			
ACOT12		Yes				
ACPP				Yes		
ACSM3					Yes	
ACSM4			Yes			
ACSM5			Yes			
ACTB	Yes					
ACTG2		Yes				
ACTL7A				Yes		
ACTL7B					Yes	
ACTN2					Yes	
ACTR8				Yes		
ACVR1		Yes			Yes	
ACVR1C			Yes			
ACVR2A	Yes					
ACVR2B	Yes					
ADAM17	Yes		Yes		Yes	Yes

ADAM18				Yes		
ADAM20	Yes					
ADAM22				Yes		
ADAM28	Yes		Yes	Yes	Yes	Yes
ADAM29				Yes		
ADAMTS1				Yes		
ADAMTS12					Yes	
ADAMTS15	Yes					
ADAMTS16		Yes				
ADAMTS18		Yes				
ADAMTS19			Yes			
ADAMTS3		Yes				
ADAMTS6			Yes			
ADAMTSL2				Yes		
ADAMTSL3	Yes		Yes			
ADAP1	Yes					
ADARB2			Yes			
ADCK1					Yes	
ADCK3				Yes		
ADCY2			Yes	Yes		
ADCY4		Yes		Yes		
ADCY6	Yes			Yes		
ADCYAP1R1	Yes					
ADD1	Yes	Yes				
ADD2			Yes			
ADD3			Yes			
ADGB		Yes		Yes		
ADK				Yes		
ADORA3	Yes					
ADPRHL1	Yes					
ADRA1D	Yes					
ADRM1		Yes				
AES	Yes					
AFF3	Yes			Yes		
AGAP1				Yes		
AGAP2			Yes			
AGBL1	Yes	Yes				
AGBL4					Yes	
AGC1				Yes		
AGER	Yes					
AGFG1		Yes				
AGPAT6			Yes			
AGPS				Yes		
AHCTF1					Yes	
AHI1			Yes			
AHNAK	Yes	Yes				
AHNAK2	Yes		Yes	Yes		
AHSA2	Yes					
AIRE			Yes			
AK5				Yes		
AK7			Yes			
AKAP10			Yes			
AKAP12	Yes		Yes	Yes	Yes	
AKAP2	Yes					
AKAP3	Yes			Yes		
AKAP6			Yes			
AKAP8L				Yes		
AKAP9	Yes			Yes		
AKR1B1				Yes		
AKR1C4				Yes		
AKT1	No	No		No	No	No
AL589988.1	Yes					
ALB			Yes			
ALCAM	Yes					
ALDH16A1			Yes			
ALDH1L1					Yes	
ALDH8A1		Yes				
ALDOB				Yes		

Gene						
ALG10B				Yes		
ALG11				Yes		
ALG3			Yes			
ALKBH1				Yes		
ALKBH3			Yes			
ALOXE3				Yes		
ALPK1	Yes	Yes				
ALPK2	Yes	Yes		Yes	Yes	Yes
ALPK3			Yes	Yes		
ALS2CL		Yes				
ALS2CR11				Yes		
ALS2CR8			Yes			
AMACR				Yes		
AMIGO2			Yes			
AMOTL1			Yes			
ANAPC1	Yes					
ANAPC4			Yes			
ANAPC5			Yes	Yes		
ANGPT1	Yes					
ANGPTL1				Yes		
ANGPTL3			Yes			
ANK2	Yes			Yes		
ANK3	Yes	Yes				
ANKFN1		Yes				
ANKHD1					Yes	
ANKLE2			Yes			
ANKRD1	Yes					
ANKRD12			Yes			
ANKRD18A			Yes			
ANKRD26	Yes			Yes		
ANKRD27			Yes			
ANKRD30A					Yes	
ANKRD30B				Yes		
ANKRD30BL			Yes			
ANKRD34A	Yes					
ANKRD36			Yes			
ANKRD36C			Yes	Yes		
ANKRD56	Yes					
ANKRD6				Yes		
ANO1		Yes	Yes			
ANO2			Yes			
ANO3	Yes		Yes			
ANO4				Yes		
ANO9			Yes			
ANTXR1			Yes			
ANXA1	Yes					
ANXA4	Yes					
AOX1	Yes					
AP1G2	Yes			Yes		
AP1M1	Yes					
AP2A1				Yes		
AP2A2				Yes		
AP2M1			Yes			
AP3B1				Yes		
APBA2					Yes	
APBB2					Yes	
APC	No	No	No	No	No	No
APC2					Yes	
APEX2	Yes					
APOA5	Yes					
APOLD1			Yes			
APP			Yes			
APPL1				Yes		
AQP8	Yes					
AR				Yes		
ARAP1				Yes		
ARFGEF1			Yes			
ARFGEF2	Yes					

ARFIP1				Yes		
ARG1			Yes			
ARG2	Yes					
ARHGAP11A					Yes	
ARHGAP17			Yes			
ARHGAP19	Yes					
ARHGAP21	Yes					
ARHGAP22			Yes			
ARHGAP25		Yes				
ARHGAP29	Yes					
ARHGAP31			Yes			
ARHGAP32			Yes			
ARHGAP39			Yes			
ARHGAP4				Yes		
ARHGEF1	Yes					
ARHGEF12	Yes					
ARHGEF15	Yes					
ARHGEF18		Yes	Yes			
ARHGEF26	Yes					
ARHGEF38				Yes		
ARHGEF40	Yes					
ARID1A		Yes	Yes		Yes	
ARID4B			Yes			
ARL13B		Yes				
ARL4C				Yes		
ARL6			Yes			
ARL6IP4	Yes					
ARMC3	Yes					
ARMC7			Yes			
ARMC9				Yes		
ARNT			Yes			
ARNTL2			Yes			
ARRDC4		Yes				
ARSB					Yes	
ARSD				Yes		
ASAP1				Yes		
ASB10			Yes			
ASB11				Yes		
ASB12				Yes		
ASB9			Yes			
ASCC1		Yes				
ASGR1		Yes				
ASH1L		Yes				
ASH2L			Yes			
ASPH				Yes		Yes
ASPM		Yes	Yes	Yes		
ASPSCR1			Yes			
ASTE1	Yes		Yes			
ASTN1				Yes		
ASTN2			Yes			
ASXL2				Yes		
ATAD2B	Yes		Yes			
ATAD5	Yes					
ATF1			Yes			
ATF7IP					Yes	
ATG13		Yes				
ATG2A		Yes				
ATG2B			Yes			
ATG4C	Yes					
ATM	Yes					
ATN1			Yes			
ATP10A					Yes	
ATP10B			Yes			
ATP10D	Yes	Yes				
ATP11B				Yes		
ATP11C				Yes		
ATP13A4			Yes			
ATP13A5			Yes			

ATP2A2	Yes				
ATP2B1	Yes				
ATP2B3		Yes			
ATP2B4		Yes			
ATP2C1				Yes	
ATP2C2			Yes		
ATP5A1				Yes	
ATP5C1	Yes				
ATP6V0A1	Yes				
ATP6V0A4				Yes	
ATP6V1A				Yes	
ATP6V1B1				Yes	
ATP8B1				Yes	
ATP8B2	Yes				
ATP8B3			Yes	Yes	
ATP9A		Yes			
ATR	No	No		Yes	
ATRIP			Yes		
ATRNL1				Yes	
ATRX			Yes		
ATXN7				Yes	
AUTS2			Yes		
AXIN1		No		No	No	No
B3GNT4				Yes	
B3GNT5			Yes		
BAALC			Yes		
BAAT		Yes			
BACH1	Yes				
BAG6			Yes		
BAI1				Yes	
BAI2			Yes		
BAI3		Yes			
BANK1			Yes		Yes
BAP1			Yes		
BASP1				Yes	
BAT2D1	Yes	Yes	Yes		
BAZ2B		Yes		Yes	
BBS10		Yes			
BBS4		Yes			
BCAP31			Yes		
BCAS1			Yes		Yes
BCKDK			Yes		
BCL11B				Yes	
BCL2L13		Yes			
BCL2L15	Yes				
BCMO1			Yes		
BCOR	Yes				
BCORL1				Yes	
BCR	Yes	Yes	Yes		Yes	Yes
BDKRB1	Yes				
BDP1					Yes
BEND3				Yes	
BET1L				Yes	
BFAR	Yes				
BICD1	Yes		Yes		
BIN1					Yes
BIN3			Yes		
BIRC6			Yes	Yes	
BLK	Yes				
BMP10				Yes	
BMP2	Yes			Yes	
BMP6	Yes				
BMPR2			Yes		
BNC1		Yes			
BOC				Yes	
BOD1L				Yes	
BOK	Yes				
BPIL3				Yes	

BPTF			Yes			
BRAF	No	No	No	No	No	No
BRAP				Yes		
BRCA1	No	No	No	No	No	No
BRCA2	Yes		Yes			
BRD2			Yes			
BRD4				Yes		
BRD7	Yes					
BRDT	Yes					
BRIP1	Yes					
BRMS1				Yes		
BRPF1		Yes				
BRWD1		Yes	Yes			
BSPH1					Yes	
BTBD11					Yes	
BTBD2			Yes			
BTBD6		Yes				
BTBD8			Yes			
BTG1			Yes			
BTG3	Yes					
BTN1A1	Yes					
BTN2A1					Yes	
BTN3A3	Yes					
BTNL8	Yes					
BUB1B	Yes					
BUB3			Yes			
BVES		Yes				
C10orf11			Yes			
C10orf112	Yes		Yes			
C10orf129				Yes		
C10orf137			Yes			
C10orf18				Yes		
C10orf31			Yes			
C10orf35				Yes		
C10orf46			Yes			
C10orf55				Yes		
C10orf6			Yes			
C10orf71				Yes		
C10orf85			Yes			
C10orf92				Yes		
C10orf96			Yes			
C11orf63	Yes					
C11orf80				Yes		
C12orf4		Yes				
C12orf51			Yes			
C12orf73				Yes		
C12orf77		Yes				
C14orf102		Yes				
C14orf118			Yes			
C14orf39			Yes			
C14orf43				Yes		
C15orf27	Yes					
C15orf43			Yes			
C15orf55		Yes	Yes		Yes	
C15orf59					Yes	
C16orf13			Yes			
C16orf46			Yes			
C16orf71			Yes			
C16orf85	Yes					
C16orf88	Yes					
C16orf93	Yes					
C17ORF37				Yes		
C17orf39	Yes					
C17orf46	Yes					
C17orf57				Yes		
C17orf63	Yes					
C17orf70		Yes				
C18orf34			Yes			

Gene						
C18orf54				Yes		
C19orf28			Yes			
C19orf29			Yes			
C19orf40	Yes					
C19orf44	Yes					
C19orf55				Yes		
C19orf69			Yes			
C1orf106			Yes			
C1orf112			Yes			
C1orf127	Yes					
C1orf129				Yes		
C1orf159				Yes		
C1orf164	Yes					
C1orf173			Yes			
C1orf186				Yes		
C1orf187		Yes				
C1orf54			Yes			
C1orf86				Yes		
C1QTNF7		Yes				
C1QTNF9B				Yes		
C1R				Yes		
C20orf132			Yes			
C20orf152	Yes					
C20orf173	Yes					
C20orf195	Yes					
C20orf78	Yes					
C21orf33				Yes		
C22orf29					Yes	
C22orf42			Yes			
C22orf43	Yes					
C2CD2			Yes			
C2orf67		Yes			Yes	
C2orf71			Yes		Yes	
C2orf77			Yes			
C3AR1				Yes		
C3orf15			Yes	Yes		
C3orf21	Yes					
C3orf65					Yes	
C3orf67			Yes			
C3orf77				Yes		
C4BPA	Yes			Yes		
C4orf21	Yes			Yes		
C4orf29					Yes	
C4orf41				Yes		
C5orf28				Yes		
C5orf38			Yes			
C5orf44					Yes	
C5orf5			Yes			
C6orf103		Yes		Yes		
C6orf146	Yes					
C6orf163	Yes					
C6orf165	Yes					
C6orf168				Yes		
C6orf174				Yes		
C6orf70		Yes				
C6orf89			Yes			
C7orf31				Yes		
C7orf44				Yes		
C8A			Yes			
C8G		Yes				
C8orf48				Yes		
C8orf76	Yes					
C9orf102			Yes			
C9orf117	Yes					
C9orf125			Yes			
C9orf128			Yes			
C9orf131			Yes			
C9orf156		Yes				

Gene	C1	C2	C3	C4	C5	C6
C9orf174	Yes					
C9orf24	Yes					
C9orf48	Yes					
C9orf5			Yes			
C9orf78	Yes					
C9orf80		Yes				
C9orf84				Yes		
C9orf93				Yes		
CA2				Yes		
CAB39				Yes		
CABC1	Yes		Yes	Yes		
CACNA1A	Yes		Yes			
CACNA1B			Yes			
CACNA1C				Yes		
CACNA1E		Yes	Yes			
CACNA1I			Yes			
CACNA1S				Yes		
CACNA2D2					Yes	
CACNA2D3	Yes					
CACNA2D4	Yes			Yes		
CACNB2	Yes	Yes	Yes	Yes	Yes	Yes
CACNG7				Yes		
CAD				Yes		
CADPS2	Yes		Yes			
CAGE1					Yes	
CALM3				Yes		
CALN1	Yes					
CAMK1G			Yes			
CAMTA2			Yes	Yes		
CANX				Yes		
CAPN11		Yes				
CAPN12			Yes			
CAPN13				Yes		
CAPN14			Yes	Yes		
CAPN3	Yes					
CAPN5				Yes		
CAPRIN2				Yes		
CARD11	Yes					
CARKD			Yes			
CASC3	Yes					
CASKIN1	Yes					
CASKIN2	Yes					
CASP12		Yes				
CASP14	Yes					
CASP2	Yes					
CASP3		Yes				
CASP7			Yes			
CASP8AP2			Yes	Yes		
CAT			Yes			
CATSPER1			Yes	Yes		
CATSPER3				Yes		
CATSPERG	Yes					
CBL	Yes					
CBR3			Yes			
CBX2			Yes			
CBX5	Yes					
CBX8	Yes					
CCAR1			Yes			
CCDC116				Yes		
CCDC132				Yes		
CCDC136		Yes				
CCDC140			Yes			
CCDC163P			Yes			
CCDC165				Yes		
CCDC33	Yes					
CCDC34					Yes	
CCDC38			Yes	Yes		
CCDC39			Yes			

Gene	Col1	Col2	Col3	Col4	Col5	Col6
CCDC48			Yes			
CCDC60	Yes					
CCDC64B		Yes				
CCDC66		Yes				
CCDC67				Yes		
CCDC74A				Yes		
CCDC77			Yes	Yes		
CCDC83					Yes	
CCDC87	Yes					
CCDC88C	Yes					
CCDC89			Yes			
CCDC97			Yes			
CCK	Yes					
CCKAR		Yes				
CCNA1				Yes		
CCNB1IP1			Yes			
CCNB3			Yes			
CCND1	Yes					
CCNE2	Yes					
CCNF			Yes			
CCNL1		Yes				
CCNT1			Yes			
CCNY	Yes					
CCR8					Yes	
CCT4	Yes					
CCT8				Yes		
CD163				Yes		
CD177		Yes	Yes			
CD19			Yes			
CD1A					Yes	
CD1C			Yes			
CD1E					Yes	
CD200		Yes				
CD302		Yes				
CD3E		Yes				
CD68		Yes				
CD69			Yes			
CD80	Yes					
CD97	Yes					
CDAN1				Yes		
CDC14A	Yes					
CDC25B	Yes					
CDC40				Yes		
CDC42	No	No				
CDC42BPA		Yes				
CDC42BPB		Yes		Yes		
CDC42BPG				Yes		
CDCP1		Yes				
CDCP2			Yes			
CDH1	No	No	No	Yes	No	No
CDH10					Yes	
CDH18			Yes			
CDH2	Yes			Yes		
CDH20			Yes			
CDH22				Yes		
CDH24		Yes			Yes	
CDH3	Yes					
CDH6	Yes			Yes		
CDH7	Yes					
CDHR5			Yes			
CDK11A	Yes					
CDK11B	Yes	Yes	Yes	Yes	Yes	
CDK13						Yes
CDK18					Yes	
CDK2	No	No				
CDK4	No	No				
CDK5RAP2	Yes	Yes				
CDK6	No	No				

Gene						
CDK8	Yes					
CDKL4			Yes			
CDKN1B			Yes			
CDKN2A	No	Yes	Yes	No	No	No
CDKN2C	No	No	No	No	No	No
CDRT15				Yes		
CDS2					Yes	
CDT1				Yes		
CEACAM8				Yes		
CELF5			Yes			
CELSR1	Yes					
CELSR2	Yes					
CENPB	Yes					
CENPE	Yes		Yes			
CENPF	Yes		Yes			
CENPH				Yes		
CENPJ			Yes			
CENPT	Yes			Yes		
CEP110			Yes			
CEP112	Yes					
CEP135			Yes			
CEP170				Yes		
CEP192	Yes					
CEP290			Yes			
CEP350			Yes			
CEP68	Yes				Yes	
CEP70		Yes				
CEP78			Yes			
CEP85				Yes		
CEP95	Yes					
CERK		Yes				
CES2				Yes		
CFH			Yes			
CFI			Yes			
CGI-77	Yes					
CGN			Yes			
CHAF1B			Yes			
CHAT				Yes		
CHCHD3			Yes			
CHD1			Yes	Yes		Yes
CHD2			Yes			
CHD5	Yes					
CHD7				Yes		
CHD8			Yes	Yes		
CHD9						Yes
CHEK2	No	No				
CHKA				Yes		
CHL1	Yes			Yes		
CHMP2A					Yes	
CHN2	Yes					
CHST1				Yes		
CHST2					Yes	
CHST9		Yes				
CHSY1		Yes				
CHTF18				Yes		
CHTOP				Yes		
CIB3		Yes				
CIC	Yes					
CIITA				Yes		
CINP				Yes		
CIR1	Yes					
CIRH1A			Yes			
CIT					Yes	
CKAP2L		Yes				
CLCN2			Yes			
CLCN3		Yes		Yes		
CLDN1		Yes				
CLDN16					Yes	

CLEC4D				Yes		
CLEC4F			Yes			
CLEC4G	Yes					
CLEC7A				Yes		
CLINT1			Yes			
CLIP1		Yes	Yes			
CLIP3			Yes			
CLK1		Yes				
CLK3						Yes
CLN8	Yes					
CLPB		Yes				
CLSPN			Yes			
CLSTN3			Yes			
CLTA	Yes					
CLTCL1	Yes	Yes	Yes		Yes	Yes
CLU	Yes					
CLUL1	Yes					
CLVS1			Yes			
CMAHP	Yes					
CMKLR1			Yes			
CMTM1			Yes			
CNDP1		Yes			Yes	
CNGA1	Yes					
CNIH			Yes			
CNKSR1		Yes				
CNKSR3	Yes					
CNNM2				Yes		
CNNM4	Yes					
CNOT1	Yes					
CNTLN	Yes					
CNTN2			Yes			
CNTN3		Yes				
CNTN4		Yes		Yes		
CNTN5			Yes			
CNTN6	Yes			Yes	Yes	
CNTNAP2	Yes		Yes			
CNTNAP3	Yes					
CNTNAP4		Yes				
CNTNAP5			Yes			
CNTRL			Yes	Yes		
COG4	Yes		Yes			
COG8	Yes					
COL11A1			Yes			
COL12A1			Yes			
COL14A1				Yes		
COL17A1	Yes					
COL18A1					Yes	
COL19A1			Yes			
COL1A1			Yes			
COL1A2		Yes			Yes	
COL22A1			Yes			
COL24A1			Yes			
COL2A1	Yes					
COL3A1			Yes			
COL5A1			Yes			
COL6A2			Yes			
COL6A6	Yes		Yes			
COL7A1	Yes					
COL8A1	Yes		Yes			
COL9A1		Yes				
COL9A3	Yes					
COMP	Yes					
COPB1	Yes					
COPS4					Yes	
COPS5			Yes	Yes		
CORO1A	Yes					
CORO2B	Yes					
CORO6	Yes					

CORO7			Yes	Yes		
COX10	Yes					
COX7C			Yes			
CP			Yes	Yes		
CPAMD8		Yes				
CPB2				Yes		
CPD	Yes					
CPE				Yes		
CPLX2				Yes		
CPLX4		Yes				
CPNE1			Yes	Yes		
CPNE9					Yes	
CPO				Yes		
CPS1	Yes					
CPXM2					Yes	
CR1			Yes			
CR1L				Yes		
CRB2				Yes		
CREB1	Yes			Yes		
CREB3L2	Yes		Yes			
CREBBP	No	No	No	Yes		No
CREBZF				Yes		
CREG2	Yes					
CRIP1			Yes	Yes		
CROT			Yes			
CRSP3					Yes	
CRTAC1					Yes	
CRYBG3			Yes			
CRYGN	Yes					
CSDC2		Yes				
CSE1L	Yes					
CSF1R			Yes		Yes	
CSGALNACT2			Yes			
CSMD1	Yes		Yes			
CSMD2				Yes		
CSMD3	Yes		Yes			
CSNK1G3	Yes					
CSPG4		Yes			Yes	
CSRP1			Yes			
CSTF1		Yes				
CSTF2T		Yes		Yes		
CTBP2		Yes		Yes	Yes	
CTC1	Yes					
CTCFL			Yes			
CTDSP2	Yes					
CTNNA1	No	No	No	No	No	No
CTNNA2	Yes					
CTNNB1	No	No	No	No	No	No
CTPS	Yes					
CTPS1	Yes					
CTR9	Yes					
CTTNBP2			Yes			
CTU1				Yes		
CTU2	Yes				Yes	
CUBN			Yes	Yes		
CUL7		No				
CUL9	Yes	Yes		Yes		
CUX1	Yes					
CUX2	Yes					
CWC22			Yes			
CWF19L1	Yes					
CXCL12					Yes	
CXCR1				Yes		
CXorf57			Yes			
CXorf58				Yes		
CXorf67	Yes					
CXXC4			Yes			
CYB5R4			Yes			

Gene						
CYC1	Yes					
CYFIP2		Yes				
CYP11B1			Yes			
CYP11B2	Yes	Yes				
CYP17A1			Yes			
CYP27B1				Yes		
CYP2A6				Yes		
CYP2A7		Yes		Yes		
CYP3A43			Yes			
CYP3A5	Yes					
CYP3A7	Yes					
CYP7A1	Yes					
CYTH1		Yes				
DAAM1			Yes			
DAB1			Yes			
DACH1	Yes		Yes			
DAG1	Yes					
DAGLA		Yes				
DAGLB			Yes			
DAPK1	Yes	Yes				
DARS2	Yes					
DAXX			Yes			
DBF4B					Yes	
DBN1	Yes					
DBT			Yes			
DCAF12			Yes			
DCAF17	Yes					
DCAF4			Yes			
DCAKD				Yes		
DCDC1			Yes			
DCDC5				Yes	Yes	
DCHS2	Yes	Yes		Yes		
DCLK2			Yes	Yes		Yes
DCLRE1A			Yes			
DCTN2			Yes			
DCTN4	Yes					
DCXR			Yes		Yes	
DDHD1		Yes				
DDR2					Yes	
DDRGK1	Yes					
DDX1	Yes				Yes	
DDX12			Yes			
DDX27	Yes					
DDX39B		Yes				
DDX4	Yes					
DDX50				Yes		
DDX51	Yes					
DDX55			Yes			
DDX60			Yes			
DDX60L	Yes	Yes				
DEDD2					Yes	
DEFB118					Yes	
DEFB133					Yes	
DENND1C		Yes				
DENND2A			Yes			
DENND4A			Yes			
DENND4C	Yes			Yes		
DEPDC1				Yes		
DFFA				Yes		
DFNB31	Yes		Yes			
DGKA			Yes			
DGKB				Yes		
DGKI		Yes	Yes			
DHCR24			Yes			
DHFRL1				Yes	Yes	
DHX16		Yes				
DHX34				Yes		
DHX38			Yes			

DIAPH2	Yes		Yes		
DIDO1		Yes	Yes	Yes	
DIP2C			Yes		
DIS3L				Yes	
DISC1	Yes			Yes	
DISP2			Yes		
DKFZP434P1750			Yes		
DKK1	Yes				
DLAT			Yes		
DLEU7				Yes	
DLG2			Yes		
DLGAP1			Yes		
DLL3				Yes	
DLX3					Yes
DMBT1			Yes	Yes	Yes
DMD	Yes		Yes		
DMPK	Yes					Yes
DNAAF1			Yes		
DNAH1			Yes		
DNAH10	Yes		Yes		
DNAH11	Yes		Yes		Yes
DNAH12	Yes				
DNAH14	Yes				
DNAH17	Yes		Yes	Yes	
DNAH2		Yes	Yes		
DNAH5	Yes	Yes			
DNAH7	Yes		Yes	Yes	Yes
DNAH8	Yes	Yes	Yes		Yes
DNAJB7			Yes		
DNAJC10			Yes		
DNAJC11				Yes	
DNAJC16	Yes				
DNAJC22			Yes		
DNAJC7					Yes
DNASE1L1		Yes			
DNASE1L2			Yes		
DNHD1		Yes			
DNM2		Yes			
DNM3					Yes
DNPEP				Yes	
DOC2A		Yes			
DOCK6	Yes				
DOCK8		Yes			
DOCK9			Yes		
DOK4			Yes		
DOT1L			Yes		
DPAGT1		Yes			
DPCR1			Yes	Yes	
DPH1			Yes		
DPP10	Yes				
DPP4			Yes		
DPP6	Yes			Yes	
DPP7			Yes		
DPT			Yes		
DPYD				Yes	
DRD2	Yes		Yes		
DRD3				Yes	
DSCAML1		Yes			Yes
DSCC1		Yes			
DSEL	Yes		Yes	Yes	
DSG1					Yes
DSG3	Yes				
DSG4	Yes		Yes		
DST		Yes	Yes		Yes
DSTYK	Yes			Yes	
DTHD1		Yes			
DUOX2			Yes		
DUPD1	Yes				

Gene						
DUSP4		Yes			Yes	
DVL1				Yes		
DYDC1				Yes		
DYNC1H1	Yes					
DYNC1I2				Yes		
DYSF		Yes	Yes			
DZIP1L			Yes	Yes		
DZIP3			Yes			
E2F7			Yes			
EBF3				Yes		
EBI3				Yes		
EBNA1BP2	Yes					
ECI1			Yes			
ECM2			Yes			
ECT2L		Yes				
EDAR			Yes			
EDC4			Yes			
EDN1			Yes			
EDNRA			Yes			
EEA1		Yes				
EEF1D	Yes					
EEPD1				Yes		
EFCAB3		Yes				
EFCAB4A				Yes		
EFCAB5			Yes			
EFR3B	Yes					
EFTUD2			Yes			
EGFR		No	Yes	No	No	
EHD1			Yes			
EHMT1					Yes	
EID1	Yes					
EIF2AK4		Yes	Yes		Yes	
EIF3D			Yes			
EIF3G			Yes			
EIF4A1		Yes				
EIF4B	Yes					
EIF4E	Yes					Yes
EIF4G1		Yes	Yes			
ELAC2			Yes			
ELFN2					Yes	
ELOVL1				Yes		
ELOVL4			Yes			
ELP2	Yes					
ELP4				Yes		
ELTD1		Yes				
EMB					Yes	
EME2			Yes			
EMP3			Yes			
EMR2		Yes				
EMR4			Yes			
ENDOU		Yes				
ENGASE	Yes					
ENO1	Yes					
ENOX1	Yes					
ENPEP				Yes		
ENPP3	Yes					
ENSG00000103472				Yes		
ENSG00000121031			Yes	Yes		
ENSG00000132356			Yes			
ENSG00000156509				Yes		
ENSG00000172261			Yes			
ENSG00000174450	Yes					
ENSG00000174501			Yes	Yes		
ENSG00000182053			Yes			
ENSG00000183059			Yes			
ENSG00000187686	Yes					
ENSG00000196960	Yes					
ENSG00000204815		Yes				

ENSG00000206044			Yes			
ENSG00000213993			Yes			
ENSG00000214581			Yes			
ENSG00000215019			Yes			
ENSG00000226232		Yes				
ENTHD1	Yes	Yes				
ENTPD7				Yes		
ENTPD8			Yes			
EP300			Yes		No	No
EP400	Yes	Yes				
EPB41L1		Yes		Yes		
EPB41L4A		Yes				
EPC1				Yes		
EPG5			Yes			
EPHA2	Yes					
EPHA5		Yes	Yes	Yes		
EPHA6	Yes		Yes		Yes	Yes
EPHA8	Yes					
EPHB1				Yes		
EPHB3				Yes		
EPHB6						Yes
EPS15	Yes	Yes			Yes	
ERAP2				Yes		
ERBB2	No	No		Yes		
ERBB3	No	No				
ERBB4	No	Yes				
ERCC5			Yes			
ERCC6	Yes	Yes				
ERCC8		Yes				
ERF	Yes					
ERI2			Yes			
ESPL1		Yes	Yes			
ESR1	No	No		No	No	No
ESR2	Yes			Yes		
ESRP2				Yes		
ESRRB			Yes			
ESRRG	Yes					
ESYT1				Yes		
ESYT3			Yes			
ETAA1		Yes				
ETHE1	Yes		Yes			
ETV1	Yes			Yes		Yes
ETV4			Yes			
ETV5			Yes			
ETV6			Yes			
ETV7	Yes					
EVC		Yes				
EVC2			Yes			
EVPL		Yes	Yes			
EXD3	Yes					
EXOC1	Yes		Yes			
EXOC2				Yes		
EXOC3			Yes			
EXOC3L1				Yes		
EXOC6B		Yes				
EXOSC10	Yes					
EYS		Yes	Yes			
F12			Yes			
F13A1			Yes			
F2				Yes		
FABP6			Yes			
FADS1	Yes					
FAM100A			Yes			
FAM105A			Yes			
FAM116A	Yes					
FAM117B	Yes					
FAM123A				Yes		
FAM125B			Yes			

FAM126A	Yes					
FAM131A		Yes				
FAM134A			Yes			
FAM135A				Yes		
FAM135B				Yes		
FAM13A		Yes				
FAM13B			Yes			
FAM151A		Yes				
FAM160A1			Yes			
FAM166B				Yes		
FAM175B			Yes			
FAM178A			Yes			
FAM184A			Yes			
FAM184B			Yes			
FAM186B		Yes				
FAM188A				Yes		
FAM190A				Yes		
FAM193A				Yes		
FAM204A			Yes			
FAM20A		Yes				
FAM221B			Yes			
FAM46A			Yes			
FAM47A		Yes				
FAM48A					Yes	
FAM5B				Yes		
FAM65B		Yes			Yes	
FAM71A	Yes					
FAM71F1			Yes			
FAM81A				Yes		
FAM81B			Yes			
FAM82B	Yes					
FAM83D			Yes			
FAM83H	Yes					
FAM86C1				Yes		
FAM90A1			Yes			
FAM98A			Yes			
FANCA	Yes					
FANCB				Yes		
FANCE				Yes		
FANCM			Yes			
FAP			Yes			
FASN				Yes		
FAT2				Yes		
FAT4			Yes	Yes		
FBF1			Yes			
FBLN2	Yes					
FBLN7			Yes			
FBN2		Yes	Yes	Yes	Yes	
FBN3		No	Yes	Yes		
FBXL14			Yes			
FBXL16		Yes				
FBXL19			Yes			
FBXL5	Yes					
FBXL6	Yes					
FBXL8	Yes					
FBXO27	Yes					
FBXO3			Yes			
FBXO4			Yes			
FBXO42				Yes		
FBXO43				Yes		
FBXO47	Yes					
FBXW10				Yes		
FBXW12				Yes		
FBXW7	No	No			No	No
FCF1				Yes		
FCGR2B	Yes			Yes		
FCHO1			Yes			
FCHO2	Yes					

Gene						
FCRL4			Yes	Yes		
FCRL5	Yes				Yes	
FDXR				Yes		
FES			Yes			
FETUB			Yes		Yes	
FGD4				Yes		
FGF10			Yes			
FGF19			Yes			
FGFR1OP			Yes		Yes	
FGFR3	No	No		No	No	No
FGFR4				Yes		
FHOD3			Yes			
FIG4	Yes					
FIGNL1		Yes				
FILIP1					Yes	
FIZ1		Yes	Yes			
FKBP15			Yes			
FKBP6				Yes		
FKBP8				Yes		
FKRP				Yes		
FLG				Yes	Yes	
FLG2			Yes			
FLI1	Yes		Yes	Yes	Yes	
FLII				Yes		
FLJ13236			Yes			
FLJ14075	Yes					
FLJ16369					Yes	
FLJ45224			Yes			
FLNA					Yes	
FLNC	Yes	Yes	Yes			
FLT1	No	No		Yes		
FLT3		Yes				
FLT4	Yes	No				
FMNL2			Yes			
FMO3	Yes					
FMO5		Yes				
FMR1			Yes			
FMR1NB	Yes					
FN1	Yes		Yes		Yes	Yes
FN3KRP		Yes				
FNBP1			Yes			
FNDC1			Yes			
FOLH1B	Yes					
FOPNL					Yes	
FOS			Yes			
FOSL1		Yes				
FOXD4		Yes				
FOXL2		No			No	
FPGS	Yes					
FRAS1			Yes	Yes		
FREM1		Yes				
FREM3		Yes	Yes			
FRMD3	Yes					
FRMD8			Yes			
FRMPD2		Yes		Yes		
FRMPD3		Yes				
FRS3		Yes				
FRY				Yes	Yes	
FRYL			Yes	Yes		
FSCB	Yes					
FSIP2				Yes		
FTSJ3				Yes		
FTSJD2				Yes		
FUK				Yes		
FUS		Yes				
FUSIP1	Yes					
FUT10				Yes		
FUT9	Yes					

FXR1	Yes					
FYB			Yes			
FYN		Yes				
FZD1			Yes			
FZD10			Yes			
FZD6			Yes			
FZD7			Yes			
G2E3	Yes					
G6PC2		Yes				
GAA				Yes		
GAB3			Yes			
GAB4					Yes	
GABBR1	Yes			Yes		
GABRA6	Yes		Yes			
GABRB1				Yes	Yes	
GABRB2				Yes		
GABRG1			Yes			
GABRP	Yes					
GABRR3	Yes					
GADL1				Yes		
GAL3ST4	Yes					
GALNT16	Yes					
GALNT6		Yes				
GALNTL1	Yes					
GAPDH			Yes			
GAPDHS	Yes					
GAPVD1					Yes	
GARNL3	Yes					
GATA2			Yes			
GATA3		Yes			No	No
GBA2				Yes		
GBF1			Yes			
GCLC			Yes			
GCLM			Yes			
GCNT7			Yes			
GDAP1L1	Yes					
GDF3			Yes			
GDPD3			Yes		Yes	
GDPD5	Yes					
GEM	Yes					
GEN1	Yes					
GFPT2			Yes			
GGA2					Yes	
GGCX			Yes			
GGT5		Yes				
GHR	Yes		Yes			
GHRHR	Yes					
GHSR				Yes		
GIGYF1		Yes				
GIMAP4	Yes					
GIN1	Yes					
GIT1			Yes			
GJA10	Yes					
GJC2				Yes		
GJC3			Yes			
GJD2	Yes					
GKAP1					Yes	
GLB1L3			Yes	Yes		
GLI1		Yes			Yes	
GLIPR1L1			Yes			
GLOD5			Yes	Yes		
GLS2	Yes					
GLTPD1			Yes			
GMEB1			Yes			
GMFB				Yes		
GNA11	Yes					
GNAI3	Yes					
GNAO1			Yes			

Gene						
GNG8	Yes					
GNL1			Yes			
GNS			Yes			
GOLGA4				Yes		
GOLGA8H				Yes		
GOLGA8J					Yes	
GOLGB1	Yes		Yes			
GOLPH3L		Yes				
GP2				Yes		
GP5			Yes		Yes	
GPAA1	Yes					
GPAM				Yes		
GPATCH2L			Yes			
GPATCH3			Yes			
GPC6		Yes	Yes			
GPHN			Yes		Yes	
GPNMB			Yes			
GPR111					Yes	
GPR112		Yes	Yes	Yes	Yes	
GPR113	Yes					
GPR12	Yes					
GPR125		Yes				
GPR128				Yes		
GPR132			Yes			
GPR133	Yes	Yes				
GPR144		Yes				
GPR172A		Yes				
GPR20				Yes		
GPR26			Yes			
GPR3			Yes			
GPR56		Yes	Yes			
GPR64			Yes			
GPR68		Yes				
GPR77			Yes			
GPR85	Yes	Yes				
GPR89A			Yes			
GPR98		Yes	Yes	Yes		
GPRASP1				Yes		
GPRC5B		Yes				
GPRIN2			Yes			
GPRIN3			Yes			
GPT		Yes				
GPX5			Yes			
GRAMD2					Yes	
GRAMD4			Yes			
GRIA1	Yes					
GRIA3	Yes	Yes	Yes	Yes	Yes	Yes
GRID1					Yes	
GRID2			Yes			
GRIK1		Yes			Yes	
GRIK2			Yes			
GRIK4				Yes		
GRIN2A			Yes			
GRIN3A			Yes			
GRM1			Yes			
GRM4			Yes			
GRM7			Yes			
GRM8		Yes		Yes		
GSDMB					Yes	
GSG1	Yes					
GSK3A	Yes					
GSTA4			Yes			
GSTO1					Yes	
GSTO2				Yes		
GTF2H3	Yes					
GTF2IRD1			Yes			
GTF3C1				Yes		
GTPBP3			Yes			

Gene						
GTSE1			Yes			
GUCY1B3			Yes			
GUCY2C		Yes	Yes	Yes	Yes	Yes
GUF1			Yes			
GULP1	Yes					
GYPB				Yes		
HABP4	Yes					
HACE1	Yes					
HAPLN1				Yes		
HAPLN3					Yes	
HAUS4	Yes	Yes				
HAVCR1					Yes	
HCN2			Yes			
HCRTR2				Yes		
HDAC6				Yes		
HDAC7	Yes		Yes			
HDAC9		Yes				
HEATR1				Yes		
HEATR5B			Yes			
HEATR7B2	Yes		Yes			
HECTD3	Yes					
HEG1		Yes				
HELLS				Yes		
HELQ	Yes					
HEPH						Yes
HERC1			Yes			
HERC2	Yes	Yes	Yes		Yes	
HERC4	Yes	Yes				
HERC5			Yes			
HERC6			Yes	Yes		
HES1	Yes					
HES5		Yes				
HEXDC			Yes			
HFM1			Yes			
HHLA1		Yes				
HIF1A	Yes		Yes	Yes	Yes	Yes
HIP1R					Yes	
HIST1H1C		Yes				
HIST1H2AM		Yes				
HIST1H2BB	Yes					
HIST1H2BC					Yes	
HIST1H2BH		Yes	Yes			
HIST1H2BL		Yes				
HIST1H2BO			Yes			
HIST1H3B	Yes					
HIST1H4D				Yes		
HIVEP1			Yes			
HK1	Yes					
HK2					Yes	
HKR1			Yes			
HLA-A			Yes			
HLA-DPA1			Yes			
HLA-G			Yes			
HLTF		Yes				
HM13			Yes			
HMCN1		Yes		Yes		
HMGCR			Yes			
HMGXB3					Yes	
HNF1A	Yes		Yes	Yes		Yes
HNF1B			Yes			
HNRNPR			Yes			
HNRPR			Yes			
HOXA13	Yes					
HOXB1		Yes		Yes		
HP		Yes				
HPCAL1	Yes					
HPX	Yes					
HRAS	No	No	No	No	No	No

HRASLS			Yes			
HRCT1			Yes			
HRH3				Yes		
HS2ST1			Yes			
HS3ST3B1		Yes				
HSD17B2			Yes			
HSD3B7				Yes		
HSP90B1	Yes				Yes	Yes
HSPA12A	Yes					
HSPA4L	Yes					
HSPD1	Yes					
HSPG2	Yes	Yes		Yes		
HTT	Yes					
HVCN1		Yes				
HYAL1				Yes		
HYDIN				Yes	Yes	
HYOU1	Yes					
IBTK			Yes			
ICAM1	Yes					
ICAM2					Yes	
ICAM5			Yes			
IDE	Yes					
IDH1		No				No
IDH3G				Yes		
IDI2				Yes		
IDO2	Yes					
IER5			Yes			
IFI16			Yes	Yes		
IFI44			Yes			
IFNK		Yes				
IFT122				Yes		
IFT140			Yes	Yes		
IFT88			Yes			
IGF1R	No	No				
IGF2BP2				Yes		
IGF2R				Yes		
IGFALS			Yes			
IGFBP6	Yes					
IGFN1			Yes			
IGSF10			Yes			
IGSF6			Yes			
IKZF4	Yes					
IKZF5			Yes			
IL16			Yes			
IL17RA				Yes		
IL17RC	Yes					
IL17RE			Yes			
IL17REL			Yes			
IL5RA					Yes	
IL6R		Yes				
ILDR2			Yes			
ILK	Yes	Yes	Yes	Yes	Yes	Yes
ILVBL		Yes	Yes			
IMMT			Yes			
IMPG1	Yes					
INCENP			Yes			
INHBA			Yes			
INO80D			Yes			
INPP4B					Yes	
INPP5B		Yes				
INPP5E			Yes			
INPP5J				Yes		
INPPL1			Yes			
INTS1	Yes					
INTS2	Yes					
INTS3	Yes					
INTS5			Yes			
IPO13	Yes					

IPPK				Yes		
IQCC					Yes	
IQCH				Yes		
IQCK					Yes	
IQGAP2		Yes				
IQSEC3			Yes			
IQUB			Yes			
IREB2			Yes			
IRF1		Yes				
IRG1	Yes					
IRS1	Yes	Yes	Yes		Yes	
IRX1			Yes			
IRX2	Yes					
IRX6	Yes					
ISL1		Yes				
ISM1		Yes				
ISYNA1	Yes					
ITFG3			Yes			
ITGA10	Yes					
ITGA11	Yes					
ITGA2			Yes		Yes	
ITGA4				Yes		
ITGA6			Yes			
ITGAD	Yes					
ITGAM		Yes				
ITGAV	Yes			Yes		
ITGAX		Yes			Yes	
ITGB1BP3		Yes				
ITGB4				Yes		
ITIH1	Yes					
ITIH5				Yes		
ITIH5L			Yes			
ITPR1	Yes					
ITPR2	Yes	Yes	Yes	Yes	Yes	Yes
ITPR3	Yes			Yes	Yes	
IVNS1ABP					Yes	
IZUMO3	Yes					
JAKMIP2	Yes	Yes	Yes			
JKAMP	Yes					
JMJD6			Yes			
JMJD7-PLA2G4B	Yes					
JMJD8	Yes					
JPH1			Yes			
KALRN	Yes					
KANK2			Yes			
KANK4			Yes			
KARS			Yes			
KAT2A	Yes					
KAT6B				Yes		
KATNB1				Yes		
KBTBD2			Yes			
KBTBD5			Yes		Yes	
KCNA6				Yes		
KCNA7				Yes		
KCNG1	Yes					
KCNH5				Yes		
KCNJ14		Yes				
KCNJ8			Yes			
KCNK13		Yes		Yes		
KCNK6			Yes			
KCNK9					Yes	
KCNMA1	Yes				Yes	
KCNMB3				Yes		
KCNN2		Yes				
KCNN3			Yes			
KCNQ3			Yes			
KCNQ5	Yes					
KCNS2			Yes			

KCNT1			Yes			
KCNU1			Yes			
KCNV1				Yes		
KCTD1				Yes		
KDELR2	Yes					
KDM2A	Yes					
KDM2B			Yes			
KDM3A		Yes				
KDM3B			Yes	Yes		
KDM4E				Yes		
KDM5A			Yes			
KDM5C			Yes		Yes	
KDM6A	No	No	No		No	
KDR	No	No	Yes			
KHDRBS2			Yes			
KHSRP		Yes				
KIAA0100	Yes					
KIAA0182			Yes			
KIAA0195	Yes			Yes		
KIAA0247				Yes		
KIAA0319					Yes	
KIAA0408				Yes		
KIAA0586		Yes			Yes	
KIAA0649		Yes				
KIAA0701	Yes		Yes	Yes		
KIAA0746	Yes	Yes				
KIAA0802				Yes		
KIAA0913	Yes					
KIAA0922			Yes			
KIAA0947	Yes					
KIAA0953	Yes					
KIAA1009			Yes	Yes		
KIAA1109			Yes	Yes		
KIAA1161	Yes					
KIAA1199		Yes				
KIAA1211					Yes	
KIAA1217				Yes		
KIAA1239			Yes			
KIAA1244	Yes			Yes		
KIAA1377		Yes				
KIAA1404				Yes		
KIAA1429	Yes					
KIAA1462		Yes				
KIAA1467			Yes			
KIAA1468	Yes					
KIAA1529	Yes					
KIAA1530	Yes					
KIAA1671		Yes				
KIAA1731				Yes		
KIAA1755			Yes			
KIAA1804	No	No				
KIAA1841			Yes			
KIAA1967		Yes	Yes		Yes	
KIAA2022				Yes		
KIF13A	Yes					
KIF19				Yes		
KIF1C			Yes			
KIF20B			Yes	Yes		
KIF21B	Yes					
KIF22			Yes			
KIF24	Yes					
KIF26A	Yes					
KIF26B	Yes					
KIF27	Yes					
KIF2B		Yes				
KIF3B			Yes	Yes		
KIF5B	Yes					
KIF5C	Yes					

KIF6		Yes	Yes			
KIFC1			Yes			
KIR2DL4	Yes		Yes			
KIR3DL1	Yes					
KIRREL					Yes	
KLHDC2				Yes		
KLHDC7A				Yes		
KLHL1	Yes			Yes		
KLHL10					Yes	
KLHL18				Yes		
KLHL24	Yes					
KLHL25			Yes			
KLHL26	Yes					
KLHL29	Yes					
KLK6	Yes		Yes			
KLK7			Yes		Yes	Yes
KMO				Yes		
KMT2C		Yes			Yes	
KMT2D	Yes					
KNG1	Yes					
KNTC1			Yes			
KPNA2			Yes			
KPNB1					Yes	
KPRP	Yes					
KRAS	No	No	No	No	No	No
KRT19			Yes			
KRT2		Yes	Yes			
KRT28	Yes					
KRT40	Yes					
KRT5			Yes			
KRT6B			Yes			
KRT6C				Yes		
KRT7			Yes			
KRT74	Yes					
KRT80					Yes	
KRTAP10-3	Yes					
KRTAP10-8	Yes					
KRTAP1-5					Yes	
KRTAP26-1				Yes		
KRTAP5-1				Yes		
KRTAP5-6	Yes					
KSR1			Yes			
KTN1			Yes			
KY	Yes					
L1CAM	Yes		Yes			
L3MBTL1			Yes			
L3MBTL2				Yes		
LAD1			Yes			
LAIR1		Yes				
LAMA1	Yes		Yes	Yes		
LAMA2			Yes			
LAMA3		Yes	Yes			
LAMA5			Yes			
LAMB1	Yes					
LAMB3			Yes			
LAMB4				Yes		
LAMC2				Yes		
LAMC3		Yes				
LANCL1			Yes			
LAPTM4B	Yes					
LARP1B				Yes		
LARP4B					Yes	
LARP6				Yes		
LASS2	Yes					
LASS5			Yes			
LATS1	Yes	Yes				
LBR				Yes		
LCE1B				Yes		

LCE2A				Yes		
LCE2B					Yes	
LCE2C	Yes					
LCE5A					Yes	
LCK				Yes		
LCMT1			Yes	Yes		
LCP2			Yes			
LCT	Yes		Yes			
LCTL					Yes	
LDHA	Yes					
LDHAL6A	Yes					
LDHAL6B				Yes		
LDHD			Yes			
LDLR	Yes					
LELP1				Yes		
LETM2			Yes			
LGALS3BP		Yes				
LGALS8			Yes			
LGR4				Yes		
LIAS		Yes				
LIG1	Yes		Yes			
LILRA2			Yes			
LILRA5			Yes			
LILRB2	Yes		Yes			
LILRB5			Yes			
LIMA1				Yes		
LIMK2		Yes	Yes			
LIPC			Yes			
LIPF		Yes				
LIX1					Yes	
LMBR1L			Yes			
LMLN	Yes	Yes				
LMOD3	Yes					
LMX1A	Yes					
LNPEP	Yes			Yes		
LNX2			Yes			
LOC283849				Yes		
LOC440345				Yes		
LOC51059				Yes		
LOC541473				Yes		
LOC652153				Yes	Yes	
LOC652737					Yes	
LOC728194	Yes					
LOC731173			Yes			
LOC91807			Yes			
LONP1				Yes		
LONRF2	Yes					
LOXHD1				Yes		
LOXL4		Yes				
LPA	Yes		Yes		Yes	
LPCAT1					Yes	
LPCAT4				Yes		
LPGAT1				Yes		
LPHN1	Yes					
LPHN3				Yes		
LRBA				Yes		
LRCH3				Yes		
LRFN2				Yes		
LRGUK			Yes			
LRIG2	Yes					
LRIG3			Yes			
LRP1		Yes				
LRP11			Yes			
LRP12			Yes			
LRP1B				Yes		Yes
LRP2	Yes			Yes		
LRP5	Yes					
LRR1	Yes					

Gene						
LRRC14		Yes				
LRRC16A	Yes				Yes	
LRRC16B	Yes		Yes			
LRRC2			Yes			
LRRC33		Yes				
LRRC36				Yes		
LRRC37A3		Yes		Yes		
LRRC37B				Yes		
LRRC4	Yes					
LRRC41	Yes					
LRRC43			Yes			
LRRC48	Yes					
LRRC4C			Yes			
LRRC8A		Yes	Yes			
LRRIQ1				Yes	Yes	
LRRK2		Yes	Yes	Yes		
LRRTM1			Yes			
LRRTM3				Yes		
LRRTM4		Yes				
LSP1	Yes					
LTA			Yes			
LTB4R2			Yes			
LTBP2			Yes			
LTBP4	Yes					
LTN1				Yes		
LUZP1			Yes			
LY6G6D			Yes			
LY6G6E	Yes					
LY75	Yes					
LYN		Yes				
LYPD2			Yes			
LYRM2	Yes					
LYST	Yes	Yes		Yes		
LZTR1			Yes			
LZTS2	Yes					
MAATS1			Yes			
MACF1			Yes	Yes		
MAD1L1				Yes		
MAD2L1	Yes					
MAD2L2				Yes		
MADD				Yes		
MAEL					Yes	
MAF	Yes		Yes	Yes		
MAGEB3			Yes			
MAGEC1					No	
MAGED2	Yes					
MAML2		Yes	Yes			
MAML3	Yes	Yes	Yes		Yes	
MAMSTR			Yes			
MAN1A1				Yes		
MAN2A2	Yes					
MAN2B2			Yes			
MAP1B			Yes	Yes		
MAP1S					Yes	
MAP2		Yes				
MAP2K1	No	No				
MAP2K2	No	No				
MAP2K3			Yes			
MAP2K4	No	No	No	No	No	No
MAP2K5			Yes			
MAP2K6	Yes		Yes			
MAP3K1	Yes	Yes	Yes	Yes	Yes	Yes
MAP3K10				Yes		
MAP3K11	Yes					
MAP3K13		Yes	Yes			
MAP3K14	Yes	Yes	Yes	Yes	Yes	Yes
MAP3K15			Yes			
MAP3K3			Yes			

MAP3K5		Yes				
MAP3K6					Yes	
MAP3K7				Yes		
MAP3K8	Yes					
MAP4K3	Yes	Yes	Yes			
MAP4K4	No	No				
MAP6				Yes		
MAP7			Yes			
MAPK1	Yes					
MAPK10	Yes					
MAPK12	Yes					
MAPK13	Yes					
MAPK4				Yes		
MAPK8IP1				Yes		
MAPK8IP3				Yes		
MAPKAPK3						Yes
MAPKAPK5			Yes			
MARCH2			Yes			
MARCKS	Yes					
MARK4		Yes				
MAST1			Yes			
MAST2				Yes		
MAST4			Yes			
MAT1A				Yes		
MATN2		Yes				
MATN4			Yes			
MATR3			Yes			
MAZ	Yes					
MB21D2	Yes					
MBD1			Yes			
MBTPS1	Yes					
MCF2		Yes				
MCF2L			Yes			
MCM3AP			Yes			
MCM3APAS					Yes	
MCM5			Yes			
MDH1B	Yes					
MDN1		Yes	Yes	Yes		
MECOM		Yes				
MED13L			Yes			
MED16	Yes					
MED21				Yes		
MED23					Yes	
MEFV			Yes			
MEGF10			Yes	Yes		
MEGF6	Yes					
MEGF8	Yes					
MEI1				Yes		
MEIG1	Yes					
MEIR5			Yes			
MERTK	Yes		Yes			
MET	No	No			No	No
METTL13	Yes			Yes		
METTL16	Yes					
METTL20	Yes					
METTL22				Yes		
METTL2A	Yes					
METTL3	Yes					
MFGE8					Yes	
MFI2		Yes		Yes		
MFN1			Yes		Yes	
MFSD1				Yes		
MFSD6	Yes					
MGA	Yes			Yes		
MGAM	Yes		Yes			
MGRN1				Yes		
MGST3				Yes		
MICAL1			Yes			

MICALL2	Yes			Yes	Yes	
MID1				Yes		
MIIP	Yes					
MINK1	Yes					
MINPP1			Yes			
MIOS	Yes					
MIPEP		Yes				
MIS18BP1				Yes		
MKI67			Yes			
MKI67IP		Yes				
MKKS				Yes		
MKL1			Yes			
MKLN1	Yes			Yes		
MKNK1				Yes		
MKX	Yes					
MLH1	Yes					
MLKL				Yes		
MLL	Yes					
MLL2	Yes					
MLL3		Yes		Yes	Yes	
MLL4	Yes					
MLLT10			Yes			
MLLT4		Yes			Yes	
MLLT6	Yes			Yes		
MLXIP				Yes		
MMP1		Yes				
MMP20				Yes		
MMS19			Yes			
MMS22L	Yes			Yes		
MNDA				Yes		
MON1B				Yes		
MON2			Yes			
MORN1	Yes			Yes		
MPEG1	Yes					
MPND	Yes					
MPP2			Yes			
MPP4			Yes			
MPRIP	Yes			Yes		
MR1	Yes					
MRAP				Yes		
MRC2			Yes			
MRGPRX2			Yes			
MRPL4			Yes			
MRPS16		Yes		Yes		
MRPS34			Yes			
MRVI1			Yes			
MS4A1		Yes				
MS4A4E				Yes		
MSH3					Yes	
MSH6		No	No	No	No	No
MSL2				Yes		
MSMB				Yes		
MSRA			Yes			
MST1			Yes			
MST1R	No	No				
MT1E			Yes			
MTA1	Yes					
MTA3					Yes	
MTCP1					Yes	
MTCP1NB		Yes			Yes	
MTERFD2				Yes		
MTF1			Yes			
MTHFD2L		Yes			Yes	
MTIF2			Yes			
MTMR3				Yes		
MTMR8				Yes		
MTOR	No	No	Yes			
MTX1				Yes		

Gene						
MUC12	Yes					
MUC16	Yes		Yes	Yes		
MUC17				Yes	Yes	
MUC20	Yes					
MUC4			Yes	Yes		
MUC5AC					Yes	
MUC5B					Yes	
MUC6				Yes	Yes	
MUT		Yes	Yes		Yes	
MYBBP1A		Yes				
MYBL2	Yes					
MYBPC1		Yes				
MYBPC2	Yes					
MYCBP				Yes		
MYCBPAP			Yes			
MYH1			Yes			
MYH11				Yes		
MYH13	Yes		Yes			
MYH14				Yes		
MYH15	Yes					
MYH2		Yes	Yes	Yes		
MYH3		Yes				
MYH4	Yes	Yes				
MYH6			Yes			
MYH8	Yes					
MYH9		Yes	Yes			
MYLK		Yes	Yes			Yes
MYLK3			Yes			
MYLK4				Yes		
MYLPF	Yes					
MYO10					Yes	
MYO15A	Yes					
MYO18A	Yes					
MYO18B	Yes		Yes			
MYO1A				Yes		
MYO1C					Yes	
MYO1D	Yes					
MYO1G			Yes			
MYO3B		Yes				
MYO5A				Yes		
MYO5C				Yes		
MYO7B			Yes			
MYO9A			Yes			
MYOC					Yes	
MYOCD				Yes		
MYOF			Yes		Yes	
MYOM1	Yes					
MYPN				Yes		
MYST3			Yes			
MYST4				Yes		
MYT1L				Yes		
MZF1					No	
N4BP2	Yes		Yes			
N6AMT1				Yes		
NAALADL2				Yes		
NAG6		Yes				
NALCN	Yes					
NAP1L4				Yes		
NAPA				Yes		
NAT10				Yes		
NAV3	Yes			Yes		
NBAS	Yes					
NBEA		Yes				
NBEAL1	Yes					
NBEAL2				Yes		
NBN		Yes				
NCAM1	Yes	Yes	Yes	Yes	Yes	Yes
NCAM2				Yes		

NCAPD2	Yes		Yes			
NCAPD3	Yes		Yes			
NCAPG				Yes		
NCBP1		Yes				
NCBP2L		Yes				
NCEH1		Yes				
NCKAP1L			Yes			
NCKAP5			Yes			
NCL		Yes				
NCOA3					Yes	
NCOA5				Yes		
NCOR1	Yes					
NCR1			Yes			
NCR2		Yes				
NDST2		Yes				
NDUFV1			Yes			
NEB	Yes			Yes		
NEDD1	Yes					
NEDD4		Yes	Yes			
NEIL3			Yes			
NEK11			Yes			
NEK3			Yes			
NEK6	No	No				
NEK8			Yes			
NELF			Yes			
NELL1		Yes				
NELL2			Yes			
NES		Yes	Yes			
NEU2	Yes					
NFASC				Yes		
NFATC2				Yes		
NFATC4				Yes		
NFE2					Yes	
NFE2L2		No				
NFIA	Yes		Yes			Yes
NFIX	Yes		Yes			
NFKB1	Yes					
NFKB2	Yes					
NFS1					Yes	
NFX1			Yes	Yes		
NGEF				Yes		
NHS					Yes	
NIPAL2			Yes			
NIPBL	Yes	Yes				
NIPSNAP3B		Yes				
NIT2	Yes					
NKD2	Yes					
NLGN4X			Yes			
NLRC3	Yes			Yes		
NLRC5			Yes			
NLRP1		Yes				
NLRP12	Yes					
NLRP13				Yes		
NLRP7			Yes	Yes		
NLRP8	Yes					
NLRP9			Yes		Yes	
NLRX1	Yes		Yes			
NMNAT2				Yes		
NOL10	Yes					
NOL11				Yes		
NOL6	Yes		Yes			
NOL8					Yes	
NOM1	Yes					
NOMO1	Yes					
NOP14					Yes	
NOP16			Yes	Yes		
NOP58			Yes			
NOS1				Yes		

NOS2				Yes		
NOTCH1			Yes			
NOTCH4	Yes					
NOX3					Yes	
NPAS4			Yes			
NPAT		Yes				
NPEPPS		Yes			Yes	
NPHP4	Yes					
NPL	Yes					
NPM1	Yes			Yes	Yes	
NPR1	Yes				Yes	
NPSR1	Yes					
NPVF					Yes	
NR0B1			Yes			
NR1H2	Yes		Yes	Yes	Yes	Yes
NR3C1				Yes	Yes	
NR5A1		Yes				
NRAS	No	No	No	No	No	No
NRCAM		Yes				
NRD1	Yes					
NRP2			Yes			
NRXN2	Yes	Yes				
NSD1	Yes					
NSMAF			Yes			
NSMCE4A				Yes		
NT5C1A	Yes	Yes				
NTSR1			Yes			
NUAK1			Yes			
NUCKS1	Yes					
NUDT3		Yes				
NUMA1	Yes					
NUMB	Yes					
NUP133			Yes			
NUP155	Yes					
NUP188	Yes		Yes			
NUP210			Yes			
NUP214					Yes	
NUP54	Yes	Yes				
NUP98				Yes		
NUPR1				Yes		
NXF1				Yes		
NXF3				Yes		
NYNRIN				Yes		
OAS3			Yes			
OBFC2B				Yes		
OBSCN	Yes		Yes	Yes		
OBSL1	Yes					
ODF2L			Yes			
ODZ1		Yes				
ODZ2	Yes					
ODZ4				Yes		
OFCC1	Yes					
OIP5		Yes				
OLR1				Yes		
ONECUT1		Yes				
OPCML	Yes			Yes		
OPHN1				Yes		
OPTN	Yes					
OR10AD1			Yes			
OR10G2	Yes					
OR10G6	Yes				Yes	
OR10G9					Yes	
OR10H5			Yes	Yes		
OR10S1				Yes		
OR10X1	Yes					
OR13A1	Yes					
OR13C5					Yes	
OR13C8			Yes			

OR14J1				Yes		
OR1E1	Yes					
OR1F1				Yes		
OR2B2			Yes			
OR2B3		Yes				
OR2B6				Yes		
OR2D3	Yes	Yes				
OR2G2					Yes	
OR2G6	Yes					
OR2L2	Yes					
OR2S2			Yes			
OR2T11	Yes					
OR2T12			Yes			
OR2T33			Yes			
OR2W3					Yes	
OR2Z1		Yes				
OR4A16			Yes			
OR4C15	Yes					
OR4C5_HUMAN		Yes				
OR4F6	Yes					
OR4K14			Yes			
OR4N4			Yes			
OR51L1				Yes		
OR51M1				Yes		
OR51Q1		Yes	Yes			
OR52I1				Yes		
OR52I2				Yes		
OR52N1					Yes	
OR5AK2			Yes			
OR5AN1			Yes			
OR5H14			Yes			
OR5I1			Yes			
OR5K2				Yes		
OR5M9			Yes			
OR7A2P	Yes	Yes				
OR7C2	Yes					
OR8D2					Yes	
OR8J1				Yes		
OR9G1			Yes			
ORC1				Yes		
ORC3			Yes			
OS9		Yes				
OSBPL11				Yes		
OSBPL1A	Yes		Yes	Yes		
OSBPL3				Yes		
OSBPL6				Yes		
OSCP1			Yes			
OSGEPL1		Yes				
OSR2				Yes		
OSTbeta			Yes			
OSTM1	Yes					
OTC	Yes					
OTOF			Yes			
OTOGL				Yes		
OTUD6B	Yes					
OTX2	Yes		Yes			
OXR1	Yes					
P2RX5			Yes			
P2RY13				Yes		
PAAF1			Yes			
PABPC1	Yes					
PABPC3	Yes					
PACS2	Yes					
PACSIN3					Yes	
PADI2				Yes		
PADI3	Yes					
PAGE3			Yes			
PAIP1				Yes		

PAK1				Yes		
PAK4		No				
PAK6		Yes				
PALB2		Yes				
PALM2-AKAP2	Yes					
PAN2			Yes			
PAPLN	Yes				Yes	
PARD3				Yes		
PARD3B		Yes				
PARN	Yes					
PARP1	Yes					
PARP12			Yes			
PARP15			Yes			
PARP9			Yes			
PARVB	Yes					
PASD1				Yes		
PASK		Yes				
PAX1			Yes			
PAX3	Yes		Yes	Yes	Yes	Yes
PAX4			Yes			
PAX5		Yes		Yes		
PAX6				Yes		
PAX9			Yes			
PBX2	Yes		Yes			
PBX3			Yes			
PBXIP1		Yes				
PCBD1			Yes			
PCBP2	Yes					
PCBP4		Yes				
PCDH1			Yes			
PCDH15	Yes		Yes			
PCDH17				Yes		
PCDH19			Yes			
PCDH9				Yes		
PCDHA1				Yes		
PCDHA12	Yes		Yes			
PCDHA13	Yes					
PCDHA8	Yes					
PCDHA9	Yes					
PCDHB13	Yes					
PCDHB16	Yes					
PCDHB3				Yes		
PCDHB5	Yes					
PCDHB6				Yes		
PCDHB7			Yes			
PCDHGA10					Yes	
PCDHGA12		Yes				
PCDHGA5					Yes	
PCDHGA7	Yes					
PCDHGA9		Yes				
PCDHGB1	Yes					
PCDHGC5			Yes			
PCLO			Yes			
PCNT				Yes	Yes	
PCNX		Yes	Yes	Yes		
PCYOX1			Yes			
PDCD6	Yes					
PDE10A			Yes			
PDE1C				Yes		
PDE4DIP	Yes		Yes	Yes	Yes	
PDE8A			Yes			
PDE8B		Yes				
PDGFA			Yes			
PDGFRA	Yes				Yes	
PDGFRL	Yes					
PDIA3				Yes		
PDK1	No	No				
PDLIM5				Yes		

PDP2			Yes			
PDZD2			Yes			
PDZD4			Yes			
PDZD7	Yes		Yes			
PDZD8			Yes			
PDZRN4			Yes			
PEG3	Yes	Yes	Yes			
PELI3	Yes					
PER1	Yes					
PEX14		Yes				
PEX5L	Yes		Yes			
PFAS			Yes			
PFKP			Yes			
PGAP3	Yes					
PGBD4				Yes		
PGBD5		Yes				
PGC			Yes			
PGGT1B			Yes			
PGLS	Yes					
PGLYRP2	Yes					
PGM2L1			Yes			
PHACTR1					Yes	
PHF1	Yes					
PHF11	Yes					
PHF16	Yes					
PHF19					Yes	
PHF20				Yes		
PHF20L1	Yes					
PHF3				Yes		
PHKB				Yes		
PHKG1			Yes			
PHLDA3				Yes		
PHLDB1			Yes			
PHLPP	Yes					
PHLPP1	Yes					
PHOX2A			Yes			
PI16					Yes	
PIEZO1			Yes			
PIEZO2		Yes				
PIF1	Yes			Yes		
PIGF		Yes				
PIGG		Yes				
PIGR		Yes				
PIGZ				Yes		
PIK3AP1		Yes				
PIK3C2A	No	No			Yes	
PIK3C2G			Yes	Yes	Yes	Yes
PIK3CA	Yes	Yes	Yes	Yes	Yes	No
PIK3CD				Yes		
PIK3CG			Yes	Yes		
PIK3R1						Yes
PILRB			Yes			
PION			Yes			
PITPNM1	Yes					
PITRM1			Yes			
PIWIL4		Yes				
PJA2				Yes		
PKD1	Yes				Yes	
PKD1L1					Yes	
PKD1L2			Yes			
PKDCC				Yes		
PKHD1			Yes	Yes		
PKIA			Yes			
PKMYT1	Yes					Yes
PKN1	Yes	Yes				
PLA2G3					Yes	
PLAG1	Yes					
PLAU				Yes		

Gene						
PLB1	Yes					
PLCB1	Yes					
PLCB2			Yes			
PLCE1			Yes			
PLCG2					Yes	
PLCL1	Yes					
PLCZ1	Yes					
PLD1	Yes					
PLD2		Yes				
PLEC		Yes			Yes	
PLEKHA5	Yes					
PLEKHA6	Yes					
PLEKHG1			Yes			
PLIN4	Yes					
PLIN5				Yes		
PLK1			Yes			
PLK3	Yes					
PLUNC	Yes					
PLXDC2			Yes			
PLXNA2			Yes			
PLXNB2		Yes	Yes			
PLXND1		Yes				
PMFBP1	Yes					
PML	Yes			Yes		
PMS1	Yes					Yes
PNPLA6		Yes				
PNPLA7			Yes			
PNPT1				Yes		
PNRC1	Yes					
POC1A				Yes		
POC1B			Yes			
PODN			Yes			
POLA1			Yes			
POLB			Yes			
POLD2			Yes			
POLG				Yes		
POLK	Yes					
POLM	Yes					
POLQ	Yes			Yes		
POLR1B				Yes		
POLR2B				Yes		
POLR2D			Yes			
POLR2E				Yes		
POLR2G		Yes				
POLR3E			Yes			
POLR3K			Yes			
POM121			Yes			
POM121C				Yes		
POP1	Yes					
POSTN					Yes	
POU5F1				Yes		
PPAPDC1A		Yes				
PPAPDC3			Yes			
PPARA	Yes					
PPARG			Yes			
PPEF2			Yes			
PPFIA2			Yes			
PPFIA3			Yes			
PPFIA4			Yes			
PPFIBP1	Yes		Yes			
PPIA	Yes					
PPL		Yes	Yes			
PPM1E	Yes					
PPM1G			Yes			
PPM1H			Yes			
PPP1R12B	Yes	Yes				
PPP1R14D		Yes				
PPP1R16B					Yes	

PPP1R2	Yes					
PPP1R3A			Yes	Yes		
PPP1R3D				Yes		
PPP1R7				Yes		
PPP2CB	Yes					
PPP2R1A		No	No	No	No	
PPP2R2B				Yes		
PPP2R3A	Yes					
PPP3CA				Yes		
PPP4R4		Yes		Yes		
PPP6R1			Yes	Yes		
PRAMEF10				Yes		
PRB3			Yes			
PRCC				Yes		
PRDM15	Yes	Yes	Yes			
PRDM5	Yes					
PRDX2					Yes	
PREX2			Yes			
PRG4	Yes	No				
PRICKLE2	Yes					
PRKAA1			Yes			
PRKAG2			Yes			
PRKCA					Yes	
PRKCH						Yes
PRKCSH			Yes			
PRKD3					Yes	
PRKDC	Yes	Yes	Yes		Yes	
PRKG1				Yes		
PRLHR			Yes			
PRMT10	Yes	Yes		Yes		
PROKR2	Yes					
PROM2			Yes	Yes		
PROS1		Yes				
PROX1		Yes				
PRPF3			Yes			
PRPF39				Yes		
PRPF40A			Yes			
PRPF40B	Yes		Yes			
PRPF4B			Yes			
PRR14L			Yes			
PRR22				Yes		
PRR3		Yes				
PRR5L	Yes					
PRRC2B	Yes	Yes				
PRSS21			Yes			
PRSS23		Yes				
PRSS3	Yes					
PRSS58		Yes				
PRUNE2	Yes					
PRX	Yes					
PSD					Yes	
PSD3				Yes		
PSG6				Yes		
PSG9	Yes					
PSKH1	Yes					
PSKH2	Yes		Yes			
PSMA8	Yes					
PSMB11					Yes	
PSMC4			Yes			
PSMC5		Yes				
PSMD1	Yes					
PSMD11			Yes			
PSMD14			Yes			
PSME4				Yes		
PTCH1			Yes			
PTCHD3				Yes		
PTDSS1			Yes			
PTDSS2	Yes			Yes		

PTEN	No	No	No	Yes	No	Yes
PTGS1			Yes			
PTH2R	Yes					
PTK2			Yes			
PTK2B	Yes					
PTK6			Yes			
PTPN21	Yes					
PTPN22			Yes			
PTPN23		Yes				
PTPN3		Yes				
PTPN4			Yes			
PTPRA			Yes	Yes		
PTPRB			Yes			
PTPRD		Yes				
PTPRE			Yes			
PTPRG				Yes		
PTPRK	Yes		Yes			
PTPRM		Yes				
PTPRN2			Yes			
PTPRO				Yes		
PTPRR			Yes			
PTPRS			Yes	Yes		
PUS10		Yes				
PUS3				Yes		
PVRIG	Yes					
PVRL1		Yes				
PWP1			Yes			
PXDNL			Yes	Yes		
PXN	Yes					
PYCR1	Yes					
PYGL				Yes		
PZP		Yes				
Q5I0X0_HUMAN			Yes			
Q5JXA8_HUMAN				Yes		
Q5VZ43_HUMAN			Yes			
Q6YL47_HUMAN				Yes		
Q6ZQU9_HUMAN				Yes		
Q6ZUG5_HUMAN	Yes					
Q86U89_HUMAN	Yes					
Q8IVF9_HUMAN				Yes		
Q8N1G8_HUMAN		Yes				
Q8NGM0_HUMAN			Yes			
Q8NH46_HUMAN	Yes					
Q96NP5_HUMAN		Yes				
Q9HAD2_HUMAN	Yes					
Q9NSQ0_HUMAN	Yes					
Q9NT31_HUMAN			Yes			
Q9NW32_HUMAN					Yes	
Q9Y6V0-3			Yes			
QDPR		Yes				
QPCT			Yes			
QRICH2		Yes				
QSER1			Yes			
RAB11FIP1	Yes					
RAB11FIP4			Yes			
RAB13	Yes					
RAB22A	Yes					
RAB3B	Yes					
RAD21	Yes		Yes			
RAD50				Yes		
RAD51B			Yes			
RAD51L1			Yes			
RAD54B	Yes					
RADIL			Yes			
RAG2			Yes			
RALGAPA2	Yes					
RALGPS1	Yes					
RALY					Yes	

RAP1GDS1	Yes					
RAPGEF2			Yes			
RAPGEF4	Yes	Yes	Yes			
RAPGEF5	Yes			Yes		
RAPGEF6	Yes			Yes		
RARB						Yes
RARRES3	Yes					
RASA2			Yes		Yes	
RASAL3				Yes		
RASEF	Yes					
RASGRF1			Yes			
RASGRP2		Yes		Yes		
RASIP1			Yes			
RASSF3	Yes					
RASSF6					Yes	
RAVER2		Yes				
RB1	No	No	No	No	No	No
RB1CC1				Yes		
RBBP8	Yes					
RBCK1			Yes			
RBFOX2			Yes			
RBL2	Yes					
RBM14		Yes				
RBM15				Yes		
RBM23	Yes		Yes			
RBM26	Yes					
RBM43			Yes			
RBMS2			Yes			
RC3H1			Yes			
RC3H2	Yes					
RCCD1	Yes					
RDH10	Yes					
RDH12				Yes		
RECQL4	Yes			Yes	Yes	
REG4			Yes			
RELB				Yes		
RELL2					Yes	
RELN			Yes			
REPIN1		Yes				
RFPL2			Yes			
RFT1		Yes				
RFWD3					Yes	
RFX6	Yes		Yes			
RGL2	Yes					
RGL3	Yes		Yes			
RGMB		Yes				
RGS20	Yes					
RGS22	Yes			Yes		
RGS3			Yes			
RGS7BP	Yes					
RHBDF2			Yes			
RHOA	Yes					
RHOT2			Yes			
RHPN2				Yes		
RIC8A					Yes	
RIMBP2				Yes		
RIMKLB		Yes				
RIMS2			Yes			
RIN1	Yes					
RIOK1			Yes			
RIOK2	Yes					
RIT1	Yes					
RLTPR	Yes					
RMI1		Yes			Yes	
RMND5A	Yes					
RNASEH1				Yes		
RNASEL				Yes		
RNF10	Yes					

RNF111	Yes					
RNF123				Yes		
RNF13	Yes					
RNF148		Yes				
RNF150					Yes	
RNF151	Yes					
RNF165				Yes		
RNF166			Yes			
RNF168			Yes			
RNF186			Yes			
RNF19A		Yes				
RNF20	Yes					
RNF213	Yes			Yes		
RNF216	Yes					
RNF220	Yes					
RNF31			Yes			
RNF39		Yes				
RNF40			Yes			
ROBO2	Yes					
ROCK1			Yes			
ROPN1	Yes					
RORB			Yes			
ROS1	Yes		Yes	Yes		
RP1			Yes			
RP11-1220K2.2			Yes		Yes	
RP11-366L20.2				Yes		
RP11-551L14.1				Yes		
RP11-694I15.6			Yes			
RPF2				Yes		
RPGR		Yes				
RPL10L		Yes				
RPL3L				Yes		
RPL4			Yes	Yes		
RPS11			Yes			
RPS19			Yes			
RPS6KA1				Yes		
RPS6KA2			Yes			
RPS6KA3	Yes					
RPS6KA5	Yes					Yes
RPS6KB1			Yes			
RPS6KL1						Yes
RRAD					Yes	
RRN3			Yes			
RRP7A	Yes					
RSBN1L					Yes	
RSF1	Yes			Yes		
RSPH1				Yes		
RSPH10B2	Yes					
RSPH4A			Yes			
RSPO3			Yes			
RSPRY1			Yes			
RTCA					Yes	
RTCD1					Yes	
RTDR1			Yes	Yes		
RTEL1	Yes		Yes			
RTL1				Yes	Yes	
RTN4IP1			Yes	Yes		
RTTN	Yes			Yes		
RUFY2			Yes		Yes	
RUNX1	No	No				
RUNX1T1	Yes					
RUSC2	Yes					
RWDD1			Yes			
RXFP1	Yes					
RYR1	Yes					
RYR2	Yes		Yes	Yes		
RYR3	Yes		Yes			
S100A10		Yes				

S100A16			Yes		
S1PR2	Yes				
SAAL1	Yes		Yes		
SACM1L	Yes				
SACS			Yes		
SAE1	Yes				
SALL1				Yes	
SALL3		Yes	Yes		
SAMD13			Yes		
SAMD14				Yes	
SAMD15				Yes	
SAMD9			Yes		
SAP30	Yes				
SARDH			Yes		Yes
SART3				Yes	
SASH1	Yes				
SAT1	Yes				
SBF2				Yes	
SBNO1			Yes		
SCAF1			Yes		
SCAI			Yes		
SCAPER			Yes		
SCD					Yes
SCGB1A1			Yes		
SCIN		Yes	Yes		
SCML2					Yes
SCN10A			Yes		
SCN2A			Yes		
SCN3B	Yes				
SCN5A				Yes	
SCN7A			Yes		
SCN8A		Yes		Yes	
SCNN1B			Yes		
SCO2			Yes		
SCRIB	Yes		Yes		
SCRN3					Yes
SCTR			Yes		
SCYL3			Yes		Yes
SDC4			Yes		
SDCBP			Yes		
SDCCAG3			Yes		
SDCCAG8				Yes	
SDF4	Yes				
SDHA			Yes		
SDK1			Yes		
SDK2					Yes
SDR42E1			Yes		
SEC23A					Yes
SEC24B			Yes	Yes	
SEC24C				Yes	
SEC31B			Yes		Yes
SEC61A2				Yes	
SEC62			Yes		
SECISBP2L			Yes		
SECTM1	Yes				
SEL1L			Yes		
SEL1L2		Yes			
SEL1L3	Yes	Yes	Yes		
SELE	Yes				
SELENBP1	Yes		Yes		
SELP			Yes		
SEMA3E	Yes				
SEMA3G				Yes	
SEMA4B			Yes		
SEMA4C				Yes	
SEMA4G				Yes	
SEMA5A		Yes	Yes		
SEMA5B	Yes				

SEMG2			Yes		
SENP1					Yes
SENP2	Yes				
SENP7			Yes		
SEPT6				Yes	
SERP1			Yes		
SERPINA2	Yes				
SERPINA6			Yes		
SERPINB2		Yes			
SERPINB3			Yes		
SERPINB4				Yes	
SERPINB7			Yes		
SERPIND1	Yes				
SERPINI2				Yes	
SERTAD1			Yes		
SETBP1		No			No
SETDB1			Yes		
SETX			Yes		
SEZ6L		Yes			
SF1		Yes			
SF3A1			Yes		
SF3A2			Yes		
SF3A3	Yes				
SF3B1	No			No	
SF3B2	Yes		Yes		
SFR1			Yes		
SFRP2	Yes				
SG269_HUMAN			Yes		
SGCG				Yes	
SgK085				Yes	
SGK1			Yes		
SgK269			Yes		
SGK3	Yes				
SGOL2			Yes		
SGSM2					Yes
SH2B1	Yes				
SH2B3			Yes		
SH3BGR		Yes			
SH3BP5L				Yes	
SH3D21	Yes				
SH3GL3				Yes	
SH3KBP1			Yes		
SHC3		Yes			
SHE				Yes	
SHKBP1					Yes
SHOX2	Yes				
SHPRH				Yes	
SI			Yes		
SIDT1	Yes				
SIGLEC10	Yes				Yes
SIGLEC12	Yes			Yes	Yes
SIGLEC14	Yes				
SIK3			Yes	Yes	
SIL1			Yes		
SIM1			Yes		
SIPA1L1		Yes	Yes		
SIPA1L2	Yes				Yes
SIPA1L3		Yes			
SIRPB2				Yes	
SIRT2				Yes	
SIRT3	Yes				
SIT1	Yes				
SKAP2			Yes		
SKOR1				Yes	
SLC10A3	Yes				
SLC12A2		Yes			
SLC12A6			Yes		
SLC13A4			Yes		

SLC15A2			Yes			
SLC16A12			Yes			
SLC17A4	Yes					
SLC17A5				Yes		
SLC17A7	Yes					
SLC1A3	Yes					
SLC1A6			Yes			
SLC22A11				Yes		
SLC22A13			Yes			
SLC22A17					Yes	
SLC22A25					Yes	
SLC22A3			Yes			
SLC22A4			Yes			
SLC22A8	Yes					
SLC22A9	Yes					
SLC23A3	Yes					
SLC24A5				Yes		
SLC25A2					Yes	
SLC25A20		Yes				
SLC25A23					Yes	
SLC25A25	Yes					
SLC25A27		Yes				
SLC25A29		Yes				
SLC25A38	Yes				Yes	
SLC25A40				Yes		
SLC25A44	Yes					
SLC25A46			Yes			
SLC26A1				Yes		
SLC26A10	Yes					
SLC26A4				Yes		
SLC26A7		Yes				
SLC26A8	Yes	Yes		Yes		
SLC26A9				Yes		
SLC28A1				Yes		
SLC28A2				Yes		
SLC28A3			Yes			
SLC2A1	Yes					
SLC2A3			Yes			
SLC30A3				Yes		
SLC30A9	Yes					
SLC35B1				Yes		
SLC35B4		Yes				
SLC35F1	Yes		Yes			
SLC36A3	Yes					
SLC37A4	Yes					
SLC38A4				Yes		
SLC38A6				Yes		
SLC39A12		Yes				
SLC39A6				Yes		
SLC44A2	Yes					
SLC44A4	Yes			Yes		
SLC45A4	Yes		Yes			
SLC4A11			Yes		Yes	
SLC4A3			Yes			
SLC4A4	Yes					
SLC4A5	Yes					
SLC5A1		Yes				
SLC5A11				Yes		
SLC5A3				Yes		
SLC5A7		Yes				
SLC5A8			Yes			
SLC5A9	Yes			Yes		
SLC6A15			Yes			
SLC6A18			Yes			
SLC6A20					Yes	
SLC6A8			Yes			
SLC6A9	Yes		Yes			
SLC7A11			Yes			

Gene	Col1	Col2	Col3	Col4	Col5	Col6
SLC7A7	Yes		Yes			
SLC7A8			Yes			
SLC9A11		Yes				
SLC9A5		Yes				
SLCO1B3	Yes					
SLCO5A1			Yes			
SLFN11	Yes					
SLFN5	Yes					
SLITRK2			Yes			
SLITRK5				Yes		
SLX4	Yes					
SMAD2			Yes			
SMAD4	No	No	No	No	No	No
SMARCAD1	Yes		Yes			
SMARCAL1			Yes			
SMARCB1	Yes					
SMARCC1	Yes					
SMC2		Yes				
SMC2L1		Yes				
SMCP		Yes				
SMCR7				Yes		
SMCR7L			Yes			
SMG1			Yes			
SMG5	Yes					
SMO			Yes			
SMTN				Yes		
SMURF2			Yes			
SMYD1					Yes	
SNAI2			Yes			
SNAP25					Yes	
SNAPC1			Yes			
SNAPC4				Yes		
SNCG			Yes			
SND1			Yes			
SNIP1		Yes				
SNTG1			Yes			
SNX13		Yes				
SNX17			Yes			
SNX2				Yes		
SNX21		Yes				
SNX25			Yes			
SNX32			Yes	Yes		
SNX9			Yes			
SOCS1		No		No	No	No
SOCS2					Yes	
SON					Yes	
SORBS2			Yes			
SORL1	Yes					
SORT1					Yes	
SOX21			Yes			
SOX4	Yes					
SOX6			Yes			
SP100			Yes			
SP140	Yes					
SPACA1				Yes		
SPACA3				Yes		
SPACA7			Yes			
SPAG16			Yes			
SPAM1			Yes			
SPATA18				Yes		
SPATA2				Yes		
SPATA9				Yes		
SPATS1			Yes			
SPEF1	Yes					
SPEG	Yes					
SPEM1		Yes				
SPEN			Yes		Yes	
SPG11	Yes			Yes		

SPG21			Yes	Yes		
SPINK5			Yes			
SPOCD1			Yes			
SPOCK2				Yes		
SPOCK3			Yes			
SPON2	Yes					
SPP1			Yes			
SPRED2	Yes		Yes			
SPRYD3	Yes					
SPTAN1	Yes					
SPTBN1			Yes			
SPTBN2				Yes	Yes	
SPTBN4	Yes					
SPTLC3		Yes				
SQLE			Yes			
SQSTM1		Yes				
SRA1		Yes				
SRC	No	No	Yes			
SRCAP				Yes		
SRCRB4D		Yes				
SREBF2			Yes			
SRF			Yes			
SRGAP1			Yes			
SRGN			Yes			
SRRM1		Yes				
SRRM2	Yes	Yes	Yes			
SRSF5					Yes	
SRSF6			Yes			
SS18	Yes					
SSRP1				Yes		
SSX1	Yes					
SSX3	Yes					
ST3GAL4	Yes	Yes				
ST6GAL1				Yes		
ST6GAL2			Yes			
ST8SIA5	Yes					
STAB1				Yes		
STAB2		Yes		Yes		
STAG1	Yes		Yes			
STAG2				Yes		
STAM2				Yes		
STAT3	Yes				Yes	
STAT4			Yes			
STAU2				Yes		
STK11		No	No	No	No	No
STK24			Yes			
STK25			Yes			
STK31	Yes		Yes	Yes	Yes	Yes
STK36				Yes		
STK4				Yes		
STOX1			Yes			
STRADA			Yes			
STRBP	Yes					
STRC					Yes	
STT3A	Yes					
STUB1				Yes		
STX17					Yes	
SUFU			Yes			
SUGT1		Yes				
SULT1C3				Yes		
SUN2			Yes			
SUN3	Yes					
SUPT16H	Yes					
SUPT6H			Yes			
SUPT7L			Yes			
SURF6		Yes				
SUSD2		Yes				
SUSD5				Yes		

Gene						
SUZ12				Yes		
SV2A				Yes		
SVEP1		Yes	Yes			
SVOPL	Yes					
SWAP70					Yes	
SWT1			Yes			
SYBU				Yes		
SYCE1			Yes	Yes		
SYCE1L			Yes			
SYCP2		Yes				
SYDE2				Yes		
SYMPK	Yes			Yes		
SYN3	Yes					
SYNE1	Yes					
SYNGAP1	Yes			Yes		
SYNGR4				Yes		
SYNJ1				Yes		
SYNJ2			Yes			
SYNM			Yes			
SYNPO2L					Yes	
SYT10			Yes			
SYT12	Yes					
SYT4			Yes			
SYTL2			Yes			
SYTL3	Yes					
SYTL4		Yes				
TAB2			Yes			
TAF5L			Yes			
TAF7			Yes			
TANC1	Yes					
TAOK3			Yes			
TAP1	Yes					
TARBP1	Yes			Yes		
TARS			Yes			
TARSL2		Yes	Yes			
TAS1R2	Yes			Yes		
TAS2R19			Yes			
TAS2R43	Yes					
TAS2R46	Yes					
TAS2R5	Yes					
TAT		Yes				
TATDN3	Yes			Yes		
TBC1D1	Yes					
TBC1D10B			Yes			
TBC1D15		Yes				
TBC1D19		Yes				
TBC1D4			Yes		Yes	
TBC1D7	Yes		Yes			
TBC1D8B				Yes		
TBCCD1	Yes		Yes			
TBCK	Yes					
TBK1			Yes			
TBL3			Yes	Yes		
TBP				Yes		
TBX10				Yes		
TBX19			Yes			
TBX2				Yes		
TBXAS1					Yes	
TCEA1		Yes				
TCEA3			Yes			
TCEB1				Yes		
TCERG1L		Yes		Yes		
TCF12	Yes		Yes			
TCF20	Yes		Yes			
TCF25	Yes	Yes				
TCF3			Yes			
TCF7L2			Yes			
TCFL5				Yes		

TCHH	Yes				
TCP1			Yes		
TCP11	Yes				
TCTEX1D1			Yes		
TDGF1				Yes	
TDRD1			Yes		
TDRD5	Yes				
TDRD6		Yes		Yes	
TEAD3				Yes	
TEC	Yes				
TELO2			Yes		
TENC1				Yes	
TENM1		Yes			
TENM2	Yes				
TET1				Yes	
TET2				Yes	
TET3	Yes				
TEX10	Yes				
TEX13B				Yes	
TEX14	Yes		Yes		
TEX15	Yes				
TFCP2L1			Yes		
TFDP3				Yes	
TFIP11	Yes	Yes			
TFRC	Yes				
TG		Yes			
TGFBR2	Yes				
TGFBRAP1	Yes		Yes		
TGS1	Yes				
THADA				Yes	
THAP2					Yes
THAP4	Yes				
THAP9				Yes	
THBS2			Yes		
THBS3				Yes	
THOC3		Yes			
THOC6	Yes				
THRAP3				Yes	
TIAM2				Yes	
TIE1	Yes				
TIGD4			Yes		
TIMELESS			Yes		
TIRAP					Yes
TJP1					Yes
TLE1	Yes				
TLE4		Yes			
TLK2			Yes		
TLL2	Yes				
TLN1				Yes	
TLN2				Yes	
TLR2				Yes	
TLR3				Yes	
TLR6	Yes				
TLX1			Yes		
TM6SF1	Yes				
TM7SF4			Yes		
TM9SF1	Yes				
TM9SF2			Yes		
TM9SF3		Yes			
TMC1	Yes				
TMCC2				Yes	
TMCO3				Yes	
TMED5				Yes	
TMED8				Yes	
TMEFF2	Yes				
TMEM106A	Yes				
TMEM110			Yes		
TMEM123	Yes		Yes		Yes

TMEM132A	Yes		Yes			
TMEM132B			Yes			
TMEM132E		Yes	Yes			
TMEM146			Yes			
TMEM14A			Yes			
TMEM151A	Yes					
TMEM164		Yes				
TMEM174			Yes			
TMEM175		Yes				
TMEM178				Yes		
TMEM187				Yes		
TMEM19			Yes			
TMEM199				Yes		
TMEM201					Yes	
TMEM209	Yes					
TMEM39A	Yes					
TMEM55B				Yes		
TMEM57				Yes		
TMEM62	Yes					
TMEM75	Yes					
TMEM87A					Yes	
TMEM87B		Yes				
TMEM89			Yes			
TMEM8B				Yes		
TMEM91			Yes			
TMOD2				Yes		
TMPO	Yes		Yes			
TMPPE	Yes					
TMPRSS11B		Yes	Yes			
TMPRSS13					Yes	
TMTC1			Yes	Yes		
TMUB1	Yes					
TMX1				Yes		
TNC			Yes			
TNFRSF13B	Yes					
TNFRSF8	Yes					
TNFSF11			Yes			
TNIK		Yes	Yes			
TNK2	Yes			Yes		
TNKS			Yes			
TNKS2	Yes					
TNMD				Yes		
TNN			Yes			
TNNI1					Yes	
TNNI3K	Yes					
TNNT1	Yes		Yes			
TNR				Yes		
TNRC6A					Yes	
TNS1	Yes					
TNS4	Yes					
TOM1			Yes			
TOMM20L			Yes			
TOMM22			Yes			
TOP2A	Yes					
TOP3B				Yes		
TOR1AIP2	Yes					
TP53	Yes	No	Yes		Yes	No
TP53BP1			Yes	Yes		
TP53TG5				Yes		
TP63				Yes	Yes	
TP73			Yes			
TPH2			Yes			
TPK1			Yes			
TPM3	Yes					
TPP1		Yes				
TPR				Yes		
TPSAB1			Yes			
TPST1	Yes					

TPX2			Yes			
TRAD	Yes					
TRAF3IP2	Yes					
TRAF4	Yes					
TRAF5			Yes			
TRAK2				Yes		
TRAPPC13					Yes	
TRAPPC8	Yes			Yes		
TRDN		Yes				
TRIM10	Yes		Yes			
TRIM13		Yes				
TRIM29				Yes		
TRIM3				Yes		
TRIM32			Yes			
TRIM33		Yes			Yes	
TRIM36				Yes		
TRIM41	Yes		Yes			
TRIM46				Yes		
TRIM60	Yes					
TRIM64B			Yes			
TRIM64C			Yes	Yes		
TRIM66			Yes			
TRIML2	Yes					
TRIOBP					Yes	
TRIP11			Yes			
TRIQK			Yes			
TRMT1L			Yes			
TROAP				Yes		
TRPA1		Yes				
TRPC1	Yes					
TRPC4			Yes	Yes		
TRPC5	Yes	Yes				
TRPM1				Yes		
TRPM2		Yes				
TRPM3			Yes			
TRPS1	Yes	Yes	Yes		Yes	Yes
TRPV4				Yes		
TRPV6	Yes					
TRRAP			Yes			
TSGA13			Yes			
TSHR			Yes			
TSLP		Yes				
TSNARE1	Yes			Yes		
TSNAXIP1	Yes					
TSPAN2			Yes			
TSPEAR				Yes		
TSPYL2		Yes		Yes		
TSR1	Yes					
TSSK2				Yes		
TTBK1				Yes		
TTC13				Yes	Yes	
TTC16		Yes				
TTC18	Yes					
TTC21B			Yes			
TTC22				Yes		
TTC25					Yes	
TTC28					Yes	
TTC30B	Yes		Yes			
TTC31			Yes			
TTC33				Yes		
TTC39B		Yes				
TTC39C				Yes		
TTC40				Yes		
TTC5				Yes		
TTC7B				Yes		
TTL				Yes		
TTLL5				Yes		
TTN	Yes	Yes	Yes	Yes		

TTYH3	Yes				
TUBA3E	Yes				
TUBB			Yes		
TUBB8				Yes	
TUBD1				Yes	
TUBG2			Yes		
TUBGCP6				Yes	
TULP3	Yes				
TULP4			Yes		
TWF1				Yes	
TXLNG	Yes				
TYK2		Yes			
U2AF1L4				Yes	
U2AF2	Yes				
U2SURP	Yes				
UBA1			Yes		
UBA3		Yes			
UBA5				Yes	
UBASH3B	Yes				
UBE2D3		Yes			
UBE2NL				Yes	
UBE2O			Yes		
UBE3A				Yes	
UBE3B				Yes	
UBFD1			Yes		
UBN2	Yes				
UBQLN3				Yes	
UBR3				Yes	
UBTD2	Yes				
UBXN2B				Yes	
UCK1		Yes			
UFSP2				Yes	
UGP2			Yes		
UGT1A4	Yes				
UGT1A7		Yes			
UGT2A1				Yes	
UGT2B11			Yes	Yes	
UGT2B28	Yes				
UGT2B7			Yes		
UGT3A1				Yes	
UHMK1				Yes	
UHRF1BP1	Yes				
UHRF1BP1L	Yes			Yes	
ULK1				Yes	
UNC13B			Yes		
UNC13C			Yes	Yes	Yes
UNC45A			Yes		
UNC5B	Yes				
UNK		Yes		Yes	
URB1	Yes				
URB2	Yes				
USH2A			Yes		
USHBP1			Yes		
USP10			Yes		
USP11	Yes				
USP14	Yes				
USP15					Yes
USP19				Yes	
USP20			Yes		
USP28					Yes
USP31	Yes				
USP33			Yes	Yes	
USP34	Yes		Yes		
USP37		Yes		Yes	
USP38	Yes		Yes	Yes	
USP4			Yes	Yes	
USP5			Yes		
USP6	Yes				

Gene	1	2	3	4	5	6
VAC14	Yes					
VAMP2			Yes			
VARS			Yes			
VARS2				Yes		
VASN			Yes			
VAV2			Yes			
VCAN	Yes					
VCPIP1		Yes	Yes			
VCX2			Yes			
VEGFC	Yes	Yes	Yes	Yes	Yes	Yes
VEPH1		Yes				
VEZF1			Yes			
VHL	No	No	No	No	No	No
VHLL			Yes			
VIT				Yes		
VN1R2				Yes		
VN2R1P			Yes			
VNN2			Yes			
VPS13B	Yes		Yes	Yes		Yes
VPS13C				Yes		
VPS13D	Yes					
VPS25		Yes				
VPS26B			Yes			
VPS35	Yes					
VPS41					Yes	
VPS53			Yes			
VPS8	Yes		Yes			
VSIG2			Yes			
VWA2			Yes			
VWA3A			Yes			
VWA3B	Yes		Yes			
VWA5A					Yes	
VWDE		Yes				
WASF3		Yes				
WASL	Yes					
WDFY3	Yes					
WDFY4		Yes	Yes			
WDR17	Yes					
WDR19	Yes				Yes	
WDR27			Yes			
WDR3			Yes			
WDR37					Yes	
WDR44	Yes					
WDR52	Yes			Yes		
WDR64		Yes				
WDR67			Yes			
WDR87	Yes					
WDSUB1			Yes			
WFDC5				Yes		
WFIKKN2			Yes			
WHSC1L1	Yes			Yes		
WNK1	Yes					
WNK2	Yes	Yes				
WNK4	Yes					
WNT10B			Yes			
WNT2				Yes		
WNT7A			Yes			
WRN	Yes					
WSCD1	Yes					
WT1	Yes	No			No	
WWC1			Yes	Yes		
WWOX	Yes					
XAB2			Yes			
XAGE5				Yes		
XIRP1	Yes					
XIRP2		Yes	Yes	Yes		
XKR3			Yes			
XPNPEP1				Yes		

XPNPEP3			Yes		
XPO6	Yes				
XPOT				Yes	
XPR1				Yes	
XRCC4		Yes			
XRN1				Yes	
XRRA1					Yes
YARS2			Yes		
YLPM1	Yes			Yes	
YWHAB				Yes	
YY1AP1			Yes	Yes	
ZBBX			Yes		
ZBTB10	Yes				
ZBTB2			Yes		
ZBTB26				Yes	
ZBTB3			Yes		
ZBTB44			Yes		
ZBTB5	Yes				
ZBTB6			Yes		
ZBTB8B	Yes				
ZC3H11A				Yes	
ZC3H13	Yes	Yes			
ZC3H15	Yes			Yes	
ZC3H3	Yes				
ZC3HAV1				Yes	
ZCCHC11					Yes
ZCCHC14			Yes		
ZCCHC2			Yes		
ZCRB1		Yes			
ZDHHC11	Yes				
ZDHHC16			Yes		
ZDHHC5				Yes	
ZEB1			Yes		
ZEB2				Yes	
ZER1			Yes		
ZFAT		Yes	Yes	Yes	
ZFC3H1		Yes	Yes		
ZFHX4	Yes			Yes	Yes
ZFP1			Yes		
ZFP14	Yes				
ZFP36L1				Yes	
ZFP41			Yes		
ZFP42		Yes			
ZFP62			Yes		
ZFPM2		Yes	Yes		
ZFR			Yes		
ZFX				Yes	
ZFYVE1		Yes			
ZFYVE16			Yes		Yes
ZFYVE26				Yes	
ZFYVE9	Yes				
ZHX1	Yes				
ZHX2				Yes	Yes
ZIC1	Yes				
ZIK1		Yes			Yes
ZIM2	Yes	Yes	Yes		
ZMAT1	Yes				
ZMPSTE24	Yes				
ZMYM6	Yes		Yes		
ZMYND11	Yes				
ZNF10				Yes	
ZNF121				Yes	
ZNF135	Yes				
ZNF146					Yes
ZNF148			Yes		
ZNF157		Yes			
ZNF160				Yes	
ZNF167	Yes				Yes

ZNF182		Yes				
ZNF184		Yes				
ZNF189			Yes			
ZNF192	Yes					
ZNF197		Yes			Yes	
ZNF20			Yes			
ZNF215			Yes			
ZNF22		Yes		Yes		
ZNF222			Yes			
ZNF229		Yes				
ZNF233	Yes					
ZNF234	Yes					
ZNF236			Yes	Yes		
ZNF250	Yes		Yes			
ZNF254		Yes				
ZNF257				Yes		
ZNF264			Yes			
ZNF274				Yes		
ZNF292		Yes	Yes			
ZNF3		Yes			Yes	
ZNF30				Yes		
ZNF300			Yes			
ZNF318				Yes		
ZNF319		Yes				
ZNF324B				Yes		
ZNF330				Yes		
ZNF345			Yes			
ZNF346				Yes		
ZNF347		Yes				
ZNF350			Yes			
ZNF354A	Yes		Yes			
ZNF382					Yes	
ZNF385	Yes					
ZNF385A	Yes					
ZNF385D				Yes		
ZNF398	Yes					
ZNF407				Yes		
ZNF415		Yes				
ZNF423	Yes					
ZNF425		Yes				
ZNF43			Yes			
ZNF430			Yes			
ZNF432		Yes			Yes	
ZNF433				Yes		
ZNF436		Yes				
ZNF440	Yes					
ZNF441		Yes			Yes	
ZNF443		Yes		Yes		
ZNF445	Yes					
ZNF45			Yes			
ZNF451	Yes					
ZNF462	Yes					
ZNF467		Yes				
ZNF473			Yes			
ZNF480					Yes	
ZNF483			Yes			
ZNF491	Yes				Yes	
ZNF506			Yes			
ZNF511			Yes			
ZNF518B	Yes	Yes				
ZNF519			Yes			
ZNF521						Yes
ZNF527					Yes	
ZNF528			Yes			
ZNF532	Yes					
ZNF536	Yes					
ZNF540			Yes			
ZNF554	Yes					

Gene						
ZNF556		Yes				
ZNF564					Yes	
ZNF571	Yes					
ZNF572		Yes				
ZNF577		Yes			Yes	
ZNF582				Yes		
ZNF594	Yes					
ZNF598				Yes		
ZNF610		Yes			Yes	
ZNF611				Yes		
ZNF615		Yes			Yes	
ZNF619		Yes	Yes			
ZNF621		Yes				
ZNF630			Yes		Yes	
ZNF654	Yes					
ZNF660					Yes	
ZNF662				Yes		
ZNF664		Yes				
ZNF667	Yes					
ZNF668		No				
ZNF673	Yes					
ZNF677		Yes				
ZNF678		No			Yes	
ZNF688		Yes				
ZNF69			Yes			
ZNF699					Yes	
ZNF700				Yes		
ZNF705G			Yes	Yes		
ZNF708	Yes	Yes				
ZNF713			Yes			
ZNF716			Yes			
ZNF720			Yes			
ZNF729			Yes			
ZNF736	Yes					
ZNF738	Yes					
ZNF74				Yes		
ZNF746	Yes					
ZNF747	Yes					
ZNF76			Yes			
ZNF775			Yes			
ZNF781	Yes				Yes	
ZNF799			Yes			
ZNF8	Yes				Yes	
ZNF80		Yes			Yes	
ZNF81	Yes					
ZNF814			Yes			
ZNF821			Yes			
ZNF829	Yes					
ZNF839			Yes			
ZNF845	Yes					
ZNF846				Yes		
ZNF852					Yes	
ZNF860			Yes			
ZNF862		Yes				
ZNF92					Yes	
ZNF99		Yes			Yes	
ZNFX1				Yes		
ZNRF4				Yes		
ZPBP			Yes			
ZRANB3				Yes		
ZSCAN1	Yes					
ZSCAN23	Yes					
ZSCAN29		Yes				
ZSWIM5	Yes		Yes			
ZSWIM6	Yes					
ZWINT			Yes			
ZYX				Yes		

Appendix 7: A collection of gene mutation status in a panel of lung cancer cell lines. No indicates wild type gene status and Yes indicates mutant gene status.

Gene Name	A549	H226	H292	H460	H1650	HCC95	Hop92
AARS		Yes					
AATK	Yes						
ABCA3					Yes		
ABCA9	Yes						
ABCB1	Yes						
ABCC11		Yes					
ABCD2	Yes						
ABHD3							Yes
ACACA		Yes			Yes		
ACAT2				Yes			
ACER1							Yes
ACOXL	Yes						
ACTN1	Yes						
ACTRT1							Yes
ACVR2A							Yes
ADAM17	Yes	Yes			Yes		
ADAM28	Yes	Yes		Yes	Yes	Yes	
ADAM29				Yes	Yes		
ADAMTS19	Yes						
ADAMTS2					Yes	Yes	
ADAMTS7	Yes						
ADCK5						Yes	
ADCYAP1R1							Yes
ADH1A	Yes						
ADM2	Yes						
AGPAT9			Yes				
AGTR1	Yes						
AHNAK				Yes			
AIM1	Yes						
AKAP12	Yes	Yes		Yes	Yes	Yes	
AKAP9	Yes	Yes		Yes		Yes	
AKT1	No	No	No	No	No	No	
ALPK1		Yes			Yes		
ALPK2		Yes		Yes	Yes		
ALS2						Yes	
ALS2CL	Yes						
AMT	Yes						
ANGEL1				Yes			
ANKFY1		Yes					
ANKHD1							Yes
ANKHD1-EIF4EBP3							Yes
ANKIB1	Yes						
ANKMY2				Yes			
ANKZF1	Yes						
ANXA2			Yes				
AP5B1		Yes					
APC		Yes					
APLF		Yes					
APOB	Yes						
APPL2				Yes			
AQP2							Yes
ARFGEF2				No			
ARHGAP25	Yes						
ARHGEF18	Yes						
ARID1A				Yes			
ARID2					No		
ARID3A	Yes						
ARIH1				Yes			
ARMC1	Yes						
ARMC4				Yes			
ARPC1B	Yes						
ARSD		Yes					

Gene							
ARSH	Yes						
ARSI	Yes						
ARVCF							Yes
ASB4				Yes			
ASB5	Yes						
ASTN2		Yes					
ATAD3B				Yes			
ATAD5				Yes			
ATG13	Yes						
ATG4B		No					
ATN1			Yes				
ATP1A3	Yes						
ATP6V0D1							Yes
ATR				No			
ATRNL1							Yes
AURKC					Yes		
AXL		No				Yes	
BAHD1							Yes
BARD1	Yes						
BBOX1							Yes
BBS2	Yes						
BBS7				Yes			
BCAR3		Yes					
BCAT1	Yes						
BCAT2		Yes					
BCCIP	Yes						
BCL3	Yes						
BCORL1							Yes
BCR	Yes				Yes		
BCS1L				Yes			
BIN1			Yes				
BIRC6						Yes	
BMP4		Yes					
BMS1	Yes						
BNC2							Yes
BRAF	No	No	No	No	No		No
BRD3	Yes			Yes			
BRPF1				Yes			
BTBD11	Yes						
BTG1	Yes	Yes					
BTNL9				Yes			
BUD13				Yes			
BZRAP1				Yes			
C10orf112	Yes						
C10orf68	Yes			Yes			Yes
C11orf40				Yes			
C11orf83	Yes						
C12orf35							Yes
C12orf65			Yes				
C14orf126			Yes				
C14orf23	Yes						
C15orf2				Yes			
C15orf42	Yes						
C15orf55				Yes			
C16orf88		No					
C20orf107							Yes
C2orf15	Yes						
C2orf55	Yes						
C2orf62				Yes			
C2orf67	Yes						
C3orf70	Yes						
C4BPB				Yes			
C7orf31	Yes						
C7orf34							Yes
C7orf63				Yes			
C9orf174	Yes						
C9orf96							Yes
CABC1					Yes		

Gene							
CACNA1B	Yes						
CACNB2		Yes			Yes	Yes	
CADM3	Yes						
CADPS				Yes			
CAGE1	Yes						
CALD1				Yes			
CaMK1b							Yes
CANX	Yes						
CAPN7	Yes						
CARD10		Yes		Yes	Yes		
CARD11	Yes						Yes
CASC1		Yes					
CASC5			Yes			Yes	
CASP10		Yes					
CASP4				Yes			
CASQ2				Yes			
CBL	Yes						
CBLB							Yes
CC2D2A	Yes						
CCBL1	Yes						
CCDC129	Yes						
CCDC132	Yes						
CCDC80	Yes						
CCKBR							Yes
CCNB1	Yes						
CCNT1	Yes						
CCNYL1			Yes				
CCR5					Yes		
CCR7				Yes			
CD164				Yes			
CD2AP	Yes						
CD300C	Yes						
CD83				Yes			
CDC42	No			No			
CDCP2	Yes						Yes
CDH1			Yes				
CDH2				Yes			
CDK11B				Yes	Yes	Yes	
CDK2	No			No			
CDK4	No			No			
CDK5RAP3		Yes	Yes				
CDK6	No			No			
CDKN2A	Yes	Yes	Yes	Yes	Yes		
CDKN2AIPNL		Yes					
CDKN2C	No						
CEBPE	Yes						
CEBPZ		Yes		Yes			
CEP110				Yes			
CEP250			Yes				
CHCHD2			Yes				
CHD1		Yes					
CHD7	Yes						
CHEK2	No			No			
CHRDL2	Yes						
CHRNB4	Yes						
CHST8	Yes						
CHTF18			Yes				
CIAO1	Yes						
CIITA					Yes		
CKAP5	Yes						
CKMT2		Yes					
CLCN7	Yes						
CLDN16		Yes					
CLMP							Yes
CLTC			Yes				
CLTCL1	Yes	Yes		Yes	Yes	Yes	
CNDP1				Yes			Yes
CNPPD1	Yes						

CNTLN	Yes						
CNTN6					Yes		
CNTNAP2							Yes
COL14A1				Yes			
COL18A1		Yes					
COL21A1	Yes						
COL4A1							Yes
COL4A6	Yes						
COL6A3		Yes					
COQ2			Yes				
CPO							Yes
CPPED1	Yes						
CRB1							Yes
CREB3L2	Yes			Yes	Yes	Yes	
CREB3L4			Yes				
CRISP2							Yes
CRP					Yes		
CSF1R	Yes				Yes		
CSMD2		Yes					
CSNK2A1				Yes			
CTNNA1	Yes						
CTNNB1	No	No	No	No	No	No	No
CTNND2	Yes						
CTSA				Yes			
CUL2	Yes						
CUL9				Yes			
CXCL12						Yes	
CXorf22	Yes						
CYB5R4				Yes			
CYP27B1	Yes						
CYP39A1	Yes						
CYP7A1	Yes						
CYTH4				Yes			
DAB1							Yes
DACH1	Yes					Yes	
DACT1	Yes						
DAXX						Yes	
DBC1				Yes			
DCAF4L2				Yes			
DCAF5	Yes						
DCAF6	Yes						
DCHS2		Yes					
DCLK3	Yes						
DDX60L	Yes						
DFNA5		Yes			Yes		
DGKB				Yes		Yes	
DGKE	Yes						
DHCR7				Yes			
DHDH				Yes			
DHX33			Yes				
DHX58	Yes						
DIP2C				Yes			
DLG2	Yes						
DMGDH	Yes						
DMXL2		Yes					
DNAH8	Yes	Yes		Yes			
DNAH9				Yes			
DNAJB12				Yes			
DNAJC5B							Yes
DOCK7		Yes					
DOK2		Yes					
DPP10		Yes		Yes			
DPP8	Yes						
DSP			Yes				
DST			Yes				
DTNB	Yes						
DYNC1I1	Yes						
DYRK1A		Yes					

EDNRA					Yes		
EGF					Yes		
EGFR	No	No	No	No	Yes		No
EGLN1	Yes						
EIF2AK1						Yes	
EIF2AK3						Yes	
EIF2AK4		Yes					
EIF2C2	Yes						
EIF3A				Yes			
EIF4A2						Yes	
EIF4B				Yes			
EIF4E	Yes				Yes		
EIF4G3		Yes					
ELL2	Yes						
ENPEP							Yes
ENSG00000176515	Yes						
ENSG00000214944		Yes					
ENSG00000216560	Yes						
EP300		Yes		Yes			
EPB41L3	Yes						
EPHA1	Yes						
EPHA5		Yes					Yes
EPHA6	Yes					Yes	
EPHA7				Yes			Yes
EPHB6	Yes						
ERAP2				Yes			
ERBB2	No	No	No	No	No	No	
ERBB2IP			Yes				
ERBB3	No			No			
ERBB4	No			No			
ERC1	Yes						
ERCC3		Yes					
ERCC5			Yes				
ERG					Yes		
ERV3-1			Yes				
ETV1					Yes		
EXOC3			Yes	Yes			
EXT1					Yes	Yes	
F12	Yes						
FAM160B2	Yes						
FAM171A2	Yes						
FAM174A	Yes						
FAM175B	Yes						
FAM184A				Yes			
FAM40B			Yes				
FAM47B							Yes
FAM47C	Yes						
FAM54B	Yes						
FAM55B	Yes						
FAM59A							Yes
FAM5B				Yes			
FAM5C							Yes
FAM63A	Yes						
FAM63B	Yes						
FAM71B							Yes
FAM71C				Yes			
FAM75D1	Yes						
FAM75D4	Yes						
FAM82B	Yes						
FANCA						Yes	
FANCB							Yes
FANCD2					Yes		
FASTKD3		Yes					
FAT	No	Yes	Yes				
FAT1	No	Yes	Yes				
FAT2							Yes
FAT4							Yes
FBN2	Yes						

Gene	1	2	3	4	5	6	7
FBXL2	Yes						
FBXL5		Yes					
FBXO32							Yes
FBXO48	Yes						
FBXW7							No
FCGR3A	Yes						
FCHSD2	Yes						
FCRL2	Yes						
FGD1	Yes						
FGF5							Yes
FGFBP2	No						
FGFR1OP	Yes	Yes		Yes	Yes	Yes	
FGFR3	Yes				Yes		
FGL1	Yes						
FH	Yes						
FLI1	Yes	Yes		Yes		Yes	
FLJ16360	Yes						
FLNC	Yes						
FLT1	No			No			
FLT3	Yes						
FLT4				No	Yes		
FLVCR1	Yes						
FLYWCH1				Yes			
FMN1	Yes						
FMN2	Yes			Yes	Yes	Yes	
FMO5							Yes
FN1				Yes	Yes	Yes	
FNDC3A		Yes					
FOXL2	No	No		No			No
FOXQ1							Yes
FRA10AC1		Yes					
FRAS1	Yes						
FRMD6	Yes						
FRY							Yes
FSTL5	Yes						
FUBP3			Yes				
FUCA2	Yes						
FUS	Yes						
FZD10							Yes
GABRB1				Yes			
GABRP							Yes
GABRR1							Yes
GAD2	Yes						
GALNT12				Yes			
GALNT13				Yes			Yes
GAS2L3	Yes			Yes			
GAS7				Yes			
GCDH		Yes					
GDF15	Yes						
GDPD3				Yes			
GEMIN4			Yes				
GHSR				Yes			
GIMAP1	Yes						
GIMAP2		Yes					
GJB4				Yes			
GLB1			Yes				
GLB1L3	Yes						
GLE1	Yes						
GLE1L	Yes						
GNAL				Yes			
GNAQ					No		
GNAS						Yes	Yes
GNL2			Yes				
GNPTAB				Yes			
GOLGA4	Yes			Yes			
GOLGA5				Yes			
GOLGB1		Yes		Yes			
GPR112				Yes		Yes	

GPR137C	Yes						
GPR85	Yes						
GRHL1	Yes						
GRIA2	Yes						
GRIA3	Yes			Yes	Yes	Yes	
GRIA4	Yes						
GRIN2A				Yes			
GRK7							Yes
GRM2							Yes
GRM7	Yes						
GTSE1	Yes						
GUCY1A2							Yes
GUCY2C	Yes			Yes	Yes		
HBXIP				Yes			
HCLS1	Yes						
HECW1				Yes			
HEG1			Yes				
HERC1	Yes						
HFM1	Yes						
HIF1A	Yes				Yes		
HIP1	Yes						
HIP1R	Yes						
HIST1H1C							Yes
HIST1H2AJ	Yes						
HIST1H3A				No			
HIST1H3C							Yes
HIST1H3J							Yes
HLA-DOB	Yes						
HLF							Yes
HMCN1	Yes						
HMGN5	Yes						
HNF1A	Yes			Yes	Yes	Yes	
HOOK1	Yes						
HOXA2							Yes
HOXB7							Yes
HOXD9							Yes
HRAS	No			No	No		No
HS6ST2	Yes						
HS6ST3	Yes						
HSP90B1				Yes			
HSPBP1	Yes						
HSPD1		Yes					
HSPG2		Yes					
HYDIN	Yes						
ID1					Yes		
IDH1		No					No
IFIH1	Yes						
IFRD2				Yes			
IGF1R	No			No			
IKZF3	Yes						
IL1RL1				Yes			
IL22							Yes
ILK	Yes			Yes	Yes	Yes	
INPP5D	Yes						
IPO7				Yes			
IQGAP3		Yes					
ITGA11	Yes						
ITGA4							Yes
ITGB4						Yes	
ITGB5			Yes				
ITGB6				Yes			
ITIH5				Yes			
ITPKB				Yes			
ITPKC	Yes			Yes			
ITPR1	Yes						
ITPR2	Yes				Yes	Yes	
JAG1				Yes			
JAG2					Yes		

JAK1						Yes	
JAKMIP2	Yes						
JHDM1D				Yes			
KAT6B	No						
KCND2	Yes						
KCNH2							No
KCNH5							Yes
KCTD19							Yes
KDELC2			Yes				
KDM5A						Yes	Yes
KDM6A	No	No	No	No	No		No
KDR	No						
KEAP1	Yes			Yes			
KIAA0020			Yes				
KIAA0100	Yes						
KIAA0226L	Yes						
KIAA0802						Yes	
KIAA0825				Yes			
KIAA1107			Yes				
KIAA1109							Yes
KIAA1467	Yes						
KIAA1549	Yes						
KIAA1731			Yes				
KIAA1804	No			No			
KIAA1919	No						
KIF13A	Yes						
KIF14				Yes			
KIF3B		Yes					
KIF3C				Yes			
KIRREL3	Yes						
KLHL20	Yes						
KLHL34							Yes
KLHL36				Yes			
KLK13	Yes						
KLK7	Yes			Yes	Yes	Yes	
KLK9	Yes						
KMT2A			Yes				
KNTC1	Yes						
KRAS	Yes	No		Yes	No		No
KRT19		Yes					
KRT26				Yes			
KRTAP12-3				Yes			
LAMA1				Yes			
LAMA3	Yes			Yes			
LAMB3	Yes						
LAMB4							Yes
LAMC2		Yes					
LANCL1	Yes						
LAT2				Yes			
LDHA				Yes			
LEPRE1		No					
LHCGR	Yes						
LIFR				Yes			
LIM2							No
LIMK2		Yes					
LMBRD2	Yes						
LMNB1						Yes	
LMNB2			Yes				
LNPEP				Yes			
LOC223075	Yes						
LOC442444	Yes						
LOC652153	Yes						
LRBA				Yes			Yes
LRIG1				Yes			
LRIT2				Yes			
LRP1	Yes						
LRP2	Yes			Yes			
LRP5L		Yes					

Gene							
LRRC16A	Yes						
LRRC32				Yes			Yes
LRRFIP2			Yes				
LRRIQ1	Yes						
LTK						Yes	
LTN1			Yes				
LYST				Yes			
MAF	Yes				Yes	Yes	
MAG	Yes						
MAGEA11	Yes						
MAGEB10	Yes						
MAGEB16	Yes						
MAML2	Yes			Yes			
MAML3	Yes			Yes	Yes	Yes	
MAN2A1		Yes					
MAP2K1	No			Yes	No		
MAP2K2	No			No	No		
MAP2K5						Yes	
MAP3K1				Yes	Yes	Yes	
MAP3K14	Yes			Yes	Yes	Yes	
MAP3K4				Yes			
MAP3K7					Yes		
MAP4K1						Yes	
MAP4K4	No			No			
MAPK11						Yes	
MAPKAPK3					Yes	Yes	
MAPKBP1	Yes				Yes	Yes	
MARK3							Yes
MAST4			Yes		Yes		
MATN2	Yes						
MAX	Yes			No	No	No	
MBLAC1	Yes						
MC2R	Yes						
MC3R							Yes
MCM3AP		Yes					
MCM3APAS	Yes			Yes	Yes		
MCPH1		Yes					
MECOM					Yes		
MECP2	Yes						
MED28		No					
MEF2A				Yes			
MEGF10	Yes						
MET	No			No			
METTL19		No					
METTL23			Yes	Yes			
METTL5				Yes			
METTL8	Yes						
MFN2	Yes						
MFSD10			Yes				
MGAM	Yes						
MIB1	Yes						
MKI67	Yes						
MKI67IP	Yes						
MKRN3						Yes	
MKS1			Yes				
MLL			Yes			Yes	
MLLT3	Yes						
MMP15			Yes				
MMP2				Yes			
MMP25							Yes
MOGAT3	Yes						
MON1A	Yes						
MPP2	Yes						
MRE11A			Yes				
MRGPRX2	Yes						
MRPL3				Yes			
MSH2	Yes						
MSH3	Yes						

Gene							
MST1R	No			No			
MTM1		Yes					
MTMR10				No			
MTO1			Yes				
MTOR	No			No			
MTUS1			Yes				
MTUS2	Yes					Yes	Yes
MUC17							Yes
MUC4			Yes				
MUC5AC	Yes						
MUC5B	Yes						
MVP			Yes				
MXRA5	Yes						
MYCBPAP	Yes						
MYH1					Yes		
MYH11				Yes			
MYH2	Yes						
MYH3	Yes						Yes
MYH8				Yes			
MYO18A	Yes						
MYO18B	Yes						
MYO3A						Yes	
MYOM2				Yes			
MYST4					Yes		
N4BP3	Yes						
NAA10				Yes			
NAB1	Yes						
NADK		Yes					
NAMPT	Yes						
NAT6				No			
NBEAL1	No						
NBN				Yes			
NCAM1	Yes			Yes	Yes	Yes	
NCOA3				Yes	Yes		
NCOA6				Yes			
NDC80	Yes						
NDE1	Yes						
NDUFB4				No			
NEB	Yes						
NEBL	Yes						
NEK3	Yes			Yes		Yes	
NEK6	No			No			
NES	Yes						
NEUROD6				Yes			
NF2		No		No			
NFAT5		Yes					
NFATC1	Yes						
NFE2L2	No				No		
NFIA					Yes	Yes	
NIF3L1				Yes			
NIN	Yes			Yes			
NIPBL	Yes						
NKAP		Yes					
NLRP7							Yes
NMNAT1	Yes						
NOL11				Yes			
NOL4							Yes
NOTCH1	No					No	
NOTCH2	No	Yes				No	
NOTCH3				Yes			
NOTCH4					Yes		
NP_001073948_1		Yes					
NPAT			Yes				
NPM1				Yes	Yes	Yes	
NPR1	Yes			Yes			
NR1H2	Yes			Yes	Yes	Yes	
NR2F1				Yes			
NR4A2							Yes

NRAS	No	No		No	No		No
NRSN2	Yes						
NRXN3				Yes			
NTF3							Yes
NUDT6	Yes						
NUP210	Yes			Yes			Yes
NUP98					Yes		
ODZ1	Yes						
OGFRL1				Yes			
OMG					Yes		
OR10G9				Yes			
OR1B1				Yes			
OR2A5							Yes
OR2L2				Yes			
OR2T11							Yes
OR2V2		Yes					
OR2W3	Yes			Yes			
OR4A15							Yes
OR4M2	Yes						
OR4N2	Yes						
OR51A7							Yes
OR52D1	Yes						
OR52H1	Yes						
OR56A3	Yes						
OR5B17	Yes						
OR5D13		Yes					
OR5M3	Yes						
OR8U1	No						
ORMDL2							Yes
OSBP	Yes						
OSBPL5		Yes					
OSTN				Yes			
OTOA				Yes			
OTOP1	Yes						
OTUD7B				Yes			Yes
P2RX7					Yes		
PACSIN2		Yes					
PADI1	Yes						
PAF1	Yes						
PAK4	No			No			
PAK7	Yes						
PAPSS1		Yes					
PARP1				Yes			
PAX3	Yes				Yes		
PAX5	Yes						
PBRM1		Yes					
PCBP3	Yes						
PCDH15					Yes		
PCDHA4				No			
PCDHA7							Yes
PCF11		No					
PCK1	Yes						
PCSK5	Yes						
PCSK7	Yes						
PDCD11			Yes				
PDE11A	Yes	Yes		Yes			Yes
PDE4D	Yes						
PDE4DIP				Yes	Yes		
PDGFRL	Yes						
PDK1	No			No			Yes
PDP1			Yes				
PEAR1				Yes			
PEPD			Yes				
PGBD5	Yes						
PGK2				Yes			
PGR	Yes				Yes	Yes	
PHACTR3	Yes						
PHF20L1		Yes					

PHKA1	Yes						
PHKA2	Yes						
PHLDB2							Yes
PHRF1	Yes						
PIAS2	Yes						
PICALM						Yes	
PIGS		Yes					
PIK3C2A	No			No	Yes		
PIK3C2G	Yes				Yes		
PIK3CA	No	No	No	Yes	No		No
PIP4K2A	Yes						
PKD1L2	Yes						Yes
PKHD1L1	Yes						
PLBD1	Yes						
PLCH1	Yes						
PLEC				Yes			Yes
PLEK2	Yes			Yes			
PLEKHA4	Yes						
PLEKHG6		Yes					
PLEKHH3	Yes						
PLK1	Yes						
PLOD3							Yes
PLUNC	Yes						
PMPCB	Yes						
PNCK							Yes
PNPLA5	Yes						
POFUT1				Yes			
POLE	Yes						
POLM			Yes				
POLN		Yes					
POLR1A		Yes					
POLR3B	Yes						
POM121L12							Yes
POMZP3	Yes						
POP1	Yes						
POU5F1B	Yes						
PPA1		Yes					
PPARA						Yes	
PPFIBP2	Yes						
PPP1R3A		Yes					Yes
PPP1R9A	Yes						
PPP2R1A	No						
PPP4R1							Yes
PPP5C	Yes						
PRAME							Yes
PRG2				No			
PRIM2	Yes						
PRKAA1						Yes	
PRKCG				Yes			
PRKCH	Yes						
PRKD3				Yes			
PRKDC	Yes			Yes	Yes	Yes	
PRKG2						Yes	
PRKRA			Yes				
PRLR				Yes			
PRMT2			Yes				
PRPF39			No				
PRR11							Yes
PRTG			Yes				
PRUNE2	Yes						
PSKH2				Yes			
PTCH1		No					
PTEN	No	No	No	No	Yes		
PTGFRN		Yes					
PTK2				Yes			
PTK6						Yes	
PTPN13				No			
PTPN14	Yes						

PTPN21	Yes						
PTRF				Yes			
PTX4				Yes			
PVRL3	Yes						
PWP2		Yes					
PYCRL	Yes						
Q6ZSY1_HUMAN	Yes						
Q86U89_HUMAN		Yes					
Q8N1G8_HUMAN	Yes						
Q96RI3_HUMAN	Yes						
QRFPR	Yes						
QRICH2	Yes						
R3HDML				Yes			
RAB10							Yes
RAB3GAP1	Yes						
RABGAP1		Yes					
RAD21				Yes			
RADIL							Yes
RAI16	Yes						
RALGAPA1	Yes						
RALY	No						
RALYL	Yes						
RANBP2							Yes
RAPGEF4	Yes						
RARG						Yes	
RASA1							Yes
RASGRF1	Yes				Yes		
RASSF3	Yes						
RAVER1	Yes						
RB1	No						
RBFOX1	Yes						
RBM12B	Yes						
RBM24	Yes						
RBM6		Yes					
RBPJ					Yes		
RECQL4	Yes				Yes	Yes	
REM1						Yes	
REN				Yes			
REV3L				Yes			
RGPD3	Yes						
RHAG							Yes
RHOA						Yes	
RIMBP2							Yes
RNF111				Yes			
RNF133	Yes						
RNF181				Yes			
RNF213		Yes					
RNF214				Yes			
RNF41		Yes					
RNGTT	Yes						
ROBO1		Yes		Yes			
ROBO2				Yes	Yes		
ROCK2				Yes			
ROR1	Yes						
RP1L1	Yes						
RPGRIP1L							Yes
RPL3	Yes						
RPL7L1				Yes			
RPS6KA5						Yes	
RPS6KB1	Yes						
RPS6KB2		No					
RPTOR				Yes			
RTEL1							Yes
RTF1							Yes
RTTN	Yes						
SALL3				Yes			
SAMD9L	Yes						
SBDS					Yes		

Gene							
SBF2	Yes						
SCAF4				Yes			
SCN4A	Yes						
SCN9A	Yes						
SDAD1	Yes						
SDK1	Yes						
SEMA3C				Yes			
SEPT9	Yes						
SEPX1			Yes				
SERPINB10				Yes			
SERPINI2	Yes						
SETBP1	No	No	No	No	No		No
SF3B1			No		No		Yes
SF3B3	Yes						
SG269_HUMAN	Yes						
SgK269	Yes						
SGMS2				Yes			
SH3GL2	Yes						
SH3PXD2B				Yes			
SH3RF2	Yes						
SIGLEC14							Yes
SIPA1	Yes						
SIPA1L2	Yes						
SIPA1L3				Yes			
SLC17A4				Yes			
SLC22A1							Yes
SLC2A13	Yes						
SLC30A10	Yes						
SLC30A8	Yes						
SLC35B2				Yes			
SLC35F3							Yes
SLC35F5							Yes
SLC45A2				Yes			
SLC5A8		Yes					
SLC6A8	Yes				Yes		
SLC7A6	Yes		Yes				
SLC7A8	Yes						
SLCO1B1							Yes
SLIT2	Yes						
SLITRK3							Yes
SLITRK4	Yes						
SLX4	Yes						
SMAD4					No		
SMAP1	No						
SMARCA4	Yes	No		No	Yes	No	
SMARCB1	No						
SMC6		Yes					
SMO		No					
SMOC1		Yes					
SMPDL3A		Yes					
SMURF2				Yes			
SMYD2	Yes						
SNAPC2		Yes					
SNX11			Yes				
SNX19						Yes	
SORCS1							Yes
SORL1							Yes
SOS1							Yes
SOX6		Yes					
SPACA1		Yes					
SPAG9	Yes						
SPATA13	Yes						
SPATA5	Yes						
SPEG				Yes			
SPRY2					Yes		
SPRY4				Yes			
SPTBN2	Yes						
SRC	No			No			

SRPK2							Yes
SRRM2				Yes			
ST6GAL2				Yes			
STAT1					Yes		
STAT2				Yes			
STAT5B	Yes						
STK11	Yes	No		Yes	No		
STK25						Yes	
STK31	Yes			Yes	Yes		
STK32A	Yes					Yes	
STOX1			Yes				
STXBP5	Yes						
STXBP5L	Yes						
SUFU	Yes	No					
SUGP2				Yes			
SULF1					Yes		
SUPT4H1	Yes						
SUPT6H		Yes					
SUZ12							Yes
SYDE2				Yes			
SYK						Yes	
SYNE1	Yes						Yes
SYT15			No				
TAAR1				Yes			
TAAR5	Yes						
TADA2A							Yes
TAF13	Yes						
TAOK2			Yes				
TAOK3							Yes
TAS2R1	Yes						
TAS2R3							Yes
TBC1D23	Yes						
TBL1XR1	Yes						
TCEA2	Yes						
TCEB3B	Yes						
TCF21					Yes		
TCHP	Yes						
TCTE1	Yes						
TDRD12	Yes						
TEC	Yes			Yes	Yes		
TECTA							Yes
TENM1	Yes						
TERT	No						
TFRC						Yes	
TGFB2	Yes						
TGOLN2			Yes				
TH1L				Yes			
THAP2	Yes						
THNSL2	Yes						
THOP1	Yes						
THRA				Yes			
TIAM1				Yes			Yes
TIAM2	Yes						
TINAGL1			Yes				
TJP1	Yes						
TLL1							Yes
TLL2				Yes			
TLR3						Yes	
TM2D1			Yes				
TMC2	Yes						
TMC7	Yes						
TMEFF2							Yes
TMEM123	Yes			Yes	Yes	Yes	
TMEM181		Yes					
TMEM218			Yes				
TMEM33				Yes			
TMEM41A			Yes				
TMEM57				Yes			

Gene							
TMEM63A	Yes						
TMEM99			Yes				
TNC	Yes						
TNFAIP2				Yes			
TNFAIP3				Yes			
TNK2						Yes	
TNN	Yes						
TNNI3K						Yes	
TNR	Yes						
TNRC6B	Yes				Yes		
TNS3	Yes						
TOP1MT			Yes				
TP53	No		No	No	Yes		Yes
TP63	Yes						
TPI1				Yes			
TPP1				Yes			
TPX2							Yes
TRAK1		Yes					
TRAPPC8			Yes				
TRHDE				Yes			
TRIM13				Yes			
TRIM22				Yes			
TRIM33	Yes			Yes			
TRIM60							Yes
TRIM65	Yes						
TRPM1	Yes						
TRPM2	Yes						
TRPM4			Yes				
TRPM7				Yes			
TRPS1				Yes	Yes		
TRPV5	Yes						
TRPV6							Yes
TRUB1	Yes						
TSC2						Yes	
TSEN34			Yes				
TSHZ3	Yes						Yes
TSSK2				Yes			
TTBK1	Yes					Yes	
TTC26				Yes			
TTC28	Yes						
TTN	Yes			Yes	Yes	Yes	
TYR				Yes			
UBQLNL							Yes
UCMA		Yes					
UGP2		Yes					
UNC5D	Yes						
URB1				Yes			
USH1C	Yes						
USH2A							Yes
USP10	Yes						
USP30				Yes			
USP34	Yes						
USP6	Yes						
UTP6			Yes				
VARS2	Yes						
VCL	Yes						
VEGFC	Yes			Yes	Yes	Yes	
VHL		No		No			
VPS41		Yes					
WAPAL	Yes						
WDR72				Yes			
WDR73			Yes				
WEE2				Yes		Yes	
WISP2				Yes			
WNK1						Yes	
WNT5B				Yes			
XIRP2	No						
XPO1		Yes					

XYLT2		No				
YES1					Yes	
YSK4	Yes					
YTHDF2		Yes				
ZAP70	Yes					
ZBBX				Yes		
ZBED5	Yes					
ZBTB25						Yes
ZBTB43						Yes
ZC3H11A		Yes				
ZC3H14						Yes
ZCCHC4				No		
ZCWPW2	Yes					
ZDHHC20			Yes			
ZFHX3	Yes			Yes		
ZFP62			Yes			
ZFP64				Yes		
ZFPM2	Yes					
ZFY		Yes				
ZFYVE27	Yes					
ZMIZ1						Yes
ZNF135	Yes					
ZNF14	Yes					
ZNF215	Yes					
ZNF217				Yes		
ZNF22	Yes					
ZNF251		No				
ZNF354B						Yes
ZNF419				Yes		
ZNF425	Yes					
ZNF480		Yes				
ZNF498				Yes		
ZNF507		Yes				
ZNF516	No					
ZNF559		Yes				
ZNF560				Yes		
ZNF587			Yes			
ZNF589		Yes				
ZNF638				Yes		
ZNF652				Yes		
ZNF654	Yes					
ZNF655				Yes		
ZNF729	Yes					
ZNF737	Yes					
ZNF749	Yes					
ZNF766	Yes					
ZNF774				Yes		
ZNF845			Yes			
ZPLD1				Yes		
ZSCAN22			Yes			
ZSCAN29	Yes					
ZW10		Yes				
ZZEF1						Yes

Drug	Original source	TDP2 ρ-value	TOP2A/ TDP2 ρ-value	TOP2B/ TDP2 ρ-value	TDP2 p-value	TOP2A/ TDP2 p-value	TDP2B/ TDP2 p-value
681640	COSMIC	0.385	-0.290	-0.688	0.522	0.636	0.199
17AAG	COSMIC	-0.036	-0.416	-0.228	0.954	0.486	0.712
17AAG	CCLE	0.246	-0.162	-0.231	0.690	0.795	0.709
ABT263	COSMIC	0.207	0.657	-0.447	0.738	0.229	0.450
ABT888	COSMIC	-0.508	0.934	0.513	0.382	0.020	0.376
AEW541	CCLE	0.967	-0.571	-0.749	0.007	0.314	0.145
AG014699	COSMIC	0.633	-0.378	-0.694	0.252	0.530	0.193
AICAR	COSMIC	0.345	0.600	0.111	0.570	0.284	0.859
AKT inhibitor VIII	COSMIC	-0.859	0.434	0.896	0.062	0.466	0.039
AMG706	COSMIC	0.324	0.814	-0.217	0.595	0.094	0.726
AP24534	COSMIC	0.972	-0.059	-0.782	0.006	0.925	0.118
AS601245	COSMIC	-0.004	-0.027	-0.418	0.994	0.966	0.483
ATRA	COSMIC	0.069	0.843	-0.271	0.913	0.073	0.660
AUY922	COSMIC	-0.092	0.364	-0.434	0.883	0.547	0.466
Axitinib	COSMIC	-0.113	0.984	0.062	0.856	0.002	0.921
AZD0530	CCLE	-0.659	0.205	0.225	0.227	0.741	0.716
AZD2281	COSMIC	0.508	0.526	-0.689	0.382	0.362	0.198
AZD6244	COSMIC	-0.489	0.793	0.131	0.403	0.110	0.834
AZD6482	COSMIC	-0.826	0.020	0.911	0.085	0.975	0.032
AZD7762	COSMIC	0.003	0.461	0.494	0.996	0.435	0.397
AZD8055	COSMIC	0.283	0.837	-0.032	0.645	0.077	0.959
BAY613606	COSMIC	-0.156	0.580	-0.322	0.803	0.305	0.597
Bexarotene	COSMIC	0.172	0.391	0.333	0.782	0.515	0.583
BIBW2992	COSMIC	-0.930	0.277	0.616	0.022	0.652	0.269
Bicalutamide	COSMIC	0.692	-0.061	-0.178	0.196	0.922	0.774
BID1870	COSMIC	-0.283	0.463	-0.163	0.645	0.432	0.794
BIRB0796	COSMIC	-0.065	0.541	0.568	0.917	0.346	0.318
Bleomycin	COSMIC	0.344	0.491	0.139	0.571	0.401	0.824
BMS708163	COSMIC	-0.781	-0.233	0.843	0.119	0.706	0.073
BMS754807	COSMIC	0.392	0.685	-0.033	0.514	0.202	0.958
Bosutinib	COSMIC	-0.830	0.628	0.705	0.082	0.257	0.184
Bryostatin1	COSMIC	0.397	0.733	-0.094	0.509	0.159	0.881
BX795	COSMIC	0.411	0.763	-0.213	0.491	0.134	0.731
Camptothecin	COSMIC	0.344	0.516	0.147	0.570	0.374	0.813
CCT007093	COSMIC	0.595	0.384	-0.743	0.290	0.524	0.150
CCT018159	COSMIC	-0.347	-0.480	0.417	0.567	0.413	0.485
CEP701	COSMIC	-0.067	-0.419	-0.006	0.915	0.483	0.993
CHIR99021	COSMIC	0.603	0.284	-0.383	0.282	0.643	0.524
CI1040	COSMIC	0.398	0.581	0.056	0.507	0.305	0.928
Cisplatin	COSMIC	0.348	0.480	0.132	0.566	0.413	0.832
Cytarabine	COSMIC	0.370	0.152	0.010	0.539	0.807	0.987
DMOG	COSMIC	-0.253	0.636	-0.210	0.682	0.248	0.735
Docetaxel	COSMIC	0.337	0.520	0.152	0.579	0.369	0.807
Doxorubicin	COSMIC	-0.079	0.653	-0.345	0.900	0.232	0.570
EHT1864	COSMIC	0.731	-0.520	-0.449	0.160	0.369	0.448
Elesclomol	COSMIC	-0.279	0.332	-0.143	0.649	0.585	0.819
Embelin	COSMIC	-0.149	0.413	-0.382	0.811	0.490	0.526
EpothiloneB	COSMIC	0.032	-0.407	-0.321	0.959	0.497	0.599
Etoposide	COSMIC	-0.091	0.701	-0.304	0.885	0.187	0.619
FH535	COSMIC	0.016	-0.334	-0.314	0.980	0.583	0.607
FTI277	COSMIC	-0.724	0.125	0.924	0.166	0.841	0.025
GDC0449	COSMIC	-0.662	0.807	0.543	0.224	0.098	0.345
GDC0941	COSMIC	-0.106	0.773	-0.243	0.865	0.125	0.694
Gefitinib	COSMIC	-0.192	-0.005	-0.414	0.757	0.993	0.488
Gemcitabine	COSMIC	0.294	0.743	0.002	0.631	0.150	0.998
GSK650394	COSMIC	0.360	0.373	0.090	0.551	0.536	0.885
GW441756	COSMIC	0.076	0.772	-0.052	0.903	0.126	0.934
IPA3	COSMIC	0.160	0.200	-0.679	0.797	0.747	0.208
JNJ26854165	COSMIC	-0.204	0.769	0.236	0.742	0.128	0.702
JNK9L	COSMIC	0.003	-0.343	-0.315	0.996	0.572	0.606
JNK Inhibitor VIII	COSMIC	0.738	0.180	-0.439	0.155	0.771	0.459
KU55933	COSMIC	0.662	0.517	-0.345	0.224	0.372	0.569
L685458	CCLE	-0.659	0.205	0.225	0.227	0.741	0.716

Appendix 8: Estimates and statistical significance of correlations between drug sensitivities and TDP2 protein levels, TOP2A/TDP2 protein ratios or TOP2B/TDP2 protein ratios in a panel of breast cancer cell lines

Lapatinib	CCLE	-0.431	0.516	0.484	0.468	0.374	0.408
Lenalidomide	COSMIC	-0.667	0.642	0.356	0.219	0.242	0.557
LFMA13	COSMIC	-0.197	0.744	-0.046	0.751	0.149	0.941
Metformin	COSMIC	-0.707	-0.074	0.874	0.182	0.906	0.053
Methotrexate	COSMIC	-0.120	-0.680	0.201	0.848	0.206	0.746
Midostaurin	COSMIC	0.553	0.603	-0.188	0.333	0.282	0.761
MitomycinC	COSMIC	0.238	-0.323	-0.269	0.700	0.596	0.662
MK2206	COSMIC	-0.177	0.854	-0.112	0.776	0.065	0.858
Nilotinib	COSMIC	-0.343	0.947	0.432	0.572	0.015	0.468
NSC87877	COSMIC	-0.325	0.022	0.025	0.594	0.972	0.968
NU7441	COSMIC	-0.240	0.895	-0.014	0.697	0.040	0.982
Nutlin3	CCLE	0.643	-0.995	-0.922	0.242	0.000	0.026
Nutlin3	COSMIC	0.445	0.760	-0.389	0.453	0.136	0.517
NVPBEZ235	COSMIC	-0.145	0.647	-0.293	0.815	0.238	0.632
Obatoclax Mesylate	COSMIC	0.178	0.111	-0.708	0.775	0.859	0.181
OSI906	COSMIC	0.410	0.607	-0.021	0.492	0.278	0.973
PAC1	COSMIC	-0.212	0.391	-0.340	0.732	0.516	0.575
Paclitaxel	CCLE	-0.139	-0.487	-0.214	0.824	0.405	0.729
Pazopanib	COSMIC	0.410	0.276	-0.269	0.493	0.654	0.662
PD0325901	CCLE	-0.155	0.194	0.435	0.803	0.754	0.464
PD0325901	COSMIC	-0.146	0.279	-0.311	0.815	0.649	0.610
PD0332991	COSMIC	0.050	0.901	-0.193	0.936	0.037	0.755
PD173074	COSMIC	0.307	-0.283	-0.212	0.615	0.645	0.732
PF02341066	CCLE	0.512	0.269	-0.119	0.378	0.662	0.849
PF4708671	COSMIC	0.056	-0.568	0.101	0.928	0.317	0.871
PF562271	COSMIC	0.207	0.463	0.329	0.739	0.432	0.589
PHA665752	CCLE	-0.659	0.205	0.225	0.227	0.741	0.716
PLX4720	COSMIC	0.730	0.396	-0.400	0.161	0.510	0.504
QS11	COSMIC	0.601	-0.791	-0.619	0.284	0.111	0.265
RAF265	CCLE	0.835	-0.384	-0.456	0.079	0.523	0.440
RDEA119	COSMIC	-0.514	0.563	0.012	0.375	0.323	0.984
RO3306	COSMIC	0.385	-0.123	-0.341	0.522	0.843	0.575
SB216763	COSMIC	0.010	0.873	-0.208	0.987	0.053	0.738
SB590885	COSMIC	0.283	-0.116	-0.275	0.645	0.853	0.654
Shikonin	COSMIC	0.173	0.224	-0.572	0.781	0.717	0.314
SL01011	COSMIC	0.240	0.778	0.120	0.697	0.122	0.847
Sorafenib	CCLE	-0.434	0.284	0.537	0.466	0.644	0.350
TAE684	CCLE	0.289	0.378	0.259	0.637	0.531	0.674
Temsirolimus	COSMIC	0.598	0.081	-0.371	0.287	0.897	0.539
Thapsigargin	COSMIC	-0.126	0.262	-0.394	0.840	0.670	0.511
Tipifarnib	COSMIC	0.150	0.671	-0.279	0.810	0.215	0.649
TKI258	CCLE	0.862	-0.932	-0.959	0.060	0.021	0.010
Topotecan	CCLE	0.730	-0.279	-0.292	0.161	0.650	0.634
TW37	COSMIC	0.326	0.450	0.167	0.592	0.448	0.788
Vinblastine	COSMIC	0.065	0.481	0.456	0.917	0.412	0.440
Vinorelbine	COSMIC	-0.045	-0.213	-0.323	0.942	0.730	0.595
Vorinostat	COSMIC	-0.751	0.146	0.972	0.143	0.815	0.006
VX702	COSMIC	-0.601	0.892	0.467	0.283	0.042	0.428
WO2009 093972	COSMIC	0.508	0.351	0.035	0.382	0.563	0.956
ZM447439	COSMIC	-0.165	0.968	0.041	0.791	0.007	0.948

Appendix 9: Estimates and statistical significance of correlations between drug sensitivities and TDP2 protein levels, TOP2A/TDP2 protein ratios or TOP2B/TDP2 protein ratios in a panel of lung cancer cell lines							
Drug	Original source	TDP2 ρ-value	TOP2A/ TDP2 ρ-value	TOP2B/ TDP2 ρ-value	TDP2 p-value	TOP2A/ TDP2 p-value	TDP2B/ TDP2 p-value
681640	COSMIC	-0.024	0.525	0.108	0.964	0.285	0.838
17AAG	COSMIC	0.177	0.105	0.157	0.737	0.844	0.767
ABT263	COSMIC	0.491	0.573	-0.248	0.322	0.235	0.636
ABT888	COSMIC	0.019	-0.440	-0.038	0.972	0.382	0.943
AG014699	COSMIC	-0.860	-0.324	-0.390	0.028	0.531	0.445
AICAR	COSMIC	0.508	0.645	0.353	0.303	0.167	0.492
AKT inhibitor VIII	COSMIC	0.251	0.972	0.213	0.684	0.005	0.731
AMG706	COSMIC	0.497	0.436	0.177	0.316	0.388	0.737
AP24534	COSMIC	0.006	0.857	-0.085	0.992	0.064	0.892
AS601245	COSMIC	0.358	0.607	0.949	0.554	0.277	0.014
ATRA	COSMIC	-0.059	-0.531	-0.541	0.912	0.279	0.268
AUY922	COSMIC	0.448	0.274	0.064	0.449	0.656	0.919
Axitinib	COSMIC	-0.689	-0.563	-0.437	0.130	0.244	0.386
AZD2281	COSMIC	0.369	-0.040	-0.547	0.541	0.949	0.341
AZD6244	COSMIC	-0.808	-0.015	-0.171	0.052	0.977	0.746
AZD6482	COSMIC	0.134	-0.125	-0.608	0.830	0.841	0.277
AZD7762	COSMIC	0.158	0.712	0.149	0.765	0.112	0.779
AZD8055	COSMIC	-0.511	0.188	-0.024	0.300	0.721	0.963
BAY613606	COSMIC	-0.844	-0.482	-0.232	0.072	0.411	0.707
Bexarotene	COSMIC	0.534	0.217	0.359	0.354	0.726	0.553
BIBW2992	COSMIC	0.156	0.097	-0.257	0.768	0.854	0.624
Bicalutamide	COSMIC	0.959	0.539	0.469	0.010	0.349	0.426
BID1870	COSMIC	-0.020	0.546	0.197	0.970	0.262	0.708
BIRB0796	COSMIC	0.260	0.309	-0.079	0.619	0.552	0.881
Bleomycin	COSMIC	0.351	0.953	0.186	0.563	0.012	0.765
BMS708163	COSMIC	-0.852	-0.486	-0.497	0.031	0.329	0.316
BMS754807	COSMIC	-0.639	-0.239	-0.917	0.245	0.699	0.028
Bosutinib	COSMIC	0.181	0.303	-0.126	0.732	0.560	0.811
Bryostatin1	COSMIC	-0.088	0.666	-0.121	0.888	0.220	0.846
BX795	COSMIC	-0.920	-0.220	-0.272	0.009	0.675	0.602
Camptothecin	COSMIC	-0.918	-0.313	-0.403	0.010	0.546	0.428
CCT007093	COSMIC	0.204	0.566	-0.514	0.699	0.241	0.297
CCT018159	COSMIC	-0.463	-0.235	0.012	0.355	0.653	0.981
CEP701	COSMIC	-0.931	-0.298	-0.327	0.007	0.567	0.527
CHIR99021	COSMIC	0.440	0.922	0.111	0.458	0.026	0.859
CI1040	COSMIC	-0.939	-0.371	-0.411	0.005	0.469	0.419
Cisplatin	COSMIC	0.061	-0.333	0.404	0.909	0.519	0.427
Cytarabine	COSMIC	-0.313	-0.343	-0.509	0.546	0.505	0.302
DMOG	COSMIC	0.361	0.906	0.172	0.550	0.034	0.782
Docetaxel	COSMIC	-0.602	-0.785	-0.537	0.282	0.116	0.350
Doxorubicin	COSMIC	0.376	0.679	0.107	0.533	0.208	0.863
EHT1864	COSMIC	-0.264	0.294	0.032	0.613	0.571	0.953
Elesclomol	COSMIC	0.058	-0.319	-0.416	0.913	0.538	0.412
Embelin	COSMIC	-0.325	0.259	0.760	0.594	0.674	0.136
EpothiloneB	COSMIC	0.385	0.019	-0.114	0.523	0.976	0.856
Etoposide	COSMIC	0.426	0.871	0.522	0.474	0.054	0.366
FH535	COSMIC	-0.146	0.728	0.686	0.815	0.163	0.201
FTI277	COSMIC	0.293	0.975	0.561	0.632	0.005	0.326
GDC0449	COSMIC	-0.014	0.123	-0.139	0.979	0.817	0.793
GDC0941	COSMIC	0.036	0.576	0.220	0.947	0.231	0.676
Gefitinib	COSMIC	0.296	-0.280	-0.541	0.569	0.590	0.268
Gemcitabine	COSMIC	0.442	-0.072	-0.091	0.457	0.909	0.884
GSK650394	COSMIC	0.156	-0.453	-0.230	0.802	0.444	0.710
GW441756	COSMIC	-0.555	0.370	-0.224	0.253	0.470	0.670
IPA3	COSMIC	-0.375	-0.473	0.290	0.534	0.421	0.636
JNJ26854165	COSMIC	0.543	0.619	-0.322	0.265	0.191	0.533
JNK Inhibitor VIII	COSMIC	0.512	0.646	-0.028	0.299	0.166	0.958
JNK9L	COSMIC	0.406	0.094	-0.031	0.497	0.881	0.960
KU55933	COSMIC	0.066	0.362	0.383	0.901	0.481	0.454
Lenalidomide	COSMIC	0.398	0.024	-0.453	0.435	0.964	0.366
LFMA13	COSMIC	0.809	-0.219	0.251	0.097	0.723	0.684
Metformin	COSMIC	-0.100	-0.407	-0.149	0.851	0.423	0.778
Methotrexate	COSMIC	0.707	-0.053	-0.165	0.182	0.933	0.791

Midostaurin	COSMIC	0.316	0.355	-0.365	0.605	0.558	0.546
MitomycinC	COSMIC	0.460	-0.062	0.303	0.436	0.921	0.620
MK2206	COSMIC	-0.428	0.491	-0.048	0.398	0.323	0.927
Nilotinib	COSMIC	0.310	0.054	-0.836	0.550	0.919	0.038
NSC87877	COSMIC	-0.028	-0.468	0.369	0.965	0.427	0.541
NU7441	COSMIC	-0.055	0.485	0.070	0.918	0.329	0.895
Nutlin3	COSMIC	0.167	-0.415	-0.097	0.752	0.413	0.855
NVPBEZ235	COSMIC	-0.119	0.498	0.168	0.823	0.315	0.750
Obatoclax Mesylate	COSMIC	-0.983	-0.329	-0.418	0.003	0.589	0.484
OSI906	COSMIC	-0.537	-0.454	-0.822	0.351	0.442	0.088
PAC1	COSMIC	0.491	-0.576	0.032	0.401	0.310	0.959
Pazopanib	COSMIC	0.365	-0.099	0.010	0.546	0.874	0.988
PD0325901	COSMIC	-0.335	0.384	0.083	0.517	0.453	0.875
PD0332991	COSMIC	0.079	0.578	0.257	0.881	0.229	0.623
PD173074	COSMIC	0.105	-0.021	0.544	0.842	0.968	0.265
PF4708671	COSMIC	-0.437	-0.706	0.057	0.386	0.117	0.915
PF562271	COSMIC	-0.064	0.918	0.183	0.918	0.028	0.768
PLX4720	COSMIC	0.184	0.289	0.719	0.726	0.579	0.107
QS11	COSMIC	-0.938	-0.288	-0.442	0.018	0.639	0.456
RDEA119	COSMIC	-0.924	-0.247	-0.292	0.008	0.637	0.575
RO3306	COSMIC	-0.156	-0.378	-0.484	0.768	0.460	0.331
SB216763	COSMIC	-0.430	-0.145	-0.455	0.395	0.784	0.365
SB590885	COSMIC	-0.312	0.094	-0.556	0.547	0.860	0.252
Shikonin	COSMIC	-0.068	-0.764	-0.565	0.913	0.132	0.321
SL01011	COSMIC	0.302	0.943	-0.049	0.561	0.005	0.927
Temsirolimus	COSMIC	0.057	-0.323	-0.416	0.915	0.532	0.412
Thapsigargin	COSMIC	-0.994	-0.356	-0.367	0.001	0.556	0.543
Tipifarnib	COSMIC	0.170	0.964	0.205	0.785	0.008	0.741
TW37	COSMIC	-0.873	-0.102	-0.171	0.023	0.848	0.746
Vinblastine	COSMIC	-0.889	-0.514	-0.562	0.018	0.297	0.246
Vinorelbine	COSMIC	-0.251	-0.110	-0.361	0.684	0.860	0.550
Vorinostat	COSMIC	-0.843	-0.102	-0.324	0.073	0.870	0.595
VX702	COSMIC	-0.178	-0.666	-0.300	0.735	0.149	0.564
WO200 9093972	COSMIC	-0.862	-0.203	-0.487	0.027	0.700	0.327
ZM447439	COSMIC	-0.694	-0.431	-0.328	0.126	0.394	0.526