Targeting DNA repair for personalisation of glioblastoma therapy

Dr Christina Joy Perry

MBChB (Hons), BSc, MSc, FRCR

Academic Unit of Clinical Oncology,

University of Nottingham City Hospital Campus,

Nottingham, UK



Thesis submitted to the University of Nottingham

for the degree of Doctor of Philosophy

August 2016

Abstract

Background

Glioblastomas (GBM) are the most common primary brain tumour in adults in the UK and are associated with a poor prognosis. High expression of MGMT (a direct DNA repair protein) has been shown to decrease response to temozolomide and radiotherapy in GBM. Little, however, is known of the effects of other DNA repair genes on survival and response to treatment in GBM.

PTEN is a tumour suppressor gene commonly mutated in high grade glioma. Tumours deficient in PTEN have been shown in some, but not all, studies to have defective DNA double strand break repair. PTEN deficient tumours may therefore be amenable to treatment targeting with a DNA repair inhibitor, exploiting a strategy known as synthetic lethality.

In this study the prognostic significance of DNA repair status in high grade glioma will be explored, alongside the effects of ATM inhibitor treatment in PTEN proficient and deficient glioblastoma cell lines.

Methods

Kaplan Meier survival analysis was performed for the mRNA expression levels of 248 probes (representing 157 DNA repair genes) in two adult glioblastoma datasets (Test dataset [n=191] and TCGA dataset [n=508]). After adjustment for multiple comparisons, multivariate analysis was performed, including genes significantly associated with survival on Kaplan Meier analysis. A DNA repair

gene prognostic index was generated incorporating DNA repair genes whose expression was associated with survival on multivariate analysis in both datasets. Artificial neural network analysis (ANN) was conducted in the TCGA dataset to investigate global gene interactions of the five key DNA repair genes included in the prognostic index. The prognostic significance of key genes included in the prognostic index was explored at the protein level in a cohort of 61 adult high grade gliomas (Nottingham cohort) by immunohistochemistry.

To explore synthetic lethality in PTEN proficient/deficient glioblastoma, Kaplan Meier survival analysis was performed for both PTEN and ATM mRNA (TCGA cohort) and protein expression (Nottingham cohort). Differences in DNA repair gene expression between PTEN proficient/deficient glioblastoma cell lines was assessed using the RT^2 DNA repair PCR array. Sensitivity of these cell lines to the ATM inhibitor KU55933 was assessed using a cell proliferation (MTS) assay. The functional consequences of this treatment were evaluated using the neutral comet assay and flow cytometric analysis of γH2AX accumulation, apoptosis and cell cycle progression. To further validate these findings, doxycycline-inducible PTEN knock-down glioblastoma cell lines were generated using lentiviral delivery of shRNA against PTEN.

Results

In both the Test and TCGA cohorts, a 14 DNA repair gene expression panel was associated with poor survival. On cox multivariate regression analysis *APE1*, *NBN*, *PMS2*, *MGMT* and *PTEN* were found to be independently

associated with poor prognosis. A DNA repair prognostic index incorporating *APE1*, *NBN*, *PMS2*, *MGMT* and *PTEN* stratified patients into three distinct prognostic sub-groups with worsening survival ($p<0.000001$). ANN analysis of *APE1*, *NBN*, *PMS2*, *MGMT* and *PTEN* revealed interactions with genes involved in the response to hypoxia (*HIF1AN*), fatty acid metabolism (*ACACB*, *ACSL4*), transcriptional response (*FOXG1*, *THRA*), and nucleotide metabolism (*HPRT*). At the protein level, low *PTEN* ($p=0.042$) and low *APE1* ($p=0.031$) retained a significant association with poor prognosis.

In the subgroup of GBMs with high expression of PTEN, high expression of ATM mRNA (probe 208442_s_at $p=0.03$, probe 210858_x_at $p=0.025$) and protein ($p=0.022$) was associated with worse overall survival than in patients with low ATM expression. In cell lines studies PTEN deficiency was not associated with sensitivity to an ATM inhibitor; however the PTEN proficient LN18 cell line exhibited increased sensitivity to KU55933 treatment. Higher rates of apoptosis and G1 cell cycle arrest were demonstrated in the LN18 cell line in response to 10μM KU55933. KU55933 treatment did not result in an increase in DNA damage (including double strand breaks), as assessed by the neutral comet assay and γH2AX accumulation. Knock-down of PTEN did not significantly influence sensitivity to KU55933.

Conclusion

DNA repair status clearly has prognostic significance in high grade glioma. There is, however, no evidence to support the use of ATM inhibitors in PTEN deficient GBM. Further work is required to identify why the LN18 PTEN proficient glioblastoma cell line is selectively sensitive to ATM inhibitor treatment.

Publications arising from this thesis

Perry C, Agarwal D, Abdel-Fatah TMA, Lourdusamy A, Grundy R, Auer DT, Walker D, Lakhani R, Scott IS, Chan S, Ball G, Madhusudan S.

Dissecting DNA repair in adult high grade gliomas for patient stratification in the post-genomic era.

Oncotarget 2014; 5(14): 5764-5781

Robertson E, **Perry C**, Doherty R, Madhusudan S.

Transcriptomic profiling of Forkhead Box transcription factors in adult glioblastoma mulitforme.

Cancer Genomics and Proteomics 2015; 12(3): 103-112

Other related publications

Eleftheriou M, Jimenez Pascual A, Wheldon LM, **Perry C,** Abakir A, Arora A, Johnson AD, Auer DT, Ellis IO, Madhusudan S, Ruzov A.

5-Carboxylcytosine levels are elevated in human breast cancers and gliomas.

Clinical Epigenetics 2015; 7(1): 88

Abdel-Fatah TMA, Arora A, Moseley PM, **Perry C**, Rakha EM, Green AR, Chan SYT, Ellis IO, Madhusudan S.

DNA repair prognostic index modelling reveals an essential role for base excision repair in influencing clinical outcomes in ER negative and triple negative breast cancers.

Oncotarget 2015; 6(26): 21964-78

Abdel-Fatah TMA, Middleton FK, Arora A, Moseley PM, Agarwal D, **Perry C**, Chan S, Green AR, Ball G, Ellis IO, Curtin NJ, Madhusudan S.

Untangling the ATR-CHEK1 network for prognostication, prediction and therapeutic target validation in breast cancer.

Molecular Oncology 2015; 9(3): 569-85

Abdel-Fatah TM, **Perry C**, Arora A, Thompson N, Doherty R, Moseley PM, Green AR, Chan SY, Ellis IO, Madhusudan S.

Is there a role for base excision repair in estrogen driven breast cancers?

Antioxidants & Redox Signaling 2014; 21(16): 2262-8

Abdel-Fatah TMA, Arora A, Moseley P, Coveney C, **Perry C**, Johnson K, Kent C, Hammed K, Ball G, Chan S, Madhusudan S.

Ataxia-telangiectasia mutated (ATM), ataxia-telangiectasia mutated and Rad3 related kinase (ATR) and DNA-dependent protein kinase (DNA-PK) expression correlate to adverse clinical outcomes in human epithelial ovarian cancers.

BBA Clinical 2014; 2: 10-17

Albarakati N, Abdel-Fatah TMA, Doherty R, Russell R, Agarwal D, Moseley P, **Perry C**, Arora A, Alsubhi N, Seedhouse C, Rakha EM, Green A, Ball G, Chan S, Caldas C, Ellis IO, Madhusudan S.

Targeting BRCA1-BER deficient breast cancer by ATM or DNA-PKcs blockade either alone or in combination with cisplatin for personalized therapy.

Molecular Oncology 2014; 9(1): 204-17

Abdel-Fatah TMA, Russell R, Agarwal D, Moseley P, Abayomi A, **Perry C**, Albarakati N, Ball G, Chan S, Caldas C, Ellis IO, Madhusudan S.

DNA polymerase β deficiency is linked to aggressive breast cancer: A comprehensive analysis of gene copy number, mRNA and protein expression in multiple cohorts.

Molecular Oncology 2014; 8(3): 520-32.

Abdel-Fatah TMA, **Perry C**, Moseley P, Johnson K, Arora A, Chan S, Ellis IO, Madhusudan S.

Clinicopathological significance of human apurinic/apyrimidinic endonuclease 1 (APE1) expression in oestrogen-receptor positive breast cancer.

Breast Cancer Research and Treatment 2014; 143(3): 411-421.

Abdel-Fatah TMA, Arora A, Alsubhi N, Agarwal D, Moseley PM, **Perry C**, Doherty R, Chan SYT, Green AR, Rakha, Ball G, Ellis IO, Madhusudan S.

Clinicopathological significance of ATM-Chk2 expression in sporadic breast cancers: a comprehensive analysis of large cohorts.

Neoplasia 2014; 16(11): 982-91.

Sultana R, Abdel-Fatah T, **Perry C**, Moseley P, Albarakti N, Mohan V, Seedhouse C, Chan S, Madhusudan S.

Ataxia telangiectasia mutated and Rad3 related (ATR) protein kinase inhibition is synthetically lethal in XRCC1 deficient ovarian cancer cells.

PLoS One 2013; 8(2):e57098. doi: 10.1371/journal.pone.0057098.

Abdel-Fatah TMA, Arora A, Agarwal D, Moseley PM, **Perry C**, Thompson N, Green AR, Rakha E, Chan S, Ball G, Ellis IO, Madhusudan S.

Adverse prognostic and predictive significance of low DNA-dependent protein catalytic subunit (DNA-PKcs) expression in early-stage breast cancers.

Breast Cancer Research and Treatment 2014; 146(2): 309-20

Gossage L, **Perry C**, Abbotts R, Madhusudan S.

Base excision repair factors are promising prognostic and predictive markers in cancer.

Current Molecular Pharmacology 2012; 5(1): 115-124

Abbotts R, **Perry C**, Madhusudan S.

Human apurinic/ apyrimidinic endonuclease is a novel drug target in cancer.

In: *DNA Repair and Human Health*. Rijeka, Croatia: In Tech; 2011

Perry C, Sultana R, Madhusudan S.

Personalized cancer medicine: DNA repair alterations are promising predictive biomarkers in cancer.

In: Kelley, MR ed. *DNA Repair in Cancer Therapy: Molecular Targets and Clinical Applications*. 1st ed. London: Academic Press; 2012

Other publications

Perry C, Soomro I, Kaye P, Hardy E, Parsons SL, Ragunath K, Lobo DN, Martin SG, Madhusudan S.

Analysis of lymphatic and blood vessel invasion biomarkers in T1 esophagogastric adenocarcinomas for improved patient prognostication.

Diseases of the Esophagus 2014; 28(3): 262-8

Abdel-Fatah T, **Perry C**, Dickinson P, Ball G, Moseley P, Green A, Madhusudan S, Ellis I, Chan S.

Bcl2 is an independent prognostic marker of triple negative breast cancer (TNBC) and predicts response to Anthracycline combination (ATC) chemotherapy (CT) in adjuvant and neoadjuvant settings.

Annals of Oncology 2013; 24(11): 2801-7

Prizes

Shortlisted for Leonardo Da Vinci Prize – Royal Society of Medicine, London UK, November 2014

NCRI Student Prize & poster presentation – NCRI Annual Conference, Liverpool UK, November 2014

Sue Watson Presentation Prize – University of Nottingham, May 2014

Poster presentations

Perry C, Agarwal D, Abdel-Fatah, Auer DT, Scott IS, Chan S, Ball G, Madhusudan M.

Dissecting DNA repair in adult high grade glioma for patient stratification and synthetic lethality application.

5th US-EU International Conference on Repair of Endogenous DNA damage. Santa Fe, New Mexico, USA, 2014.

Perry C, Agarwal D, Abdel-Fatah T, Lourdusamy A, Grundy R, Auer D, Walker D, Lakhani R, Scott I, Chan S, Ball G, Madhusudan S.

Dissecting DNA repair in adult high grade gliomas for patient stratification and synthetic lethality application in the post-genomic era.

NCRI Cancer Conference, Liverpool, UK, 2014.

Acknowledgements

I would like to thank my supervisors Dr Srinivasan Madhusudan and Dr Stewart Martin. Dr Madhusudan's involvement has been key in devising this project. Moreover, I wish to thank the Medical Research Council (MRC) for their financial support.

Additionally, I wish to thank Professor Graham Ball and Ms Devika Agarwal of Nottingham Trent University for their expert bioinformatics input. Further thanks go to Dr Anna Grabowska's group, and especially Mrs Pamela Collier, for her technical skills and support with the lentiviral studies. Mr Paul Moseley and Mr Chris Nolan's support with the immunohistochemistry component of this study has been invaluable, as has the expert histopathological input of Dr Tarek Abdel-Fatah.

I wish to thank all members of the Department of Academic Clinical Oncology for their encouragement and support over the years. My special thanks extend to Dr Richard McIntosh for his help in proof reading this thesis. In addition, I wish to thank past and present members of Dr Madhusudan's group for their friendship and support. My sincere thanks go to Dr Rachel Doherty for her technical and morale support during the later stages of my PhD.

Finally, I wish to thank my friends and family for their support and interest in my research over the years. I am especially grateful to my partner Dr Alasdair Miller for his patience and tireless support for the duration of my research and during the writing of this thesis.

Table of Contents

1 Background 28
1.1 Introduction 28
1.2 Central Nervous System tumours 29
1.2.1 Classification of glioma 29
1.2.2 Genetic alterations in glioblastoma 31
1.3 DNA repair 32
1.3.1 Base excision repair pathway 34
1.3.2 Direct repair 37
1.3.3 Double strand break repair 38
1.3.4 Nucleotide excision repair 43
1.3.5 Mismatch repair 45
1.4 DNA damage response 47
1.4.1 ATM is a key component of the DNA damage response 48
1.4.2 ATM signalling in response to DNA damage 50
1.4.3 ATM and the cell cycle 51
1.5 DNA repair in human cancer 54
1.5.1 Polymorphisms within DNA repair genes and cancer risk 56
1.5.2 DNA repair gene expression as a prognostic and predictive marker in cancer 57
1.5.3 Significance of DNA repair in glioma 58
1.6 Targeting DNA repair for therapy 63
1.6.1 PARP inhibitors 63
1.6.2 MGMT inhibitors 65
1.6.3 Inhibitors of the DNA damage response 65
1.6.4 Clinical trials of DNA repair inhibitors 66
1.6.5 Challenges of inhibiting DNA repair 68
1.7 Synthetic lethality 70
1.7.1 PARP inhibition and synthetic lethality 71
1.8 PTEN and cancer 73
1.8.1 Functions of the PI3K pathway 75
1.8.2 The PI3K family 79

1.8.3 Functions of PTEN 80
1.8.4 PTEN in the nucleus 81
1.8.5 PTEN and p53 85
1.8.6 PTEN and ATM 87
1.8.7 PTEN expression in glioma 87
1.8.8 PTEN and synthetic lethality 89
1.8.9 ATM inhibition as a synthetic lethality strategy in PTEN-deficient tumours 93
1.9 Conclusion 94
1.10 Aims of study 97
2 Materials and Methods 99
2.1 DNA repair gene expression profiling in adult glioblastoma 99
2.1.1 Data acquisition for the Test dataset 99
2.1.2 Data acquisition for the Validation dataset 99
2.1.3 Probeset selection 100
2.1.4 Patient demographics for the Test dataset 100
2.1.5 Patient demographics for the TCGA dataset 100
2.1.6 Survival analysis based on DNA repair gene expression level 102
2.1.7 DNA repair gene prognostic index 102
2.2 Artificial neural network (ANN) analysis 104
2.2.1 Top 200 APE1, MGMT, NBN, PMS2 and PTEN interactors .. 104
2.2.2 Gene-gene interactions for APE1, MGMT, NBN, PMS2 and PTEN 107
2.3 DNA repair gene expression profiling by immunohistochemistry .. 110
2.3.1 Immunohistochemistry technique 110
2.3.2 Primary antibodies 111
2.3.3 Antibody optimisation 112
2.3.4 Inclusion of both positive and negative control sections 113
2.3.5 Histopathological scoring of glioma sections 113
2.3.6 Patient demographics 114
2.3.7 Ethical approval 114
2.3.8 Statistical analysis of IHC data 116
2.4 Cell line studies 117

2.4.1 Cell lines and culture requirements 117
2.4.2 DNA repair inhibitors and other compounds 118
2.4.3 Cell line characterisation studies 119
2.4.4 Western blot analysis 125
2.4.5 Cell survival assays 129
2.4.6 Functional consequences of DNA repair inhibitor treatment in glioma cell lines 131
2.4.7 Production of inducible PTEN knockdown glioma cell lines by lentiviral transduction 135
3 Results 141
3.1 Introduction 141
3.1.1 Prognostic scores in high grade glioma incorporating clinical variables 142
3.1.2 IDH mutations, MGMT promoter methylation status and 1p19q deletions as prognostic and predictive markers in GBM 143
3.1.3 Prognostic scores in high grade glioma incorporating clinical and molecular variables 145
3.1.4 Prognostic scores in high grade glioma based on molecular markers 146
3.1.5 DNA repair gene prognostic and predictive scores 147
3.1.6 Array express and The Cancer Genome Atlas (TCGA) datasets 148
3.1.7 Artificial neural network analysis techniques 149
3.1.8 Study rationale 150
3.2 Results 152
3.2.1 Exploring DNA repair gene expression in the Test and TCGA datasets 152
3.2.2 Artificial neural network analysis for APE1, MGMT, NBN, PMS2 and PTEN 173
3.3 Discussion and conclusions 206
3.3.1 APE1, MGMT, NBN, PMS2 and PTEN are identified as the most important DNA repair genes for predicting prognosis in adult GBM 206
3.3.2 Various quality assurance measures help provide confidence in the results obtained 213

3.3.3 Key drivers of the APE1, MGMT, NBN, PMS2 and PTEN networks are identified 214
3.3.4 Additional DNA repair genes are identified that may play important roles in gliomagenesis 223
3.3.5 ANN inference can be used to identify recurrent themes and generate new hypotheses 226
3.3.6 Summary 227
4 Results 230
4.1 Introduction 230
4.1.1 An overview of the DNA repair proteins studied 230
4.1.2 DNA repair and DDR protein expression in glioma 233
4.1.3 Immunohistochemistry is a widely used technique to evaluate protein expression 234
4.1.4 Study rationale 237
4.2 Results 238
4.2.1 APE1, NBN, PMS2 and PTEN antibody specificity is confirmed by western blot analysis 238
4.2.2 Control sections ensure the quality of staining in maintained 239
4.2.3 Determination of the optimum concentration of antibody for IHC analysis 240
4.2.4 Older age and higher grade are significantly associated with poor overall survival in high grade glioma 246
4.2.5 APE1 protein is widely expressed in high grade glioma 247
4.2.6 APE1 protein expression is associated with nuclear PTEN and PMS2 expression 247
4.2.7 Variable NBN protein expression is present in high grade glioma 249
4.2.8 NBN protein expression is not associated with clinicopathological features 249
4.2.9 PMS2 is widely expressed in high grade glioma 251
4.2.10 PMS2 expression was not associated with clinicopathological features, except APE1 expression 251
4.2.11 Both nuclear and cytoplasmic PTEN expression patterns are present in high grade glioma 253
4.2.12 PTEN expression is associated with grade 254

4.2.13 Low APE1 and low PTEN nuclear expression are associated with poor overall survival in high grade glioma 257
4.2.14 Combination of low nuclear PTEN and low APE1 expression is associated with poor overall survival 259
4.2.15 Low nuclear PTEN expression is associated with survival in patients treated with chemotherapy 260
4.3 Discussion 263
4.3.1 Low APE1 protein expression is associated with poor overall survival in high grade glioma 263
4.3.2 Low PTEN expression is associated with poor overall survival in high grade glioma 266
4.3.3 Low PTEN expression is associated with low APE1 expression 268
4.3.4 NBN and PMS2 protein expression are not associated with overall survival in high grade glioma patients 268
4.3.5 Limitations and future work 270
4.3.6 Summary 273
5 Results 275
5.1 Introduction 275
5.1.1 Rationale for using an ATM inhibitor in PTEN deficient glioblastoma 276
5.1.2 PTEN mutation in glioma cell lines 278
5.1.3 Overcoming genetic diversity in glioma cell lines 278
5.1.4 Gene silencing techniques 279
5.1.5 Development of isogenic cell lines by lentiviral transduction 280
5.1.6 Study rationale 284
5.2 Results 285
5.2.1 An exploration of ATM expression in high grade glioma 285
5.2.2 Exploring the potential of targeting PTEN proficient and deficient glioblastoma cell lines with an ATM inhibitor 296
5.2.3 Development of PTEN knock-down LN18 and LN229 cells using lentiviral transduction 334
5.3 Discussion and conclusions 354
5.3.1 High ATM expression is associated with poor survival in PTEN expressing high grade glioma 354

5.3.2 ATM inhibitor treatment does not result in synthetic lethality in PTEN deficient glioblastoma cell lines 357
5.3.3 ATM inhibitor treatment does not result in an increase in DNA damage and double-strand breaks 360
5.3.4 ATM inhibitor treatment results in G1/S phase cell cycle arrest in the LN18 cell line 362
5.3.5 Potential explanations for the enhanced sensitivity of the LN18 cell line to ATM inhibitor treatment 363
5.3.6 Summary 371
6 General discussion and conclusions 374
6.1 General discussion 374
6.2 Summary of key findings 378
6.3 Future developments 379
6.3.1 Further evaluation of the DNA repair gene prognostic index 379
6.3.2 Construction of a glioma TMA 379
6.3.3 Further evaluation of selective genes identified from ANN 380
6.3.4 Unravelling the mechanism behind the sensitivity of the LN18 cell line to ATM inhibitor treatment 380
7 References 384
8 Appendices 415

Table of Figures

Figure 1.1. Overview of the Base Excision Repair pathway. 35
Figure 1.2. Overview of Homologous Recombination 40
Figure 1.3. Overview of the classical Non-homologous End Joining pathway 42
Figure 1.4. Overview of the Nucleotide Excision Repair pathway 44
Figure 1.5. Overview of the Mismatch Repair pathway 46
Figure 1.6. Overview of the DNA Damage Response 48
Figure 1.7. Cell cycle in mammalian cells. 52
Figure 1.8. Diagrammatic representation of the concept of synthetic lethality 70
Figure 1.9. Diagrammatic representation of the PTEN gene and its commonly reported mutations. 73
Figure 1.10. Overview of the PI3K signalling pathway 78
Figure 2.1. Summary of the methods used to develop the prognostic index and for Artificial Neural Network (ANN) analysis 109
Figure 2.2. Chemical structure of KU55933 118
Figure 3.1. DNA repair pathways represented by the 12 genes significant in both the Test and TCGA datasets. 154
Figure 3.2. Kaplan Meier survival curves for the five DNA repair genes included in the prognostic index in the Test dataset. 156
Figure 3.3. Kaplan Meier survival curves for the five DNA repair genes included in the prognostic index in the TCGA dataset. 157
Figure 3.4. Kaplan Meier survival curves for the Combined Prognostic Index (PI_3) in both the Test and TCGA datasets. 160
Figure 3.5. Graphical representation of 1, 2 and 3 year survival as predicted by the Prognostic Index score in the Test and TCGA datasets. 162
Figure 3.6. Kaplan Meier survival curves for the five DNA repair genes included in the prognostic index in radiotherapy treated patients in the TCGA dataset. 164
Figure 3.7. Kaplan Meier survival curves for the five DNA repair genes included in the prognostic index in patients not treated with radiotherapy in the TCGA dataset. 165
Figure 3.8. Kaplan Meier survival curves for the five DNA repair genes included in the prognostic index in chemotherapy treated patients in the TCGA dataset. 167
Figure 3.9. Kaplan Meier survival curves for the five DNA repair genes included in the prognostic index in patients not treated with chemotherapy in the TCGA dataset. 168
Figure 3.10. Kaplan Meier survival curves for age at diagnosis in the Test and TCGA datasets. 169
Figure 3.11. *APE1* interactome 177

Figure 3.12. Kaplan Meier survival curves for the top 3 hubs within the APE1 interactome. 179
Figure 3.13. *MGMT* interactome. 183
Figure 3.14. Kaplan Meier survival curves for the top 3 hubs within the MGMT interactome. 185
Figure 3.15. *NBN* interactome. 189
Figure 3.16. Kaplan Meier survival curves for the top 3 hubs within the NBN interactome. 191
Figure 3.17. *PMS2* interactome. 195
Figure 3.18. Kaplan Meier survival curves for the top 3 hubs within the PMS2 interactome. 197
Figure 3.19. *PTEN* interactome. 201
Figure 3.20. Kaplan Meier survival curves for the top 3 hubs within the PTEN interactome. 202
Figure 3.21. Kaplan Meier survival curves for the 5 additional DNA repair genes identified in the interaction analyses. 205
Figure 4.1. Western blots for APE1, NBN, PMS2 and PTEN. 238
Figure 4.2. Negative and positive control sections. 239
Figure 4.3. Range of APE1 antibody concentrations used for optimisation. 241
Figure 4.4. Range of NBN antibody concentrations used for optimisation. . 242
Figure 4.5. Range of PMS2 antibody concentrations used for optimisation on colon tissue. 243
Figure 4.6. Range of PMS2 antibody concentrations used for optimisation on glioma tissue. 244
Figure 4.7. Range of PTEN antibody concentrations used for optimisation . 245
Figure 4.8. Kaplan Meier survival curves for age and grade in high grade glioma patients. 246
Figure 4.9. Examples of APE1 expression in high grade glioma sections. ... 247
Figure 4.10. Examples of NBN expression in high grade glioma sections. .. 249
Figure 4.11. Examples of PMS2 expression in high grade glioma sections. 251
Figure 4.12. Examples of PTEN expression in high grade glioma. 253
Figure 4.13. Kaplan Meier survival curves for the APE1, nuclear PTEN, cytoplasmic PTEN, NBN and PMS2 protein expression in high grade glioma. 258
Figure 4.14. Kaplan Meier survival curve for the combined expression of APE1 and nuclear PTEN. 259
Figure 4.15. Kaplan Meier survival analysis for APE1, PTEN, NBN and PMS2 expression in patients who received radiotherapy. 261
Figure 4.16. Kaplan Meier survival analysis for APE1, PTEN, NBN and PMS2 expression in patients who received chemotherapy. 262
Figure 5.1. A simplified vector map of the pTRIPZ lentiviral vector. 282
Figure 5.2. Kaplan Meier survival curves for the three probes representing *ATM* expression in the TCGA dataset. 286

Figure 5.3. Kaplan Meier survival curves for the three probes representing *ATM* expression in the TCGA cohort for patients with low *PTEN* mRNA expression. 287
Figure 5.4. Kaplan Meier survival curves for the three probes representing *ATM* expression in the TCGA cohort for patients with high *PTEN* mRNA expression. 288
Figure 5.5. Range of ATM antibody concentrations used for optimisation. . 291
Figure 5.6. Examples of ATM expression in high grade glioma sections. 292
Figure 5.7. Kaplan Meier survival curves for ATM protein expression in the Nottingham Cohort. 294
Figure 5.8. Kaplan Meier survival curves for ATM protein expression in the Nottingham Cohort, in patients with low PTEN (left) and high PTEN (right) protein expression. 295
Figure 5.9. Western blot for PTEN protein expression in glioblastoma cell lines. 297
Figure 5.10. Endpoint PCR for beta actin (A) and PTEN (B) mRNA expression in glioblastoma cell lines 298
Figure 5.11. Western blot for ATM protein expression in glioblastoma cell lines. 299
Figure 5.12. Endpoint PCR for beta actin (A) and ATM (B) mRNA expression in glioblastoma cell lines. 300
Figure 5.13. MTS assay for ATM inhibitor KU55933 in glioblastoma cell lines. 303
Figure 5.14. Growth assay with 10µM KU55933 in the LN18 and U87MG glioblastoma cell lines. 304
Figure 5.15. Example Comet images for LN18 cell line. 307
Figure 5.16. Neutral comet assay in glioblastoma cell lines 6 hours after exposure to 10µM KU55933. 308
Figure 5.17. Neutral comet assay in glioblastoma cell lines 24 hours after exposure to 10µM KU55933. 308
Figure 5.18. Neutral comet assay in LN18 glioblastoma cell lines 6 and 24 hours after exposure to 100µM MMS. 309
Figure 5.19. Example of isotype control staining in the LN18 glioblastoma cell line by flow cytometry. 312
Figure 5.20. Percentage of cells staining positive for γH2AX in cells both untreated and treated with 10 µM KU55933. 313
Figure 5.21. Representative dotplots for γH2AX staining in glioblastoma cell lines untreated and treated with 10 µM KU55933. 314
Figure 5.22. Representative cell cycle plots for glioblastoma cell lines untreated and treated with 10 µM KU55933. 316
Figure 5.23. Graphical representation of the percentage of cells in each phase of cell cycle in glioblastoma cell lines both untreated and treated with 10 µM KU55933. 317

Figure 5.24. Diagrammatic representation of apoptosis. 318
Figure 5.25. Dotplots for Annexin V-FITC and propidium iodide combination flow cytometry for glioblastoma cell lines untreated and treated with 10µM KU55933. 320
Figure 5.26. Percentage of apoptotic cells, as detected by Annexin V-FITC staining, for glioblastoma cell lines untreated and treated with 10µM KU55933. 321
Figure 5.27. Example of a real-time PCR amplification plot. 324
Figure 5.28. Melting curves for the four glioblastoma cell lines, plus melting curve for the non-specific DMC1 primer in the LN18 cell line. 325
Figure 5.29. Scatterplot comparing DNA repair gene expression in U87MG (PTEN deficient) (x axis) against LN18 (PTEN proficient) (y axis)glioblastoma cell lines. 327
Figure 5.30. Scatterplot comparing DNA repair gene expression in A172 (PTEN deficient) (x axis) against LN18 (PTEN proficient) (y axis) glioblastoma cell lines. 328
Figure 5.31. Scatterplot comparing DNA repair gene expression in LN229 (PTEN proficient) (x axis) against LN18 (PTEN proficient) (y axis) glioblastoma cell lines. 329
Figure 5.32. Western blot for p53 protein expression in glioblastoma cell lines. 332
Figure 5.33. Western blot for PI3K (p85 subunit) protein expression in glioblastoma cell lines. 333
Figure 5.34. Puromycin selection curve for the LN18 cell line..... 335
Figure 5.35. Puromycin selection curve for the LN229 cell line..... 335
Figure 5.36. TurboRFP expression in transduced glioblastoma cell lines with and without doxycycline induction of shRNA expression. 337
Figure 5.37. Western blot of PTEN expression in shRNA lentiviral transduced doxycycline inducible PTEN knock-down cell lines..... 338
Figure 5.38. Quantification of PTEN protein expression from Western blot analysis in the LN18 and LN229 transduced doxycycline-inducible PTEN knock-down cell lines. 339
Figure 5.39. Western blot for ATM protein expression in the LN18-477 and LN229-477 lentiviral transduced PTEN knock-down cell lines..... 340
Figure 5.40. Endpoint PCR for beta actin, PTEN and ATM expression in LN18-477 and LN229-477 cell lines. 341
Figure 5.41. Proliferation of LN18 and LN229 non-transduced cell lines in response to doxycycline..... 342
Figure 5.42. Proliferation of LN18-477 and LN229-477 glioblastoma cell lines in response to doxycycline removal. 344
Figure 5.43. Protocol 1: MTS for LN18-477 and LN229-477 (control and PTEN knock-down) cell lines with KU55933. 345

Figure 5.44. Protocol 2: MTS for LN18-477 and LN229-477 (control and PTEN knock-down) cell lines with KU55933. 346
Figure 5.45. Scatterplot comparing DNA repair gene expression in the LN18-477 control (x axis) against the LN18-477 PTEN knock-down (y axis) glioblastoma cell lines. 348
Figure 5.46. Scatterplot comparing DNA repair gene expression in LN229-477 control (x axis) against LN229-477 PTEN knock-down (y axis) glioblastoma cell lines. 351

Table of Tables

Table 1.1. Overview of the major DNA repair pathways in humans. 34
Table 1.2. Key cell cycle proteins that interact with ATM 53
Table 1.3. DNA repair genes associated with inherited cancer syndromes..... 55
Table 1.4. Clinical trials of DNA repair inhibitors that are currently recruiting patients..... 67
Table 1.5. PTEN associated cancer syndromes. 74
Table 1.6. Summary of the structure and function of the PI3Ks 80
Table 2.1. Baseline demographic and treatment data for the TCGA dataset 101
Table 2.2. Antibodies used for IHC analysis..... 112
Table 2.3. List of tissues used for optimisation and antibody concentrations112
Table 2.4. Baseline demographic and treatment data for the Nottingham Cohort 115
Table 2.5. Composition of genomic DNA elimination mix. 121
Table 2.6. List of QuantiTect® primers used for end-point PCR 122
Table 2.7. Components of PCR Mastermix..... 122
Table 2.8. Summary of quality assurance criteria used for real-time PCR analysis 124
Table 2.9. List of antibodies used for western blot analysis..... 127
Table 2.10. List of housekeeper antibodies used for Western blot analysis.. 127
Table 3.1. Medical Research Council (MRC) prognostic index 143
Table 3.2. DNA repair genes significantly associated with survival in both datasets after BH FDR correction..... 153
Table 3.3. Multivariate analysis results for the Test and TCGA datasets. 155
Table 3.4. Formulae for the three Prognostic Indices developed in the Test and TCGA datasets..... 158
Table 3.5. The association between APE1, PMS2, NBN, MGMT and PTEN expression and age in the Test and TCGA datasets..... 170
Table 3.6. Multivariate analysis for DNA repair gene expression and clinical variables combined. 172
Table 3.7. Top 10 probes that predict APE1 expression. 175
Table 3.8. DNA repair genes present in the top 200 APE1 predicting genes. 176
Table 3.9. Function of gene hubs within the APE1 interactome. 178
Table 3.10. Top 10 probes that predict MGMT expression. 181
Table 3.11. Function of gene hubs within the MGMT interactome. 184
Table 3.12. Top 10 probes that predict NBN expression. 187
Table 3.13. DNA repair genes present in the top 200 NBN predicting genes. 188
Table 3.14. Function of gene hubs within the NBN interactome. 190
Table 3.15. Top 10 probes that predict PMS2 expression. 193

Table 3.16. DNA repair genes present in the top 200 PMS2 predicting genes. 194
Table 3.17. Function of gene hubs within the PMS2 interactome. 196
Table 3.18. Top 10 probes that predict PTEN expression. 198
Table 3.19. Function of gene hubs within the PTEN interactome..... 199
Table 3.20. Multivariate analysis for APE1, MGMT, NBN, PMS2 and PTEN expression plus the 15 gene hubs. 203
Table 4.1. Clinicopathological associations of APE1 protein expression in high grade glioma. 248
Table 4.2. Clinicopathological associations of NBN protein expression in high grade glioma. 250
Table 4.3. Clinicopathological associations of PMS2 protein expression in high grade glioma. 252
Table 4.4. Clinicopathological associations of nuclear PTEN protein expression in high grade glioma. 255
Table 4.5. Clinicopathological associations of cytoplasmic PTEN protein expression in high grade glioma. 256
Table 5.1. An overview of the function of the pTRIPZ lentiviral vector elements. 283
Table 5.2. Clinicopathological associations of *ATM* (probe 208442_s_at) mRNA expression in the TCGA Cohort. 289
Table 5.3. Clinicopathological associations of ATM expression in the Nottingham Cohort. 293
Table 5.4. DNA repair genes at least two-fold up-regulated in LN18 as compared to A172 and U87MG and LN229. 330
Table 5.5. DNA repair genes at least two-fold up-regulated in LN18-477 PTEN knock-down as compared to LN18-477 PTEN expressing control. 349
Table 5.6. DNA repair genes at least two-fold down-regulated in LN18-477 PTEN knock-down as compared to LN18-477 PTEN expressing control. 350
Table 5.7. DNA repair genes at least two-fold up-regulated in LN229-477 PTEN knock-down as compared to LN229-477 PTEN expressing control. .. 352
Table 5.8. DNA repair genes at least two-fold down-regulated in LN229-477 PTEN knockdown as compared to LN229-477 PTEN expressing control. ... 353

List of frequently used abbreviations

AKT/PKB……...Protein kinase B

AP site……....Abasic/apurinic site

APE1……Apurinic/apyrimidinic endonuclease 1

ATM……...Ataxia telangiectasia mutated

ATR……..Ataxia telangiectasia and Rad3 related

BER…….. Base excision repair

BRCA1/2……....Breast cancer 1/2, early onset

CNS……..Central nervous system

C_T…….Threshold cycle

DNA……..Deoxyribonucleic acid

DSB……...Double strand break

Gy……..Gray

GBM……...Glioblastoma (multiforme)

GGR……Global genome repair

HR…….Homologous recombination

ICL……...Inter-strand crosslink

IR……..…Ionizing radiation

Kb……Kilobase

KDa……..Kilodalton

MGMT……O-6-methylguanine-DNA methyltransferase

MMR……...Mismatch repair

MMS……..Methyl methanesulphonate

MRN……...MRE11/RAD50/NBN complex

NBN……Nibrin

NER…….Nucleotide excision repair

NHEJ……..Non-homologous end joining

PARP……...Poly (ADP-ribose) polymerase

PCR…….Polymerase chain reaction

PCV……Procarbazine/lomustine/vincristine

PI3K…….....Phosphatidylinositide 3-kinase

PMS2……...Postmeiotic Segregation Increased 2

Polβ……DNA polymerase beta

PTEN…….....Phosphatase and tensin homolog

RNA……Ribonucleic acid

Rpm…… Revolutions per minute

ShRNA……..….Small hairpin RNA

SiRNA……...Small interfering RNA

SNP……...Single nucleotide polymorphism

SSA……..Single strand annealing

SSB…….Single strand break

TCR……..Transcription coupled repair

TMZ…….……Temozolomide

UV……Ultraviolet

XRCC…….X-ray repair cross-complementing protein

WHO…….…….World Health Organisation

µg…….Microgram

µl……...Microlitre

µM……..Micromolar

CHAPTER 1

Background

1 Background

1.1 Introduction

Gliomas are the most common type of brain tumour diagnosed in patients in the United Kingdom (UK). Glioblastoma multiforme (GBM) is the most frequently diagnosed, and most aggressive, pathological subtype of glioma (CRUK, 2013). Even with surgery and intensive chemo-radiotherapy treatment regimens GBM carries a poor prognosis, with less than 50% of patients alive 18 months after diagnosis (Stupp et al., 2005). New treatment strategies are clearly needed to improve the outcome of these patients.

The functional status of DNA repair mechanisms within cells has been shown to both influence cancer development and how tumours respond to treatment. Deficient DNA repair in normal tissues has been shown to lead to tumour development, whereas over-active DNA repair in cancer cells can result in treatment resistance (Torgovnick and Schumacher, 2015). DNA repair status has been shown to be both a prognostic and predictive marker in numerous cancer types (Madhusudan and Middleton, 2005). In addition, therapeutic strategies have been employed to target DNA repair to both enhance the effects of conventional treatment and to exploit a concept known as synthetic lethality. Many tumours already have specific DNA repair defects and hence inhibiting their backup DNA repair pathways can result in selective tumour cell killing by a mechanism known as synthetic lethality (Guo et al., 2011). This may represent a new personalised treatment approach for GBM patients.

1.2 Central Nervous System tumours

Tumours of the central nervous system (CNS), unlike many other tumours, occur relatively frequently in all age groups. The incidence of brain tumours does, however, increase with age. In the United Kingdom 9400 patients were diagnosed with CNS tumours in 2011; over 50% of these patients will die from their disease.

Malignancy is relatively uncommon in children. Central nervous system lesions are, however, the second most common childhood tumour, and the most common cause of childhood cancer death in the UK. Central nervous system tumours account for 27% of all tumours diagnosed in children, 14% of all tumours in teenagers and young adults and 10% of all tumours in individuals less than 30 years of age.

Gliomas are the most common CNS tumour in both children and adults. 34% of all CNS tumours in the UK are astrocytomas, of which 80% are GBM (CRUK, 2013).

1.2.1 Classification of glioma

Gliomas arise from glial cells, which are the supporting cells of the CNS (Schwartzbaum et al., 2006b). Three main types of glioma are defined according to the World Health Organisation (WHO) classification: astrocytoma, oligodendroglioma and mixed oligoastrocytoma (Louis et al., 2007). This classification is based on the cell type of origin with astrocytoma

derived from astrocytes, oligodendroglioma derived from oligodendrocytes and oligoastrocytoma derived from a mix of the two (Louis et al., 2007).

Gliomas can be classified as low grade (WHO grade 1 or 2) or high grade (WHO grade 3 or 4) (Behin et al., 2003). Grade four glial tumours are also known GBMs. Glioblastoma multiforme can be subdivided into primary GBM (tumours which arise *de novo*) and secondary GBM (tumours which arise in low-grade gliomas). Primary GBM tend to be more common in older patients while secondary GBM are generally diagnosed in younger patients (Behin et al., 2003).

Grade two gliomas include astrocytoma, oligodendroglioma and oligoastrocytoma. The management of low grade tumours remains controversial. Prognosis is often variable, although generally better than higher grade lesions. More than 40% of individuals diagnosed with a grade two glioma are alive 10 years after diagnosis (CRUK, 2013). Management often involves a 'watch and wait' policy with treatment initiated on progression. Treatment may be in the form of surgery, radiotherapy, or surgery followed by radiotherapy. Chemotherapy is usually reserved for further disease progression or high grade transformation. Grade one gliomas include pilocystic astrocytoma and are generally considered a disease of childhood. Surgery alone is often curative for these tumours (Pouratian and Schiff, 2010).

Grade three gliomas include anaplastic astrocytoma, anaplastic oligodendroglioma and anaplastic oligoastrocytoma and are commonly treated with surgery and adjuvant radiotherapy. Chemotherapy, with agents such as temozolomide and PCV (procarbazine, lomustine and vincristine), is generally

reserved for disease progression (Stupp et al., 2010). Prognosis of grade 3 glioma is better than that of GBMs, but only 27% of patients are alive 5 years after diagnosis (CRUK, 2012).

Glioblastoma multiforme are the most aggressive form of glioma and have the worst prognosis (Schwartzbaum et al., 2006b). The current 'gold standard' treatment for patients with a GBM is surgical debulking followed by concurrent chemo-radiotherapy with temozolomide and then 6 further cycles of temozolomide. This regimen is based on the results of a trial performed by Stupp *et al*. In this study, patients with histologically confirmed GBM were treated either with the above regimen or radiotherapy (60Gy in 30 fractions) alone. The median survival benefit of combination treatment was 2.5 months (14.6 months *vs* 12.1 months), but more importantly the 2 year survival rate increased from 10.4% with radiotherapy alone to 26.5% with combined chemo-radiotherapy (Stupp et al., 2005). If compared to survival figures for other common cancers prognosis for GBM remains poor; for example nearly 90% of breast cancer patients are alive 5 years after diagnosis (CRUK, 2011).

1.2.2 Genetic alterations in glioblastoma

Glioblastomas have been shown to have a dramatically altered genome as compared with normal brain. This is due to a combination of mutations, copy number variation and altered methylation status (Maher et al., 2006, Koschny et al., 2002, Ruano et al., 2006, TCGA, 2008). Previous studies have shown that secondary GBM have a higher burden of genetic abnormalities than primary GBM. In a study by Maher *et al* an average of 38 amplifications and

63 deletions were seen in the karyotype of primary GBMs as compared with an average of 53 amplifications and 83 deletions in secondary GBMs (Maher et al., 2006). Significant overlap, however, exists between the genetic modifications seen in primary and secondary GBM with commonly amplified and gained regions identified (Fischer and Meese, 2007). At a gene level, loss of *PTEN*, *Ink4a*, *Arf* and *p53* are amongst the most common genomic aberrations, alongside epidermal growth factor receptor (*EGFR*) amplification (Parsons et al., 2008, TCGA, 2008, Verhaak et al., 2010).

1.3 DNA repair

DNA is inherently unstable and constantly under threat from endogenous agents including oxygen free radicals. DNA can also be damaged by exogenous agents, for example ionising radiation (IR) and ultraviolet (UV). In normal cells multiple mechanisms exist to repair damage to DNA (Hoeijmakers, 2001). Defective repair of DNA damage is known to increase cancer predisposition and various syndromes have been defined which are associated with specific DNA repair defects (See Section 1.5) (Croteau, 2013).

While DNA damage in healthy cells represents a significant threat to cell survival and genomic stability, the ability to induce DNA damage in tumour cells has been exploited by the use of both chemotherapy and radiation in cancer treatment. Both of these treatment modalities act by damaging DNA, under the premise that rapidly dividing cancer cells will be more sensitive to this action than normal healthy cells. Often, however, this damage is recognised and tumour cells instigate the process of DNA repair leading to

treatment failure or resistance (Croteau, 2013). The ability to selectively block repair of DNA damage in tumour cells by pharmacological inhibition represents an exciting strategy to enhance the effectiveness of cytotoxic agents and improve patient outcomes (Madhusudan and Hickson, 2005).

Multiple DNA repair pathways exist in humans, each of which is able to process a specific type of DNA lesion. The key pathways include: base excision repair (BER); nucleotide excision repair (NER); O^6-methylguanine DNA methyltransferase (MGMT) (direct repair); mismatch repair (MMR); homologous recombination (HR); and non-homologous end-joining (NHEJ). The presence of multiple repair pathways means that should one pathway fail often a second pathway is able to take over. A summary of these key DNA repair pathways is shown in Table 1.1(Fernet, 2013, Kelley, 2012).

Table 1.1. Overview of the major DNA repair pathways in humans.

Repair mechanism	Lesion repaired	Cause of lesions
Base excision repair/single strand break repair	AP sites, single-strand breaks, modified base lesions	Oxidation, alkylation, ring saturation
Nucleotide excision repair	Bulky DNA adducts, UV-induced photoproducts, intra-strand cross-links	UV, base-damaging carcinogens, crosslinking agents
Mismatch repair	Single nucleotide mismatches, insertions, deletions and small loops	Base modifying agents, errors in normal replication
Direct repair (MGMT)	O^6-methylguanine	Alkylation
Homologous recombination and Non-homologous end joining/alternative NHEJ	DNA double strand breaks	Radiation, alkylating agents, bleomycin, topoisomerase poisons, failure to repair single strand break

AP=apurinic sites; UV=ultraviolet; MGMT=O6-methylguanine DNA methyltransferase; NHEJ=non-homologous end joining.

1.3.1 Base excision repair pathway

Base excision repair is responsible for the repair of bases damaged by oxidation, alkylation or ring saturation. In addition, BER also repairs spontaneous base damage, abasic sites and single-strand breaks (SSBs) (Abbotts and Madhusudan, 2010). An overview of the BER pathway is shown in Figure 1.1.

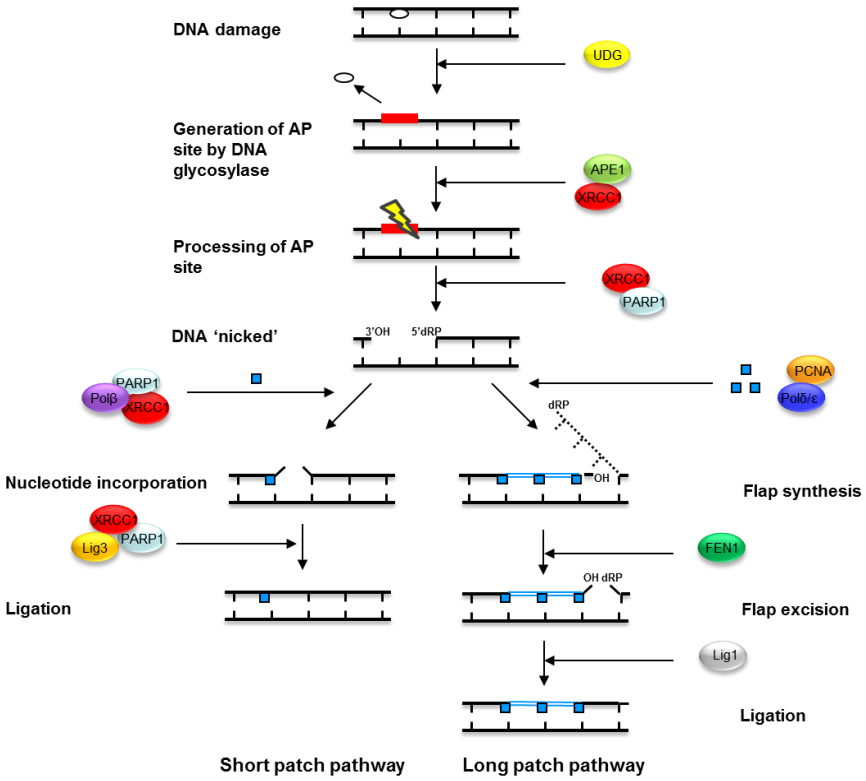


Figure 1.1. Overview of the Base Excision Repair pathway.

Firstly DNA damage is recognised by a DNA glycosylase which removes the damaged base. This creates an AP site which is normally processed by APE1, resulting in cleavage of the DNA backbone. DNA ends are then processed prior to the repair proceeding down either the short patch or long patch repair pathway. Both pathways replace the missing nucleotides by either Polβ in the short patch or Pol δ/ε in the long patch pathway. An additional flap excision step is incorporated into the long patch pathway whereby excess DNA is removed by FEN1. Finally the newly synthesised section of DNA is sealed into the DNA strand by a DNA ligase.

Base excision repair initially proceeds via a common pathway, whereby a damage specific DNA glycosylase identifies and removes the damaged base creating an abasic or apurinic (AP) site (Fishel, 2013). Apurinic sites also arise spontaneously and are reported to occur at a rate of approximately 10,000 per human cell per day (Lindahl, 1993). Irrespective of how they form, AP sites are both cytotoxic and mutagenic. An AP endonuclease cleaves AP sites, most

commonly human apurinic apyrimidinic endonuclease 1 (APE1). The phosphodiester bond immediately 5' to the AP site is hydrolysed by APE1 which results in a 3'-hydroxyl group and an abasic 5'-deoxyribose-phosphate (dRP) residue flanking the nucleotide gap (Abbotts and Madhusudan, 2010).

Repair now proceeds by either the short patch (SP-BER) or long patch (LP-BER) BER pathway. The pathway selected is dependent on a variety of factors which include: a) the type of DNA damage; b) the local availability of BER proteins; c) the stage of the cell cycle; and d) other factors. Both pathways serve to resynthesize the missing portion of DNA. The 3'-hydroxyl group acts to prime re-synthesis and the 5'-dRP must be excised before repair can complete. Short patch BER is able to replace a single nucleotide and is reliant on Polymerase beta (Polβ). The 5'-dRP is removed by the N-terminal portion of Polβ while the C-terminal region is responsible for the polymerase activity, replacing the missing nucleotide. If repair proceeds via the LP-BER pathway polymerase δ/ε adds 2-8 nucleotides. This creates a flap which is subsequently removed by Flap Structure-specific Endonuclease 1 (FEN1). Once the nucleotide gap has been filled, by either the SP-BER or LP-BER pathway, a DNA ligase joins the ends to complete the repair process. If repair has proceeded via the SP-BER pathway, DNA ligase III seals the nick, whereas DNA ligase I is responsible for this process in the LP-BER pathway (Fishel, 2013) (Abbotts and Madhusudan, 2010).

The single strand break repair pathway (SSBR) is a sub-pathway of BER and is responsible for the repair of some single strand breaks (SSBs). Poly (ADP-ribose) Polymerase 1 (PARP1) is the key damage recognition protein in this

pathway and in response to SSBs recruits end-processing proteins such as Aprataxin (APTX) and Tyrosyl-DNA Phosphodiesterase 1 (TDP1) (Caldecott, 2008).

1.3.2 Direct repair

Direct repair of DNA damage is a simple repair process, involving a single protein. DNA damaging lesions are removed in an essentially error-free process, without incision of the sugar-phosphate DNA backbone or a base excision step. Two main types of damage can be repaired in this fashion, including that caused by alkylating agents and also by UV (Eker et al., 2009).

Repair of damage caused by alkylation by MGMT is the most extensively studied pathway. MGMT removes the alkyl group from the O^6 position of guanine. The alkyl group is transferred to an internal cysteine residue on the MGMT molecule in a stoichiometric reaction. This process results in inactivation of the MGMT molecule by ubiquitination and subsequently degradation. It is due to this action of self-destruction that MGMT is commonly referred to as a 'suicide' protein. As one molecule of MGMT can only remove one alkyl group, MGMT can be depleted by an excess of DNA adducts at the O^6 position. Other DNA lesions thought to be repaired by MGMT include O^4-methylthymine, *n*-propyl-, *n*-butyl-, 2-chloroethyl, 2-hydroxyethl, *iso*-propyl and *iso*-butyl (Pegg and Byers, 1992, Pegg et al., 1995, Pegg, 2000). As alkylating agents, including temozolomide, are commonly used in the treatment of high grade glioma, it is unsurprising that MGMT expression has been shown to influence response to treatment and prognosis in

GBM (Costa et al., 2010, Hegi et al., 2005, Qiu et al., 2014, Rivera et al., 2010, Zhang et al., 2013).

1.3.3 Double strand break repair

DNA double strand breaks can occur due to a variety of causes, including ionizing radiation, alkylating agents, topoisomerase poisons and bleomycin. These DNA double strand breaks can either occur directly or indirectly. Indirect double strand breaks occur as a consequence of SSBs that are converted to DSBs during cell replication (Croteau, 2013).

Two key pathways exist to repair DNA double strand breaks; homologous recombination and NHEJ. Homologous recombination is an error-free process whereas NHEJ is error-prone. Non-homologous end-joining can also be divided into two sub-pathways; classical NHEJ and alternative NHEJ (Croteau, 2013).

1.3.3.1 Homologous recombination

Homologous recombination can only occur during the S/G2 phase of cell cycle as a homologous template is required in order to reconstruct the damaged strand of DNA (Croteau, 2013). In HR, DNA damage is first identified by the multi-protein MRN complex which consists of MRE11, RAD50 and NBN. Within this complex, MRE11 acts as a DNA binding protein and possess both endo- and exonuclease activity. RAD50 associates with MRE11 and links together two MRN complexes across the DNA strands (Moreno-Herrero et al., 2005). BRCA1 is also recruited to the damage site and activates the DNA

damage response. BRCA1, in combination with the MRN complex and a further protein called CtIP, stimulate resection of the DNA ends (Sartori et al., 2007). In turn, further proteins are recruited to the damage site including BLM, EXO1 and DNA2 which, in combination with the proteins recruited earlier, result in the formation of 3' single strand ends to the DNA. These regions of single stranded DNA are recognised by RAD51 and RPA. RAD51 is responsible for the identification of a homologous region of DNA (ideally within the sister chromatid), a process that is aided by RAD52 and RAD54 plus a group of proteins known as the RAD51 homologs. The RAD51 homologs include RAD51B, RAD51C, RAD51D, XRCC2 and XRCC3. A replacement section of DNA can now by synthesised using the sister chromatid as a template and a DNA polymerase. The process of strand invasion of the sister chromatid allows the 3' invading DNA to act as a primer for DNA synthesis while the other DNA strand is moved away forming a structure known as a 'D loop'. Repair can now proceed by two different mechanisms. A non-crossover form of repair occurs when following DNA synthesis the newly formed strand dissociates from the template strand and re-joins the original damaged DNA section. This process is known as synthesis-dependent strand annealing (SDSA). Alternatively, a more complex process can occur whereby both ends of the damaged DNA strand migrate across to the template strand resulting in the formation of a double Holliday junction (Reviewed in (Croteau, 2013)). This is shown in more detail in Figure 1.2. Holliday junctions can be resolved by a variety of mechanisms involving GEN1, MUS81 MUS81-EME1 or the BTR complex, which consists of BLM, topoisomerase 3α, RMI1 and RMI2 (Wechsler et al., 2011). Finally LIG1 completes the process by sealing the

DNA ends together (Croteau, 2013). An overview of HR is shown in Figure 1.2.

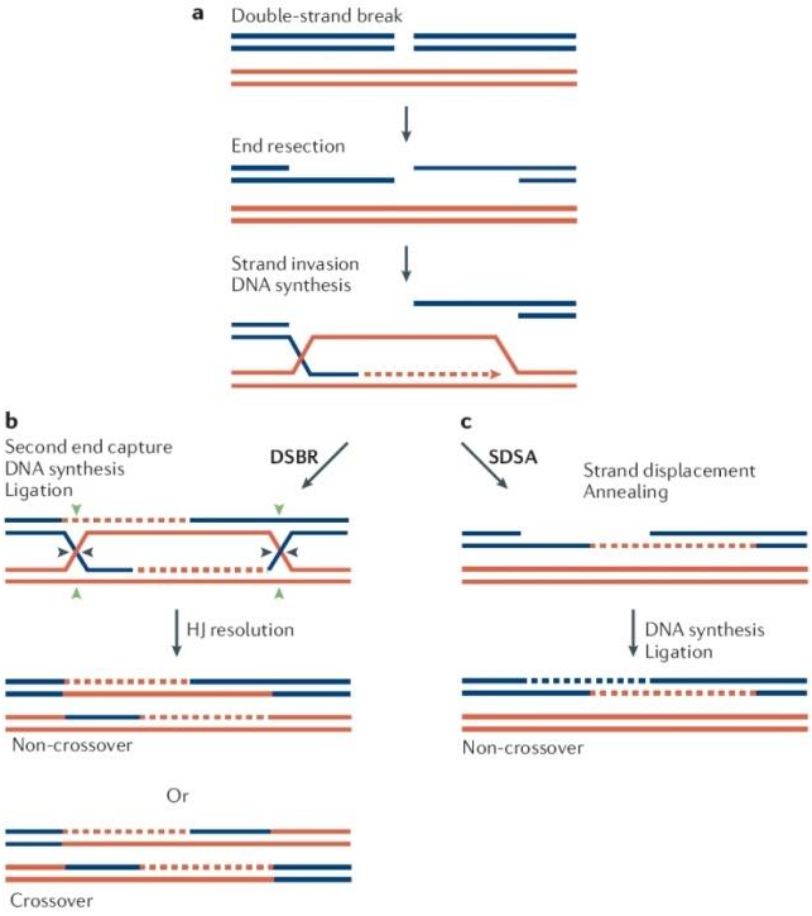


Figure 1.2. Overview of Homologous Recombination

DNA damage is first recognised by the MRN complex which subsequently recruits further proteins, including BRCA1, which are involved in the resection of DNA ends. Strand invasion and DNA synthesis then occurs prior to repair completing by either following the DSBR pathway or SDSA pathway. The DSBR pathway results in HJ formation whereby a strand of DNA from each sister chromatid crosses over to act as a template for the remaining strand of the other sister chromatid. This missing DNA can then be resynthesized and dependent on how the HJ resolves may or may not result in the cross over sections of DNA. If the SDSA pathway is followed the sister chromatid purely acts as a template and no crossover occurs. Reprinted by permission from Macmillan Publishers Ltd: Nature Reviews Molecular Cell Biology (Sung and Klein, 2006), copyright (2006). DSBR=double strand break repair; SDSA=synthesis-dependent strand annealing; HJ=Holliday junction.

1.3.3.2 Non-homologous end joining

Non-homologous end joining can occur in any stage of the cell cycle, including G1, and is able to re-join the ends of double-strand breaks in the absence of sequence homology. Despite being an error-prone pathway, NHEJ is the preferred DSB repair pathway in human cells (Croteau, 2013).

The initial step in the classical NHEJ pathway is the recognition of the DNA DSB by the Ku70/Ku80 heterodimer. This heterodimer localises to the damaged DNA and recruits DNA-PKcs, which in turn phosphorylates and recruits further proteins to the site of the damage. Often the DNA ends are incompatible for direct ligation and require further processing by either endo- or exonucleases such as FEN1, WRN, Artemis and TDP1. When end resection has occurred, a DNA polymerase will replace the missing nucleotides and finally a DNA ligase will ligate the strand back together. In the case of classical NEHJ the DNA ligation step is performed by LIG4 in complex with XRCC4 and XLF (reviewed in (Croteau, 2013, Mladenov and Iliakis, 2011)). An overview of classical NHEJ is shown in Figure 1.3.

Alternative NHEJ (Alt-NHEJ) does show some cell cycle selectivity; it mostly occurs in the G2 phase and does not occur in resting cells. This pathway typically occurs when classical NHEJ is compromised. Similar to classical NHEJ the first step in the alternative pathway involves the recognition of the DSB, which in this situation is thought to involve PARP1 and the MRN complex. During Alt-NHEJ end-processing is performed by CtIP and the MRN complex with ends resected until small regions of micro-homology (5-25 base pairs) are present between the DNA ends. After this the excess flap of

DNA is removed and DNA polymerisation is performed if necessary. Finally the DNA ends are ligated by LIG3 (reviewed in (Croteau, 2013)).

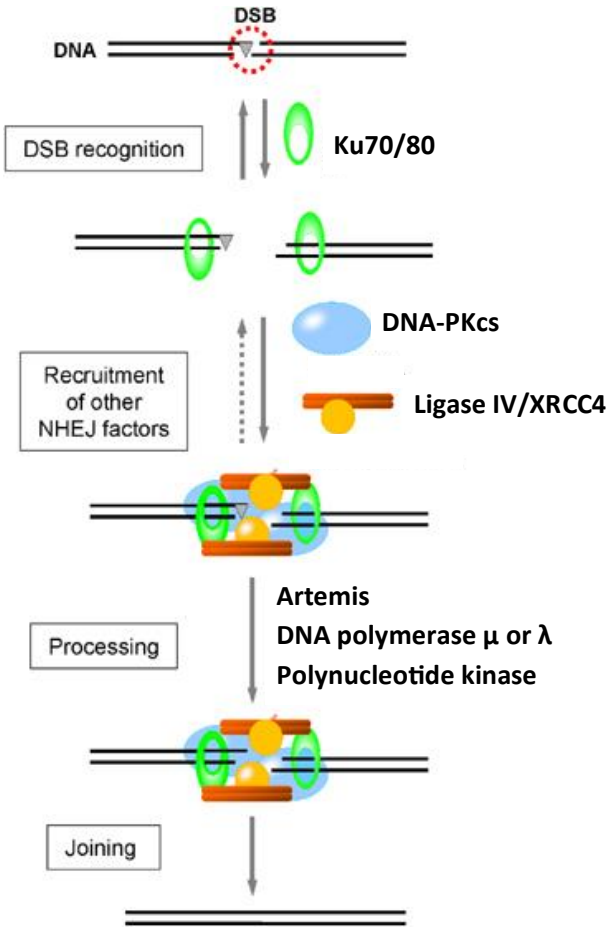


Figure 1.3. Overview of the classical Non-homologous End Joining pathway

Damage recognition is performed by the Ku70/Ku80 heterodimer and further proteins are recruited including DNA-PKcs, Ligase IV and XRCC4. The DNA ends are processed, if required, by proteins including Artemis and polynucleotide kinase. DNA synthesis follows by DNA polymerase μ/λ and finally the DNA ends are joined by DNA ligase IV. Adapted by permission from Macmillan Publishers Ltd: Oncogene (van Gent and van der Burg, 2007), copyright (2007).

1.3.4 Nucleotide excision repair

The nucleotide excision repair pathway repairs bulky DNA lesions and adducts which cause distortion of the DNA helix. Rather than identifying a specific type of DNA lesion, it is this distortion to the DNA helix that occurs as the first step in the NER process. Nucleotide excision repair is a multi-step, multi-protein process and can only occur if a single strand of DNA is damaged. Dependent on the activation step of NER the process will proceed via either the global genome repair (GGR) or transcription coupled repair (TCR) pathway (Croteau, 2013).

Global genome repair occurs when the lesion is identified within the bulk of the genome and damage recognition is performed by XPC/Cen2/HR23B or DDB1/DDB2. Transcription coupled repair occurs if the DNA damage is identified by the transcription apparatus during RNA synthesis. In the TCR pathway DNA damage is recognised by RNA polymerase II and associated factors such as CSA and CSB. After the initial recognition step, both GGR and TCR proceed via a common pathway. The next step involves the unwinding of the DNA at the damage site and incision of the damage. The transcription factor, TFIIH is recruited to the lesion and the helicases, XPB and XPD, plus XPA and RPA. Single strand incisions are performed either side of the DNA damage by ERCC1/XPF (5' end) and XPG (3' end). This generates a gap of approximately 30 nucleotides that is re-synthesised by Polymerase δ/ε with the assistance of PCNA. Finally, the DNA ends are re-joined by LIG1 or potentially Ligase 3/ERCC3 (reviewed in (Croteau, 2013, Marteijn et al., 2014,

Matsumura and Ananthaswamy, 2002). A summary of NER is shown in Figure 1.4.

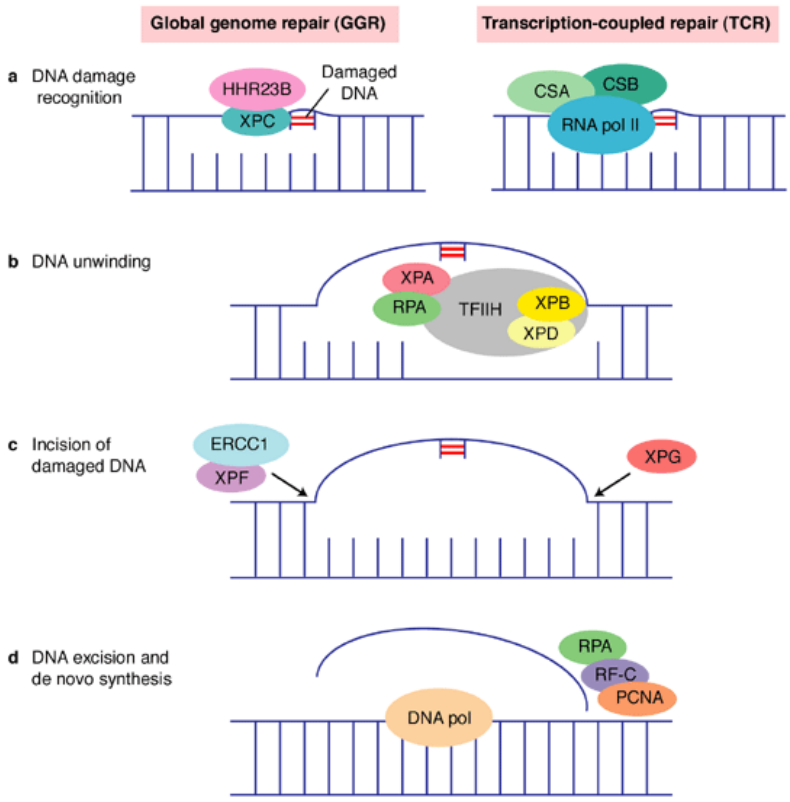


Figure 1.4. Overview of the Nucleotide Excision Repair pathway

Damage within the genome is recognised by the GGR sub-pathway and the HHR23B/XPC complex (a). Damage within the transcription apparatus, which occurs during RNA synthesis, is recognised by the TCR sub-pathway and the CSA/CSB and RNA polymerase II complex (a). After the initial recognition step, repair proceeds via a common pathway whereby DNA is first unwound (b) by the XPA/XPB/XPD/RPA/TFIIH complex. Next the region of DNA damage is incised (c) at the 5' end by ERCC1/XPF and at the 3' end by XPG. A replacement portion of DNA is resynthesized (d) by DNA polymerase δ/ε with the help of PCNA and the DNA ends re-joined by a DNA ligase. Reproduced from (Matsumura and Ananthaswamy, 2002) in accordance with the Cambridge University Press copyright regulations.

1.3.5 Mismatch repair

Single-base mismatches and misaligned short nucleotide repeats are repaired by the MMR pathway. Mismatch repair can occur before, during or after mitosis and is initiated by a variety of complexes dependent on the type of mismatch to be repaired. In MMR, DNA damage is recognised by either the MutS or MutL complexes which identify different types of DNA damage (Modrich, 2006, Hsieh and Yamane, 2008). In human cells, MutSα is the predominant MutS complex and is a heterodimer consisting of MSH2 and MSH6. MutSβ is also present and is composed of MSH2 and MSH3. MutSα is responsible for recognising base-base mismatches and base insertions or deletions of 1-2 nucleotides in length. MutSβ is able to identify insertions or deletions of 2-10 nucleotides in length. Various MutL complexes exist in humans, with the most common being MutLα. MutLα is responsible for co-ordinating MMR and consists of a heterodimer between MLH1 and PMS2 (Kadyrov et al., 2006).

Once the mismatched bases have been recognised it is essential to identify which is the newly replicated, and therefore damaged, strand. It has been shown that this can be achieved by the identification of 'nicks' within the DNA strand. Nicks are common in a newly synthesised strand of DNA due to process of DNA replication producing Okazaki fragments that then require joining together. The location of the nick in relation to the DNA damage influences how the damage is resected. For example if the nick is 3' to the damaged region the DNA can be resected directly back to the damage by the endonuclease EXO1. If, however, the nick is 5' to the damage, MLH1 is

activated and results in the recruitment of the PCNA/RFC/EXO1 complex. This then enables resection of the damaged DNA (Croteau, 2013, Genschel and Modrich, 2003). After resection, PCNA acts as a polymerase to fill the gap and the DNA ends are sealed together by LIG1 (Croteau, 2013). The MMR pathway is reviewed in Figure 1.5.

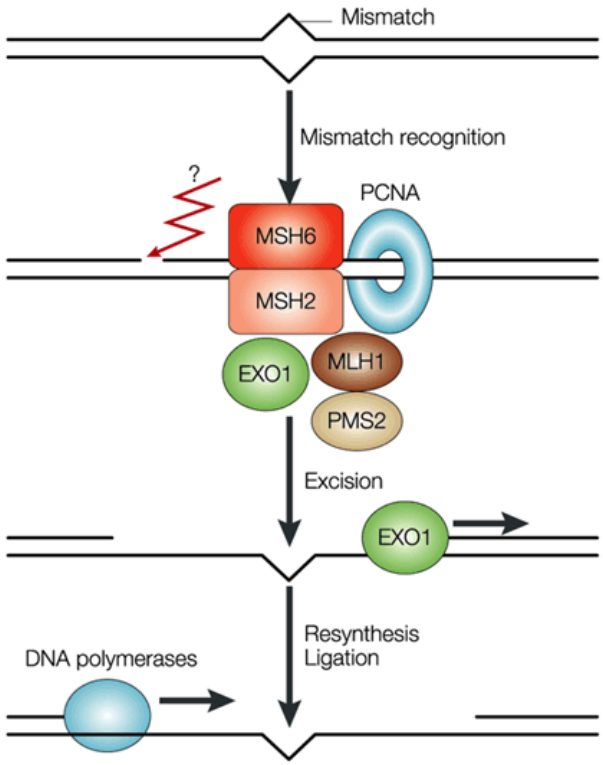


Figure 1.5. Overview of the Mismatch Repair pathway

The DNA mismatch is recognised by the MutS (MSH6/MSH2) or MutL (MLH1/PMS2) complex and the damaged strand, i.e. newly replicated strand, is identified by the presence of nicks due to Okazaki fragment formation during DNA replication. MutL co-ordinates the repair and PCNA/EXO1 are recruited to the damage site. The damaged region of DNA is now removed by EXO1, either alone or in complex with PCNA/RFC/MLH1. Finally the gap is filled with newly synthesised DNA by PCNA and the end joined by LIG1. Reprinted by permission from Macmillan Publishers Ltd: Nature Reviews Immunology (Martin and Scharff, 2002), copyright (2002).

1.4 DNA damage response

The ability of cells to repair damage to DNA involves both the individual DNA repair pathways but also the DNA damage response (DDR). The DNA damage response pathway can co-ordinate several response pathways including DNA repair, cell cycle arrest and apoptosis. Phosphoinositide 3-kinase-related protein kinases (PIKKs) are the major regulators of the DDR and include Ataxia telangiectasia mutated (ATM) and Ataxia telangiectasia and Rad3 related (ATR). ATM and ATR differ in the type of DNA damage they detect and their substrate specificity. ATM is traditionally thought to be recruited in response to DNA DSBs, typically caused by IR, and activates the checkpoint protein, checkpoint kinase-2 (Chk2) (reviewed in (Abraham, 2001, Weber and Ryan, 2015)). Recently, however, Khoronenkova *et al* have shown that ATM is activated by SSBs induced by both endogenous and exogenous agents. This results in G1 cell cycle delay and enables SSBs to be repaired prior to replication and thus prevents the formation of DSBs (Khoronenkova and Dianov, 2015).

ATR responds to UV-induced pyrimidine dimers and replication stress, and activates checkpoint kinase-1 (Chk1). Activation of both ATM and ATR results in the activation of p53 and phosphorylation of H2AX at serine 139 to form γH2AX (Fragkos et al., 2009). An overview of the DNA damage response is shown in Figure 1.6.

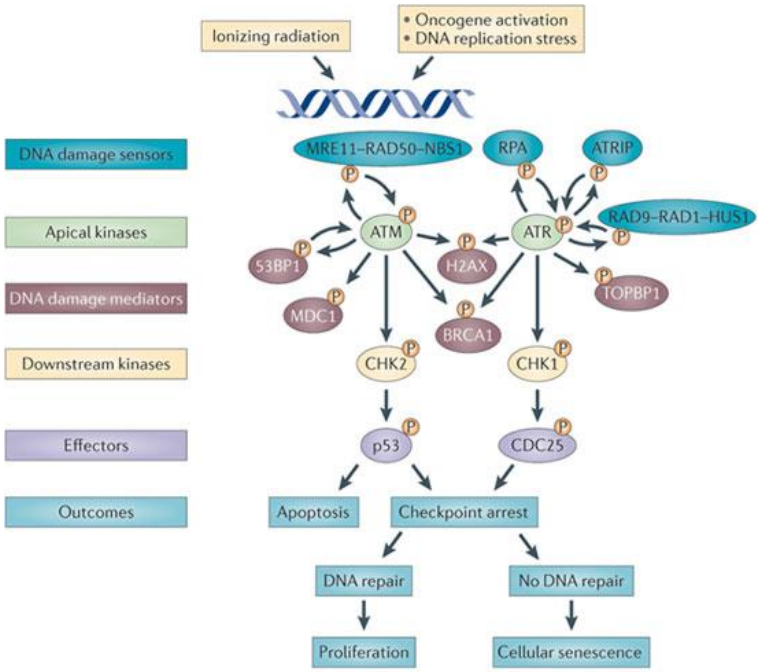


Figure 1.6. Overview of the DNA Damage Response

DNA damage is first detected by a DNA damage sensor, either the MRE11-RAD50-NBS1 complex which activates ATM or RPA/ATRIP/RAD9-RAD1-HUS1 which activates ATR. A chain of phosphorylation events is triggered and DNA damage mediators are activated. In the case of ATM these mediators are 53BP1, MDC1, H2AX and BRCA1 and for ATR are TOPBP1 and also H2AX and BRCA1. CHK2 acts as the downstream kinase for ATM and activates p53 resulting in apoptosis or checkpoint arrest. CHK1 acts as the downstream kinase for ATR and activates CDC25 which in turn can also trigger checkpoint arrest. Dependent on the outcome of checkpoint arrest cells may repair their DNA and continue to proliferate, or the DNA damage can persist and the cells enter a state of cellular senescence. Reprinted by permission from Macmillan Publishers Ltd: Nature Reviews Cancer (Sulli et al., 2012), copyright (2012).

1.4.1 ATM is a key component of the DNA damage response

The *ATM* gene is located on chromosome 11 (11q22.3) (Sanal et al., 1990, Gatti et al., 1988) and spans 150kb (66 exons) (Uziel et al., 1996). The ATM gene product is a serine/threonine protein kinase that belongs to the PIKK super family. The ATM protein is approximately 350kDa in mass and is composed of 3056 amino acids. Two domains are responsible for the kinase

activity of ATM, the Phosphatidylinositol 3/4 catalytic domain and the FRAP-ATM-TRRAP (FAT) domain found near the C-terminus (FATC) (Morgan et al., 1997).

ATM protein is constitutively expressed in cells although maintained in an inactive state in the absence of DNA damage. In response to damage to DNA, inactive ATM homodimers undergo rapid autophosphorylation at three serine residues located within the FAT domain. As a consequence active ATM monomers are released but total cellular ATM content remains unchanged (Bakkenist and Kastan, 2003). Activated ATM, alongside other PIKK family members, now can phosphorylate other downstream effectors of the DDR that are involved in DNA repair and replication, cell cycle checkpoint arrest and apoptosis. ATM therefore plays a key role in co-ordinating the DNA damage response from the initial recognition of DNA damage to activating an appropriate cellular response (Khalil, 2012).

Cells deficient in ATM are unable to recognise DNA damage. These deficient cells fail to activate the DNA damage response protein cascade and as a consequence genomic damage persists. Cells lacking ATM are unable to initiate cell cycle checkpoint activation in response to DNA damage. As a result damaged DNA is replicated until the burden of damage is so great that the cell initiates apoptosis via a p53 independent pathway (Thompson and Schild, 2002).

1.4.2 ATM signalling in response to DNA damage

In the presence of DNA damage, changes occur to chromatin structure with unwinding of the DNA super-coiled structure due to DNA strand breaks and the recruitment of RAD51. This exposes the histone variant H2AX, which is subsequently phosphorylated on Serine 139 (now referred to as γH2AX) by ATM. γH2AX now acts as a docking site for numerous components of the DDR, including the MRN complex (Carney et al., 1998).

The MRN complex (as discussed in Section 1.3.3) consists of MRE11, RAD51 and NBN and plays a pivotal role in both the repair of DSBs and DNA recombination (Carney et al., 1998, Burma et al., 2001, Hopfner et al., 2001). NBN assists the localization of the MRN complex to the site of DNA damage and binds to γH2AX. NBN also enables ATM to bind to the MRN complex and is involved in its phosphorylation and activation (Carney et al., 1998). MRE11 excises damaged DNA through its double strand exonuclease activity. RAD50 has ATPase activity and acts in combination with RAD51, RAD52 and RAD54. Together, this complex of RAD molecules acts to keep broken ends of DNA together and in addition is involved in the search for homology to enable recombination to occur (Carney et al., 1998). Another key protein involved at this stage in the DDR is BRCA1, which acts as a scaffold protein at the area of damage. The presence of BRCA1 results in the recruitment of DNA repair proteins involved down-stream in the repair process (Wang et al., 2000). This early chain of events results in further activation of ATM (Khalil, 2012).

1.4.3 ATM and the cell cycle

An overview of the cell cycle in eukaryotic cells is shown in Figure 1.7. In summary, cells cycle through four well defined phases: M Phase (mitosis), G1 (Gap 1), S phase (DNA synthesis) and G2 (Gap 2). During S phase DNA synthesis occurs and DNA within the nucleus replicates. In M phase the cell undergoes mitosis and both the nucleus and cytoplasm divide to create two new daughter cells. These two major phases are separated by Gap phases. The G1 phase signifies the transition from M phase back into S phase and allows the cell to increase in size and prepare the chromosomes for replication. The movement of the cell from S phase into M phase in separated by G2, during which the cell is able to prepare for entry into mitosis (Schafer, 1998).

Progression through the cell cycle is controlled by two families of proteins known as the cyclins and cyclin-dependent kinases (CDKs). CDKs are serine/threonine kinases while cyclins act as their regulatory partner. Various CDK/cyclin partnerships temporarily exist, and are then degraded, at different stages of the cell cycle to control movement from one phase to the next (Schafer, 1998).

CDKs are also regulated by phosphorylation status and CDK inhibitory proteins (CKIs). Phosphorylation of CDKs can either be activating, when it is as a result of the CDK activating complex (CAK), or inhibitory when initiated by Wee1 and Myt1. Dephosphorylation by the Cdc25 phosphatases results in the activation of the Cyclin-CDK complexes. Two families of CKI exist, the INK4 (INhibitors of CDK4) family and the CIP/KIP family. The INK4 family consists of $p16^{INK4a}$, $p16^{INK4b}$, $p16^{INK4c}$ and $p16^{INK4d}$ and can inhibit CDK4 and

CDK6, which are D-type cyclin-dependent kinases. The CIP/KIP family can modulate all CDKs and is made up of $p21^{CIP1/WAF1}$, $p27^{KIP1}$ and $p57^{KIP2}$ (Malumbres and Barbacid, 2005).

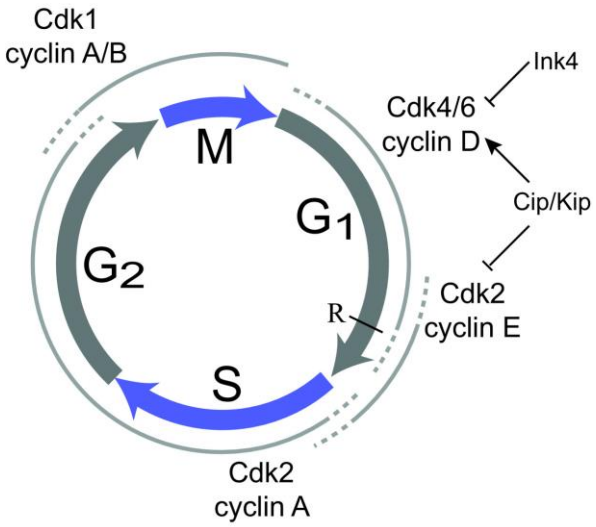


Figure 1.7. Cell cycle in mammalian cells.

The cell cycle consists of four phases and progression from one stage to the next is controlled by a series of cyclin-CDK complexes and their inhibitors and activators. Cdk4/6 and the cyclin D family are active during G1 phase of cell cycle with Cdk2 and cyclin E responsible for the transition into S phase. Cdk2 and cyclin A are activated during S phase and into G2 with Cdk1 and cyclin A/B responsible for the G2/M checkpoint. Additionally the two key families of regulators are shown, Ink4 and Cip/Kip. Modified from: (Boxem, 2006). Open access available under a Creative Commons Attribution 2.0 license.

Cell checkpoints exist to control the entry of cells into the next stage of the cell cycle. In the normal cellular environment this allows cells to proceed around the cycle in the correct order. However, at times of DNA damage, checkpoint activation allows cells time to repair the damage or, if the damage is irreparable, they can proceed via a cell death pathway (Fernet, 2013).

ATM plays a key role in cell cycle regulation and interacts with proteins involved in all stages of the cycle, including the tumour suppressor gene p53.

This means that ATM has influence over all the major cell cycle checkpoints (Table 1.2) (Khalil, 2012).

Table 1.2. Key cell cycle proteins that interact with ATM

G1	G1/S	S	G2/M
Mdm2	cAbl	BRAC1	Chk1
NBN	p53	Chk2	Chk2
p53	Rad51	CtIP	Rad17 (RFC)
		FANCD2	
		H2AX	
		MRN	
		RPA	

ATM interacts directly with the tumour suppressor gene p53. ATM phosphorylates p53 at Serine 15, resulting in p53 stabilization and activation (Banin et al., 1998). In addition, ATM phosphorylates MDM2 (a negative regulator of p53) at Serine 395, resulting in degradation of MDM2 and therefore stabilization of p53 (Gannon et al., 2012). Activation of p53 induces expression of the endogenous inhibitors of the cyclin-cdk complexes, resulting in check-point arrest in G1 (Khalil, 2012).

In addition to the phosphorylation of p53, ATM also phosphorylates the check-point kinases, Chk1 and Chk2, in response to DNA damage. Chk1 and Chk2 themselves phosphorylate p53, further contributing to its stabilization. These Serine/Threonine checkpoint kinases can subsequently phosphorylate and inactivate cdc25 phosphatases. This results in cell cycle arrest at the G2/M phases of the cell cycle due to a failure of activation of cyclin-Cdk complexes and mitosis-promoting factor (Gatei et al., 2003).

1.5 DNA repair in human cancer

Functional DNA repair is essential for the maintenance of genome integrity. It is therefore unsurprising that numerous cancer predisposition syndromes have been described in patients with defective DNA repair (Croteau, 2013). A summary of these DNA repair associated cancer predisposition syndromes is shown in Table 1.3 (Fernet, 2013). One such example is germline mutation of the *ATM* gene which results in a syndrome known as Ataxia telangiectasia (AT), or Louis-Barr syndrome (Fernet, 2013, Khalil, 2012). Ataxia telangiectasia is a rare autosomal recessive condition that is associated with both cellular and humoral immune deficiency, hypersensitivity to ionizing radiation and cerebellar atrophy. Ataxia telangiectasia is a rapidly progressive condition that presents in infancy with ataxia which manifests as the child becomes ambulant with abnormal gait and truncal movements. Telangiectasia present at a later stage, usually between two to eight years of age [Reviewed in (Lavin, 2008)]. Patients with AT are also predisposed to cancer, particularly to haematological malignancies. Of note, individuals with a heterozygous mutation of the *ATM* gene have an intermediate level of radiation sensitivity as compared with normal controls and AT homozygotes (Chen et al., 1978). In addition, AT gene carriers also have an increased risk of breast cancer (Swift et al., 1991, Renwick et al., 2006).

Table 1.3. DNA repair genes associated with inherited cancer syndromes

Gene	Pathway	Disorder	Cancer predisposition
Artemis	DSBR	Omenn syndrome	Lymphoma
ATM	DDR/DSBR	Ataxia telangiectasia	Leukaemia, lymphoma
BRCA1/BRCA2	DSBR	Hereditary breast cancer	Breast, ovarian, pancreatic
BLM	DSBR	Bloom syndrome	Various tumours, including breast
CHK2	DDR	Li-Fraumeni-like syndrome	Various tumours, including breast
FA A-P	DSBR, ICLR	Fanconi Anaemia	Leukaemia, other tumours
LIG4	DSBR	Ligase IV deficiency	Leukaemia
MLH1, MLH2, MSH2, MSH6, PMS1, PMS2	MMR	Hereditary non-polyposis colorectal cancer (HNPCC)	Colorectal, endometrial
MRE11	DSBR	AT-like disorder	Lymphoma
NBN	DSBR	Nijmegen Breakage syndrome	Lymphoma
p53	DDR	Li-Fraumeni	Various tumours, including breast
RECQL4	DSBR, BER	Rothmund-Thomson syndrome	Osteosarcoma
WRN	DSBR, BER	Werner syndrome	Various tumours
XP A-G	NER	Xeroderma Pigementosum	UV light-induced skin cancer

AT=Ataxia telangiectasia; BER=base excision repair; DSBR=double strand break repair; DDR=DNA damage response; FA=Fanconi anaemia; ICL =interstrand crosslink repair; MMR=mismatch repair; NER=nucleotide excision repair; UV=ultraviolet; XP=Xeroderma Pigmentosum.

1.5.1 Polymorphisms within DNA repair genes and cancer risk

Single nucleotide polymorphisms (SNPs) have been identified within a variety of DNA repair genes that are associated with risk of developing a range of different tumours. For example, *APE1* polymorphisms have been shown to be associated with both increased and decreased risk of a variety of tumours, including breast (Jelonek et al., 2010, Wu et al., 2012, Smith et al., 2008), bladder (Gangwar et al., 2009, Mittal et al., 2012, Wang et al., 2010a, Narter et al., 2009, Andrew et al., 2008), gastrointestinal (Jelonek et al., 2010, Canbay et al., 2011, Pardini et al., 2008, Gu et al., 2011, Zhao et al., 2011, Canbay et al., 2010) and lung cancer (Deng et al., 2011, Lu et al., 2009, Ji et al., 2011, Li et al., 2011, Agachan et al., 2009, Lo et al., 2009). Polymorphic variants of *APE1* have also been identified that may modify the risk of GBM. Zhou *et al* studied the *APE1* gene promoter -141T/G variant in 766 glioma patients and 824 healthy controls from a Chinese Han population. The results showed no association between glioma risk and the -141T/G polymorphism but on subgroup analysis demonstrated a significant association between the variant allele G and a decreased risk of GBM. The risk of GBM was reduced by 46% in patients with the homozygous GG genotype as compared to the TT genotype (Zhou et al., 2011).

ATM polymorphisms have also been found to be associated with the development of a variety of cancers including oral (Bau et al., 2010), breast (Heikkinen et al., 2005), lung (Lo et al., 2010) and prostate (Angele et al., 2004) cancer. Shen *et al* conducted a meta-analysis of 12 studies specifically looking at two *ATM* polymorphisms. They concluded that heterozygous

carriers of the rs664143 SNP had an increased risk of developing cancer, specifically lung cancer. Conversely, the rs664677 SNP was associated with a decreased risk of lung cancer but an increased risk of breast cancer (Shen et al., 2012).

Other studies have investigated a range of DNA repair polymorphisms and their association with glioma risk (Zhao et al., 2013, Adel Fahmideh et al., 2014). One such study investigated 10 SNPs in seven DSB repair pathway genes and found that *LIG4* rs1805388 and *XRCC4* rs1805377, either alone or combination, were associated glioma susceptibility. In this study *ATM* SNPs were not found to be associated with glioma risk (Zhao et al., 2013). Another study by Fahmideh *et al* performed a systematic review and meta-analysis of 27 studies covering 105 SNPs in 42 DNA repair genes. SNPs within *ERCC1* (rs3212986), *ERCC2* (rs13181), and *XRCC1* (rs25487) were found to potentially increase the risk of glioma. Decreased susceptibility to glioma was demonstrated in association with SNPs within the *PARP1* (rs1136410) and *MGMT* (rs12917) genes (Adel Fahmideh et al., 2014).

1.5.2 DNA repair gene expression as a prognostic and predictive marker in cancer

DNA repair gene expression has also been shown to be a both a prognostic and predictive marker in many tumour types. Prognostic markers attempt to assess the likely outcome for an individual with a particular condition. The level of a prognostic marker may help select patients for treatment but does not influence their response to treatment. Predictive markers, however, aim to determine the

likelihood of a patient obtaining clinical benefit from a particular intervention, including the risk of toxicity (Mehta et al., 2010).

ATM is an example of a DNA repair gene that has been shown to be both a prognostic and predictive marker. For example, mutations within the *ATM* gene have been recorded in 12% of cases of B-cell chronic lymphocytic leukaemia. Patients with a mutation in the *ATM* gene had significantly worse overall survival and treatment-free survival and were more likely to be refractory to treatment with DNA damaging chemotherapeutic agents than patients without the mutation (Austen et al., 2005). Another study found that the *ATM* T77C genotype was associated with decreased overall survival in patients with pancreatic cancer. This study also found that the *RecQ1* A159C, *RAD54L* C157T and *XRCC1* R194W genotypes were all associated with poor overall survival and that individuals with three or more of these at risk alleles had a significantly worse survival than those with only one high risk allele (Li et al., 2006). Grabsch *et al* found that both ATM and BRCA1 protein expression were associated with survival in colorectal cancer (Grabsch et al., 2006). A further study in head and neck cancer has found that the *ATM* promoter is hypermethylated in 25% of patients and that this hypermethylation is associated with a younger age at first diagnosis and decreased overall survival (Ai et al., 2004).

1.5.3 Significance of DNA repair in glioma

Temozolomide is an example of an alkylating agent used in the treatment of many tumours, including high grade gliomas (Stupp et al., 2005).

Temozolomide acts by forming 3 main types of methyl adduct, which include N7-methylguanine (the most common, ~70%), N3-methyladenine and O6-methylguanine (O6-MeG). Resistance to temozolomide is a well-recognised problem in cancer treatment and has traditionally been attributed to the formation of O6-MeG adducts, even though these only account for ~9% of all methyl adducts formed (Tentori and Graziani, 2002). MGMT is responsible for the repair of O6-MeG lesions and resistance to temozolomide is associated with increased MGMT levels, in addition to decreased MMR capacity (Zhang et al., 2010). Base excision repair is responsible for the repair of N-methylated bases which account for >80% of the DNA lesions induced by temozolomide and therefore has been shown to play a key role in temozolomide resistance (Trivedi et al., 2005). In addition, the alkylating agent procarbazine is commonly used in the treatment of high grade glioma (Kappelle et al., 2001) and also produces lesions that are repaired by the BER pathway (Pletsa et al., 1997).

Radiotherapy is widely used in the treatment of high grade glioma and induces a variety of DNA lesions by either direct or indirect interactions with genomic DNA. Lesions commonly seen include SSBs, damaged bases, abasic sites, and DSBs and therefore both BER (Vascotto, 2012) and HR (Willers, 2012) are key pathways involved in the repair of radiation induced DNA damage.

1.5.3.1 DNA repair gene expression is a predictive marker in glioma

As both direct repair and BER as key pathways involved in the repair of damage induced by temozolomide, it is not surprising that both MGMT status

and expression of components of the BER pathway are associated with survival in glioma. MGMT status in particular has been extensively studied in glioma.

1.5.3.1.1 MGMT methylation status as an indicator of prognosis and response to treatment in GBM

MGMT promoter methylation has been found to be present in just under half of all GBM patients (Hegi et al., 2005). Methylation of the *MGMT* promoter has been shown to be associated with improved outcome in GBM and also to predict sensitivity to treatment. An early study by Heigi *et al* found that *MGMT* gene silencing by methylation resulted in a significant survival benefit in patients treated with temozolomide as compared to patients with an unmethylated promoter. In this study patients with a methylated *MGMT* promoter who were treated with temozolomide plus radiotherapy had a median survival of 21.7 months as compared to a median survival of 15.3 months in patients treated with radiotherapy alone. If the *MGMT* promoter was not methylated the addition of temozolomide to radiotherapy did not improve survival (Hegi et al., 2005). These findings were confirmed in a meta-analysis by Zhang *et al* who found that *MGMT* promoter methylation was associated with both improved progression-free and overall survival in GBM patients treated with alkylating agent chemotherapy (Zhang et al., 2013). Later studies have found that *MGMT* methylation status may also influence response to radiotherapy. Rivera *et al* investigated the effect of *MGMT* methylation status on response to radiotherapy alone in patients that received post-operative radiotherapy. In this study unmethylated tumours had a worse outcome than

tumours with a methylated *MGMT* promoter and were twice as likely to progress during radiotherapy treatment (Rivera et al., 2010).

1.5.3.1.2 APE1 expression as a predictive marker in glioma

APE1 is a key component of the BER pathway and functions as an AP endonuclease. Bobola *et al* investigated the association between AP endonuclease activity and time to tumour progression (TTP) after alkylating agent chemotherapy in patients with glioma. In this study AP endonuclease activity prior to alkylating agent chemotherapy was shown to be inversely correlated with TTP after treatment in grade III gliomas. Similar results were seen in a cohort of Grade 2 and 3 gliomas treated with radiotherapy. Interestingly, no association was seen between AP endonuclease activity and TTP in GBMs after treatment with alkylating agent based chemotherapy or radiotherapy (Bobola et al., 2004).

1.5.3.2 DNA repair in glioma stem cells

Cancer stem cells are tumour-initiating cells. Human GBM have been found to be composed mostly of tumour cells but also a smaller proportion of glioma cancer stem cells. These glioma cancer stem cells appear similar to normal stem cells and express CD133, as well as other genes typical of neural stem cells. While glioma stem cells have a low proliferation rate they are capable of self-renewal and possess a high tumorigenic potential (Altaner, 2008).

$CD133^+$ glioma stem cells have been shown to be more radio-resistant than $CD133^-$ glioma cells, and hence the proportion of $CD133^+$ expressing cells in a

tumour increases after radiation. Preferential activation of the DNA damage checkpoint response has been demonstrated in glioma stem cells in response to radiation, with increased levels of activating phosphorylation of Chk1, Chk2, Rad17 and ATM. DNA damage has also been shown to be repaired more rapidly in $CD133^+$ populations than $CD133^-$ populations, as evidenced by a more rapid resolution of both comet tails and gamma H2AX foci. Targeting the DNA damage checkpoint response with an inhibitor of Chk1 and Chk2 has been shown to reverse the radioresistance seen in $CD133^+$ glioma stem cells (Bao et al., 2006).

1.6 Targeting DNA repair for therapy

As discussed earlier, defects within DNA repair pathways are associated with an increased risk of developing cancer. DNA repair status, however, is also important in modifying response to conventional cancer treatments such as chemotherapy and radiotherapy. These treatment agents act by damaging DNA and therefore intact DNA repair has the potential to confer treatment resistance (Helleday et al., 2008). Over recent years, studies have investigated the potential of inhibiting specific DNA repair pathways to enhance the effects of both chemotherapy and radiotherapy (Helleday et al., 2008). Probably the most well know class of DNA repair inhibitors are the PARP inhibitors and the story of their development will be discussed in more detail (Davar et al., 2012). Another class of inhibitor of particular relevance to glioma treatment is the MGMT inhibitors. Currently, inhibitors of the DDR are undergoing extensive investigation, including inhibitors of ATM which will be discussed in more detail (Khalil, 2012, Weber and Ryan, 2015).

1.6.1 PARP inhibitors

The discovery and development of PARP inhibitors has led the way in targeting specific DNA repair pathways for therapeutic gain. PARP1 is the most widely studied member of the PARP family and is essential for SSB repair, a subpathway of BER. PARP1 is also involved in other cellular functions, including maintaining telomere length, and has roles in other DNA repair processes (Curtin, 2012).

In terms of its role in SSB repair, PARP1 binds to DNA in the presence of DNA breaks and undergoes activation. Using nicotinamide adenine dinucleotide (NAD+) as a substrate, PARP enzymes catalyse the addition of large branched chains of poly(ADP)-ribose on to several proteins, including PARP1 itself. The presence of a large negatively charged polymer results in relaxation of the surrounding chromatin. This enables SSB repair enzymes to access the DNA break, including XRCC1 and Polβ, and aids PARP1 dissociation (Underhill et al., 2011). PARP inhibitors therefore block the SSB repair pathway at an early stage resulting in the accumulation of SSBs which can lead to double strand breaks and replication fork collapse (Bryant et al., 2005, Farmer et al., 2005).

1.6.1.1 PARP inhibition and sensitivity to DNA damaging agents

Early pre-clinical studies with PARP inhibitors, including PD128763 and NU1025, demonstrated the ability of these drugs to potentiate the effects of temozolomide and 3-methyl-(triazen-1-yl)imidazole-4-carboxamide (MTIC). More potent inhibitors of PARP were subsequently developed and demonstrated substantial enhancement of the cytotoxic effects of the alkylating agent methyl methanesulphonate (MMS). In mouse models of human GBM, GPI 15427 in combination with temozolomide was found to increase survival (Tentori et al., 2003). Similarly, PARP inhibitors have been found to potentiate the effects of radiation. For example, the PARP inhibitor E7016 was shown to enhance growth delay in a glioma mouse model when given in combination with temozolomide and radiation (Russo et al., 2009).

1.6.2 MGMT inhibitors

Lomeguatrib (6-(4-bromo-2-thienyl) methoxy purin-2-amine) is a substrate inactivator of MGMT that is administered orally. In *in vitro* studies, Lomeguatrib was shown to successfully deplete cellular levels of MGMT and to increase sensitivity to temozolomide (Barvaux et al., 2004). Phase I and II clinical trials followed (Ranson et al., 2006, Kefford et al., 2009, Khan et al., 2008, Watson et al., 2009, Ranson et al., 2007), however difficulties in selecting an appropriate dose arose; adding in Lomeguatrib reduced the dose of temozolomide that could be administered due to toxicity in the form of myelosuppression. If, however, Lomeguatrib was given for a shorter period of time prior to administration of temozolomide no improvement in response was demonstrated due to the recovery of MGMT levels. Selecting an appropriate regimen for Lomeguatrib is clearly challenging and hence there are currently no open trials of MGMT inhibitors registered at ClinicalTrials.gov (NIH, 2015).

1.6.3 Inhibitors of the DNA damage response

Recently a range of inhibitors have been developed that target the DNA damage response. Several specific inhibitors of ATM have been developed. The first potent and specific inhibitor of ATM was developed by Hickson *et al* and is known as KU55933 (2-morpholin-4-yl-6-thianthren-1-yl-pyran-4-one). KU55933 was discovered by screening a small molecule compound library that had been developed for the phosphatidylinositol 3'-kinase-like kinase family. KU55933 was found to inhibit ATM with a ki of 2.2 nmol/L and an IC_{50} of 13

nmol/L. When combined with ionising radiation (IR) or DSB inducing chemotherapeutic drugs, KU55933 was found to enhance the cytotoxic effects of these agents. Inhibition of ATM with KU55933 in combination with IR also resulted in a loss of cell cycle arrest as would normally be seen with IR alone. Neither radio-potentiation nor cell cycle arrest were seen when KU55933 was administered to ataxia telangiectasia cells. The mechanistic effects of KU55933 were demonstrated by the loss of IR-dependent phosphorylation of key downstream targets (Hickson et al., 2004).

Since the discovery of KU55933 in 2004, other specific inhibitors of ATM have been described. One such example is KU60019, an analogue of KU55933, which has an IC_{50} of 6.3 nmol/L. KU60019 has been shown to effectively radio-sensitise human glioma cell lines and to be effective at blocking the phosphorylation of downstream targets of ATM (Golding et al., 2009). CP466722 is a reversible and specific inhibitor of ATM. CP466722 has been shown to sensitize tumour cell lines to IR and induce characteristic cell cycle checkpoint defects. Removal of CP466722 rapidly reversed the inhibition of ATM activity but radio-sensitisation was seen to persist (Rainey et al., 2008).

1.6.4 Clinical trials of DNA repair inhibitors

A number of clinical trials are currently recruiting that involve the administration of DNA repair or cell cycle inhibitors either alone or in combination with other anti-cancer treatments. A summary of the clinical trials listed at www.clinicaltrials.gov (NIH, 2015) is shown in Table 1.4.

Table 1.4. Clinical trials of DNA repair inhibitors that are currently recruiting patients

Drug	DNA repair target	Tumour site	Comments
AZD6738	ATR	Advanced solid tumours	Plus chemotherapy and/or novel anti-cancer agents
AZD6738	ATR	Advanced solid tumours	Plus palliative radiotherapy
BMN 673	PARP	Advanced solid tumours	Plus deleterious BRCA mutations
Folinic acid	MGMT	GBM and non-methylated *MGMT* promoter	Plus temozolomide
LY2606368	Chk1/2	Ovarian and breast cancer	BRCA1/2 mutation or non-high risk TNBC or high-grade serous ovarian cancer at low genetic risk
Olaparib	PARP	Advanced non-small cell lung cancer	
Olaparib	PARP	Head and neck cancer	Plus radiotherapy
Olaparib	PARP	Relapsed GBM	Plus temozolomide
Olaparib	PARP	Ewings sarcoma	Plus temozolomide
Talazoparib	PARP	Children with refractory or recurrent solid tumours	With and without temozolomide
Talazoparib	PARP	BRCA1/2 mutated ovarian cancer	Received prior PARP inhibitor treatment
Veliparib	PARP	Locally advanced, unresectable pancreatic cancer	Plus gemcitabine and IMRT
VX-970	ATR	Small cell lung cancer	Plus topotecan

GBM=glioblastoma multiforme; IMRT=intensity modulated radiotherapy; TNBC=triple negative breast cancer.

Three different PARP inhibitors are currently under investigation as is the ATR inhibitor AZD6738. There are currently no clinical trials listed with ATM inhibitors. Two clinical trials that are specifically focused on GBM are currently recruiting patients. One study is investigating the combination of the PARP inhibitor Olaparib with temozolomide in patients with relapsed glioblastoma (NCT01390571). The second glioblastoma clinical trial is investigating the ability of folinic acid to modulate the *MGMT* gene in patients treated with temozolomide (NCT01700569) (NIH, 2015).

1.6.5 Challenges of inhibiting DNA repair

PARP inhibitors are the most widely studied class of DNA repair inhibitors and therefore these agents have paved the way in the identification of potential problems of DNA repair inhibition. Two key problems that have been highlighted are treatment toxicity and treatment resistance. Myelosuppression has been observed not infrequently when PARP inhibitors are used in combination with standard chemotherapy (Curtin, 2013). This has also been identified as a problem with the MGMT inhibitor Lomeguatrib, as discussed in Section 1.6.2. The combination of a DNA repair inhibitor with chemotherapeutic agents has proved challenging as this involves finding the optimal dose and timing for two agents to both increase therapeutic benefit and minimise toxicity.

Also, while PARP inhibitors have shown promise in tumours that are difficult to treat, disease control is often limited to less than 12 months. Several potential resistance mechanisms have been proposed which involve the up-

regulation of P-glycoprotein efflux pumps (Rottenberg et al., 2008), and mutational restoration of BRCA function (Edwards et al., 2008, Ashworth, 2008).

As discussed earlier, defective DNA repair in normal cells can predispose to tumour development and therefore a further concern with PARP inhibition is that prolonged inhibition may paradoxically increase the risk of second malignancy. In p53 null mice, loss of PARP1 resulted in nearly 50% developing aggressive brain tumours of similar characteristics to human cerebellar medulloblastomas (Tong et al., 2003). Whether similar effects will be seen in humans remains to be seen (Zaremba et al., 2011). Using DNA repair inhibitors in patients with advanced disease, and therefore a limited life expectancy, would be one method of reducing the potential of developing second malignancies, especially until safety data matures (Abbotts and Madhusudan, 2010).

A commonly stated problem in the treatment of patients with GBM is the inability of many systemic therapies to cross the blood-brain barrier. Little is known about the potential for small molecule inhibitors of DNA repair to breach this barrier. This, however, is currently under investigation by Chalmers *et al* who are performing a Phase I study investigating olaparib in combination with temozolomide in patients with relapsed GBM. The results of this work to date have shown that in murine studies olaparib does not cross an intact blood-brain barrier. However, in both mouse models and eight patients with resectable recurrent GBM, olaparib was able to penetrate the tumour at therapeutic levels (Chalmers, 2014).

1.7 Synthetic lethality

One potential way to overcome the toxicity of DNA repair inhibitors in combination treatments is to administer the DNA repair inhibitor as a single agent in specific patient groups. As discussed previously, tumours frequently have defects in DNA repair and this presents the opportunity to target them with a treatment strategy known as 'synthetic lethality'. Synthetic lethality, in the context of DNA repair, is observed when loss of a single DNA repair pathway is not lethal but loss of two pathways in combination is fatal (Lippert et al., 2010). Figure 1.8 provides a diagrammatic representation of synthetic lethality.

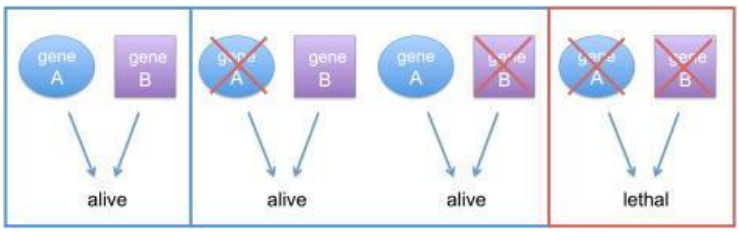


Figure 1.8. Diagrammatic representation of the concept of synthetic lethality

Gene A and gene B are both essential components of two different DNA repair pathways. If both genes, and therefore both pathways, are functioning the cell is able to repair damage to its DNA. If one or other gene is not functioning the cell is able to repair any damage via the other pathway. If however both genes are lost and therefore neither pathway is able to function, the cell is unable to repair damage to its DNA and as a consequence dies (Nijman, 2011). Open access under a Creative Commons BY-NC-ND 3.0 license .

1.7.1 PARP inhibition and synthetic lethality

In addition to their ability to potentiate chemo- and radiotherapy, PARP inhibitors have also been shown to have therapeutic potential when administered as single agents in selected patient groups by exploiting the concept of synthetic lethality. PARP inhibition has been shown to be lethal in cells lacking BRCA1 (Farmer et al., 2005) and BRCA2 (Bryant et al., 2005, Farmer et al., 2005). BRCA1 and BRCA2 are proteins involved in HR. BRCA2 binds directly to the single strand DNA overhang at DSBs as well as to RAD51 molecules, enabling the assembly of the RAD51 filament. BRCA2 interacts with BRCA1, mediated by PALB2. Tumours with deficiency of the BRCA proteins are therefore lacking in HR and the ability to repair DNA DSBs (Willers, 2012). It has been proposed that loss of PARP, through inhibition, results in SSB accumulation and replication fork collapse. The effect of this is to trigger HR (Bryant et al., 2005); however in BRCA deficient tumours this pathway is deficient and hence the cell accumulates DNA damage and dies. BRCA mutations are found in a variety of tumours including both breast and ovarian cancer (Sultana et al., 2012), and it is mostly in these tumour sites that clinical trials of PARP inhibitors have taken place.

The work with PARP inhibitors and the use of the synthetic lethality concept has highlighted opportunities to exploit other synthetically lethal relationships. One such example is demonstrated in the work by Sultana *et al* who have investigated synthetic lethality in both Chinese Hamster (CH) and human cell lines deficient and proficient in DNA DSB repair. In this study, APE1 inhibitors were shown to exhibit toxicity against BRCA2 deficient and ATM

deficient CH cells, both of which are deficient in DNA DSB repair. In CH cells expressing dominant-negative APE1, or human cancer cells depleted of APE1, the use of DSB repair inhibitors (ATM kinase inhibitor, DNA-PKcs inhibitor and Wortmannin) was able to induce lethality (Sultana et al., 2012).

Similarly, Quiros *et al* have explored RAD51 and BRCA2 as potential synthetic lethality targets in glioma. They have demonstrated that knock-down of RAD51 and BRCA2 sensitises glioma cells to DSBs and cell death when treated with nimustine or temozolomide, but not when treated with IR. MGMT expression abolished these effects. Inhibition of DNA-PKcs (a component of the NHEJ pathway) resulted in a minor increase in temozolomide sensitivity, but had a significant impact on sensitivity to IR. The addition of the PARP inhibitor olaparib to knock-down cell lines improved the effects of temozolomide further (Quiros et al., 2011). Albarakati *et al* demonstrated that BRCA1 negative tumour cells have down-regulation of BER and that these cell lines are sensitive to inhibition of ATM and DNA-PKcs alone and in combination with cisplatin (Albarakati et al., 2015).

Phosphatase and tensin homolog deleted on chromosome 10 (PTEN) is a tumour suppressor gene and a key regulator of the phosphoinositide 3 kinase (PI3K) signalling pathway (Di Cristofano et al., 1998, Cully et al., 2006). Over recent years studies have investigated the potential to target PTEN deficient tumours using a synthetic lethality strategy (Abbotts et al., 2014, Mendes-Pereira et al., 2009, Minami et al., 2012).

1.8 PTEN and cancer

The *PTEN* gene was first discovered in 1997 by two groups who were searching for tumour suppressor genes on chromosome 10q23. The newly discovered gene was named *MMAC* (mutated in multiple advanced cancers) or *PTEN* (Li et al., 1997, Steck et al., 1997). The *PTEN* gene is composed of nine exons, with exons 1-6 incorporating the N-terminal domain and exons 6-9 the C-terminal domain. The phosphatase core motif is located in exon 5 (Lee et al., 1999). The C-terminal domain is responsible for PTEN stability and protein-protein interactions (Lee et al., 1999, Georgescu et al., 1999). A graphical representation of the *PTEN* gene is shown in Figure 1.9.

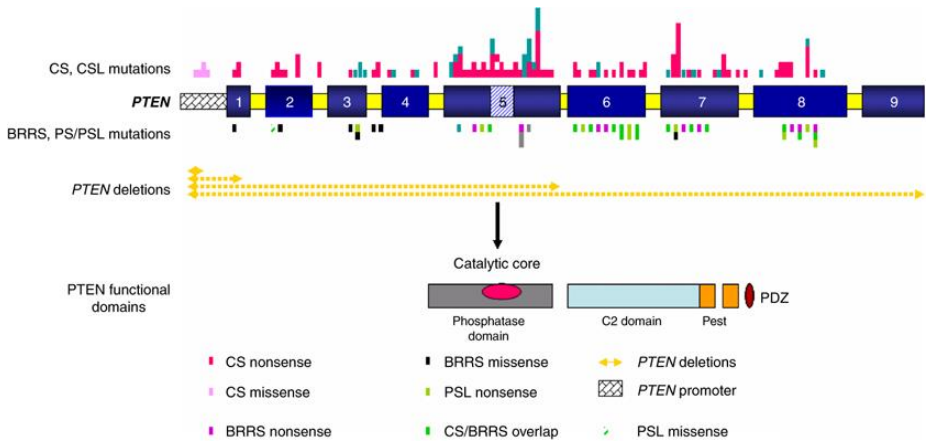


Figure 1.9. Diagrammatic representation of the PTEN gene and its commonly reported mutations.

The nine exons of PTEN are shown in blue with the phosphatase domain in exon 5. The eight introns are shown in yellow. In addition, the locations of previously recorded mutations are shown, alongside the syndrome with which they are associated. Reprinted by permission from Macmillan Publishers Ltd: Oncogene (Orloff and Eng, 2008), copyright (2008).
BRRS=Bannayan–Riley–Ruvalcaba syndrome; CS=Cowden syndrome; CSL=Cowden-like syndrome; PHTS=PTEN hamartoma tumour syndrome; PLS=Proteus-like syndrome; PS=Proteus syndrome.

Since its discovery, germline mutations in *PTEN* have been found to be linked to several hereditary cancer syndromes including Cowden syndrome and Cowden-like syndrome, Bannayan-Riley-Ruvalcaba syndrome, *PTEN* hamartoma tumour syndrome and Proteus syndrome and Proteus-like syndrome. Patients with Cowden-like syndrome have Cowden syndrome features but do not meet the diagnostic criteria for Cowden syndrome. Similarly, patients with Proteus-like syndrome have features compatible with Proteus syndrome but do not meet the diagnostic criteria. In addition to a predisposition for tumour development, these syndromes share other common characteristics including the presence of developmental defects and benign hamartomas (Orloff and Eng, 2008, Liaw et al., 1997, Marsh et al., 1997, Smith et al., 2002). A summary of the PTEN associated cancer syndromes is shown in Table 1.5.

Table 1.5. PTEN associated cancer syndromes.

Condition	Inheritance	Cancer risk	Other features
Cowden syndrome (Liaw et al., 1997)	Autosomal dominant	Breast, thyroid, skin, meningioma	Hamartomas, macrocephaly
Bannayan-Riley-Ruvalcaba syndrome (Marsh et al., 1997)	Autosomal dominant	Potential increased risk	Hamartomas, macrocephaly, cognitive impairment, penis pigmentation
Proteus syndrome (Smith et al., 2002)	Generally sporadic	Benign tumours only	Overgrowth of bones, skin and other tissues. Neurological abnormalities.

In addition, *PTEN* has been found to be mutated in a wide range of malignancies including melanoma (Guldberg et al., 1997), breast (Stemke-Hale

et al., 2008, Li et al., 1997), prostate cancer (Li et al., 1997) (Cairns et al., 1997), and GBM (Li et al., 1997, Keniry and Parsons, 2008). Moreover, in many tumours transcriptional or translational down-regulation of *PTEN* is observed in the absence of *PTEN* mutation (Wiencke et al., 2007) (Zhou et al.).

1.8.1 Functions of the PI3K pathway

PTEN has many roles within the cell, including as a negative regulator of the Phosphoinositide 3-kinase (PI3K) signalling pathway (Stambolic et al., 1998). The PI3K signalling pathway controls essential cellular processes that are involved in the development of cancer, such as cell cycle, cellular proliferation and cell survival. Components of other key cellular pathways, such as the Ras signalling pathway, can also influence the PI3K pathway. Significant cross-talk also exists between PI3K signalling and the p53 tumour suppressor pathway (Cully et al., 2006). Over-activation of the PI3K signalling pathway is common in human cancer (Liu et al., 2009).

An overview of the PI3K signalling pathway is shown in Figure 1.10. In summary (Reviewed in (Cully et al., 2006)), the PI3K signalling pathway is triggered as extracellular signals bind to receptor tyrosine kinases and result in their autophosphorylation and activation. In turn, the p110 catalytic domain of PI3K is re-located to the cell membrane assisted by the Src homology 2 (SH2) domains on the p85 regulatory subunit. PI3K can now convert phosphatidylinositol (4,5) bisphophosphate (PIP_2) to phosphatidylinositol (3,4,5) triphophosphate (PIP_3). The generation of PIP_3 results in the recruitment of other proteins, such as protein kinase B (PKB, also known as

AKT) and its activating kinase, phosphatidylinositol-dependent kinase 1 (PDK1), to the cell membrane. PKB exists in three isoforms, all of which are activated by PDK1. Activated PKB can now phosphorylate down-stream effector proteins such as p27, BAD, GSK3, TSC2 and members of the FOXO family of transcription factors (Cully et al., 2006).

p27, BAD and GSK3 are all inhibited by PKB mediated phosphorylation. p27 is a cell cycle inhibitor and therefore inactivation allows entry into the cell cycle (Liang et al., 2002, Shin et al., 2002). BAD is a pro-apoptotic molecule and hence phosphorylation results in the inhibition of apoptosis (Datta et al., 1997). GSK3 is a serine/threonine kinase that plays a complex role in the regulation of multiple cellular processes, including the regulation of glycogen synthesis, cellular proliferation and apoptosis. Inhibition of GSK3 by phosphorylation appears to promote cellular proliferation; however, the effects on apoptosis remain controversial (Jacobs et al., 2012).

Unphosphorylated TSC2 forms a heterodimer with TSC1, which together promote the GTPase activity of RHEB (a Ras homolog) (Li et al., 2004). As a consequence of PKB mediated phosphorylation of TSC2, and decreased GTPase activity of RHEB, levels of RHEB-GTP increase. GTP-bound RHEB stimulates the nutrient sensor kinase, target of rapamycin (TOR). TOR serves to integrate the formation of glucose, and amino acid availability, with PI3K growth factor signalling (Hay and Sonenberg, 2004). Activation of TOR results in the phosphorylation, and activation, of S6 kinase (S6K), which may play a role in protein translation and cell size. In addition, activated TOR also phosphorylates, and inactivates, the EIF4E inhibitory binding protein 4EBP.

This results in the activation of the translation initiation factor EIF4E, promoting protein synthesis (Fingar et al., 2002, Miron et al., 2003).

FOXO1, FOXO3A and FOXO4 are all transcription factors that are phosphorylated by activated AKT. As these proteins are mediators of apoptosis and cell cycle arrest, inhibition of the FOXO family promotes cell proliferation and survival. Nuclear FOXO3A, for example, can induce transcription of genes involved in cell cycle inhibition (p27 and cyclin G2) and apoptosis. Phosphorylation of FOXO3A by PKB results in its expulsion from the nucleus and renders it unable to perform its transcriptional role (Cully et al., 2006).

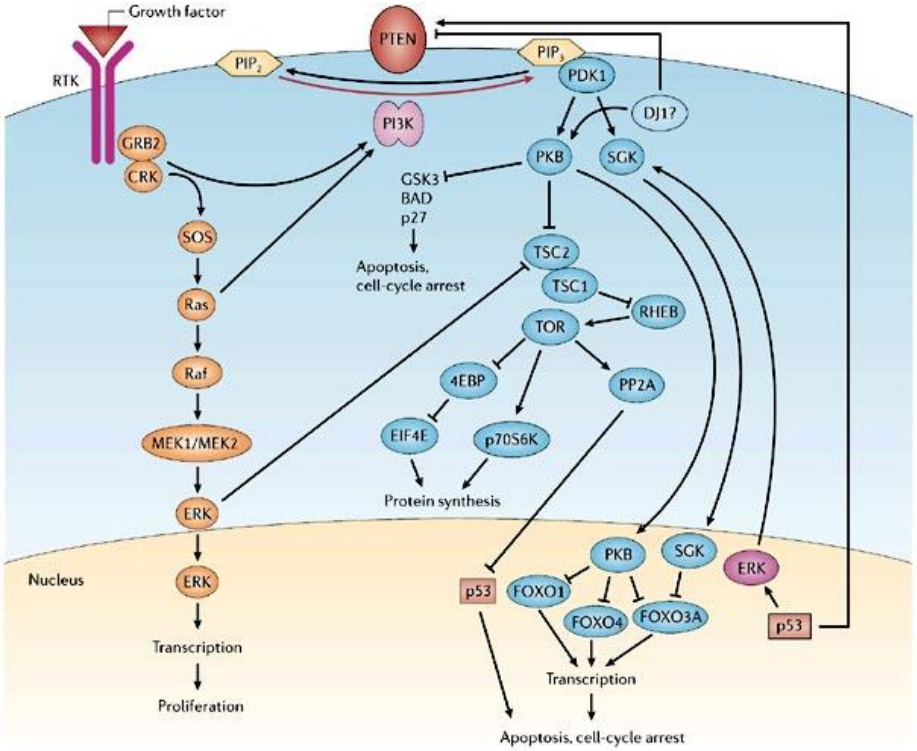


Figure 1.10. Overview of the PI3K signalling pathway

Firstly, receptor tyrosine kinases (RTKs) are activated by extracellular signals including growth factors. A chain of phosphorylation events then occur which either can be stimulatory of inhibitory, beginning with the phosphorylation and activation PI3K (shown in pink). PIP_2 is subsequently converted to PIP_3, resulting in the recruitment of PDK1 and activation of PKB (also known as AKT). Further downstream signalling results in protein synthesis through activation of S6K and the translation initiation factor EIF4E. The FOXO family of transcription factors, p27, BAD and potentially GSK3 all promote cell cycle arrest and/or apoptosis. Inhibition of these proteins by phosphorylated PKB therefore promotes cell proliferation and survival. The interaction between the Ras (orange symbols) and p53 signalling pathways and PI3K signalling (blue symbols) is also shown. PTEN is shown as the key negative regulator of the PI3K signalling pathway. Reprinted by permission from Macmillan Publishers Ltd: Nature Reviews Cancer (Cully et al., 2006), copyright (2006).

In summary, activation of the PI3K signalling pathway promotes cell proliferation and survival. Over-activation of PI3K signalling is commonly due to increased activity of PI3K due to mutation or loss of the pathway's key negative regulator, PTEN. While over-activation of PI3K results in AKT over-

activity, AKT itself can also be subject to activating mutations (Cully et al., 2006, Fruman and Rommel, 2014, Liu et al., 2009).

1.8.2 The PI3K family

While PTEN is the major negative regulator of the PI3K pathway, recent studies have identified that multiple components of the PI3K pathway are commonly mutated, or demonstrate altered expression, in a range of human malignancies (Hafsi et al., 2012). PI3Ks are a family of lipid kinases that were discovered in the 1980's (Engelman et al., 2006, Bader et al., 2005, Vivanco and Sawyers, 2002). Three classes of PI3Ks have been described based on their structure and substrate specificity (Fruman et al., 1998, Katso et al., 2001). The structure and function of these 3 classes of PI3Ks are summarised in Table 1.6 (Worby and Dixon, 2014, Fruman et al., 1998, Katso et al., 2001). Class I PI3Ks have been studied the most and are activated by cell surface receptors directly (Liu et al., 2009). Class I PI3Ks can be further sub-divided in to Class IA and Class IB enzymes (Liu et al., 2009).

Mutations of class IA PI3Ks commonly occur in human cancers and are frequently activating mutations. The *PIK3CA* gene encodes the p110α subunit and has been shown to be mutated in a wide range of tumours including breast, endometrial and colon cancers (Liu et al., 2009). In addition p110α amplifications have been demonstrated in tumours such as lung, cervix and gastric cancers, as well as over 6% of GBMs (Liu et al., 2009, Samuels et al., 2004).

Table 1.6. Summary of the structure and function of the PI3Ks

Class of PI3K	Activators	Structure
IA	Activated directly by cell surface receptors: RTKs, GPCRs, certain oncogenes e.g. small G protein Ras	Heterodimer: p110 catalytic subunit and p85 regulatory subunit
IB	Activated directly by cell surface receptors: GPCRs	Heterodimer: p110γ and p101 regulatory subunit. Also p84 and p87PIKAP regulatory subunits.
II	Activated by RTKs, cytokine receptors and integrins	Single catalytic subunit. Substrates include PI or PI(4)P
III	Amino acids	Single catalytic subunit, Vps34. Only produces PI(3)P which regulates membrane trafficking.

GPCRs=G-protein coupled receptors; PI3K=phosphoinositide 3-kinase; p87PIKAP= PI3K gamma adapter protein of 87kDa; RTKs=receptor tyrosine kinases.

Due to the importance of the PI3K signalling in human cancer, drugs have been developed to target multiple nodes within the PI3K pathway, including PI3K, AKT and mTOR. Many of these drugs are currently being tested in clinical trials or are in routine clinical use (Fruman and Rommel, 2014, Liu et al., 2009).

1.8.3 Functions of PTEN

As discussed previously, *PTEN* is commonly mutated in human tumours, however, not all mutations occur in the phosphatase domain i.e. the domain responsible for PI3K signalling regulation (Eng, 2003). This finding implies that PTEN may have tumour suppressor activity distinct from that due to the PI3K pathway functions. Over recent years PTEN has been shown to be involved in genomic integrity and DNA repair, in addition to its role as a regulator of PI3K signalling (Shen et al., 2007). Although previously considered to be a cytoplasmic protein, recent studies have demonstrated

PTEN to also be present in the nucleus. This differential sub-cellular localisation appears to be associated with a functional role of PTEN. The chromosomal stability and DNA repair functions of PTEN are manifest when PTEN is located in the nucleus, whereas the lipid-phosphatase function occurs when PTEN is present in the cytoplasm (Lian and Di Cristofano, 2005).

1.8.4 PTEN in the nucleus

In recent years various theories have been proposed as to how PTEN enters the nucleus given that it lacks a classic nuclear localisation signal (NLS) (Ming and He, 2012). It has been proposed that nuclear export of PTEN was cell cycle dependent and under control of the PI3K signalling pathway (Liu et al., 2007). Others have linked PTEN nuclear importation to Major Vault protein (MVP), a reputed nuclear-cytoplasmic transport protein. This process requires the interaction of NLS4 and either NLS2 or NLS3 but is not reliant upon PTEN phosphatase activity or phosphorylation of PTEN (Chung et al., 2005).

Bassi *et al* have recently identified an approximately 77 kDa form of PTEN as compared to the expected 55 kDa version. They conclude that this represents a SUMOylated version of PTEN, referred to as SUMO-PTEN (small ubiquitin-like modifier) and K254 was identified as a major PTEN SUMOylation site. In addition, it was shown that nuclear localization and retention of PTEN is controlled by SUMOylation and that SUMOylation, and not just nuclear localisation, is required for the DDR functions of PTEN. Treatment with genotoxic agents, including cisplatin and doxorubicin, resulted in a reduction in SUMO-PTEN. In response to genotoxic stress, and in the presence of ATM,

SUMO-PTEN is expelled from the nucleus. In the absence of nuclear PTEN cells were hypersensitive to DNA damage (Bassi et al., 2013).

1.8.4.1 PTEN in chromosomal stability

Several studies have shown a role for PTEN in chromosomal stability. Firstly, Puc *et al* found that aneuploidy in human primary breast cancers is associated with a lack of PTEN expression (Puc et al., 2005). This finding was followed by work by Shen *et al* who demonstrated that PTEN can associate with centromeres by interaction with centromere protein C (CENP-C) (Shen et al., 2007). This centromere-specific binding protein binds to the centromere, and along with CENP-A, is essential for kinetochore function and correct sister chromatid segregation during mitosis (Kalitsis et al., 1998, Regnier et al., 2005, Tomkiel et al., 1994). Loss of PTEN was also shown to be associated with extensive centromere breakage (Shen et al., 2007). These findings support a role for PTEN in chromosomal stability, whereby as a consequence of PTEN loss kinetochore assembly cannot proceed, resulting in improper sister-chromatid segregation and ultimately chromosomal instability (Kalitsis et al., 1998, Regnier et al., 2005, Tomkiel et al., 1994).

1.8.4.2 PTEN and double strand break repair

PTEN has been shown to be involved in DSB repair. Shen *et al* demonstrated in mouse embryonic fibroblasts that deletion of PTEN causes spontaneous DNA DSBs. In addition, PTEN was shown to transcriptionally regulate RAD51 (Shen et al., 2007), an important component of the HR pathway of DNA DSB repair (Tutt and Ashworth, 2002, Pastink et al., 2001). *In vivo*

PTEN has been shown to be physically associated with the chromatin that surrounds the RAD51 promoter and it has been proposed that PTEN may act on this chromatin to modify transcription of RAD51. PTEN does not act alone to transcriptionally regulate RAD51, and the transcription factor E2F-1 has also been shown to associate with the RAD51 promoter. It has been proposed that PTEN and E2F-1 act together to transcriptionally regulate RAD51 (Shen et al., 2007). In mouse cells disruption of RAD51 has been shown to result in extensive chromosomal aberrations (Smiraldo et al., 2005). As a consequence it has been proposed that tumours with loss of PTEN, and therefore lower levels of RAD51, demonstrate defects in the DSB repair system which in turn can promote chromosomal instability (Shen et al., 2007).

However, in other studies this association between PTEN and RAD51 has not been demonstrated. Gupta *et al* also found that PTEN deficient cells demonstrated genomic instability but did not find a difference in the early steps of DNA damage sensing and chromosome modification between PTEN deficient and proficient cells. The study concluded that the genomic instability observed in PTEN deficient cells was due to causes other than defective DNA repair (Gupta et al., 2009).

1.8.4.3 PTEN and nucleotide excision repair

PTEN may also play a role in the regulation of NER. A link between PTEN and NER has been shown in skin keratinocytes. Ming *et al* have demonstrated that PTEN enhances XPC transcription and therefore positively promotes GGR (Ming et al., 2011). Inhibition of PTEN results in suppression of XPC and as a consequence diminished GGR. It is likely that the PTEN/AKT/p38 axis is

responsible for the regulation of XPC expression. High levels of AKT have been shown to suppress p38 and result in decreased levels of XPC. AKT activation also results in nuclear translocation of p130, a transcriptional repressor of XPC (Ming et al., 2010, Ming and He, 2012).

1.8.4.4 PTEN and cell cycle checkpoints

PTEN has previously been shown to play a role in cell cycle regulation through Cyclin D1 suppression (Diao and Chen, 2007, Weng et al., 2001, Radu et al., 2003). Cyclin D1 forms a complex with Cyclin-dependent kinase (CDK) 4 and CDK6 which allows the cell to process through the G1/S checkpoint (Radu et al., 2003). In the presence of PTEN, progression through the G1/S checkpoint is stalled due to decreased levels, and nuclear localisation, of Cyclin D1 (Radu et al., 2003). This stalling of the cell cycle allows additional time for DNA repair to take place.

Checkpoint kinase 1 is encoded for by the *CHEK1* gene and is a serine/threonine-specific protein kinase that co-ordinates the DDR and cell cycle checkpoints. Chk1 acts to prevent damaged cells progressing through the cell cycle by activating cell cycle checkpoints causing cell cycle arrest and resulting in increased DNA repair. Chk1 has been shown to act at various points throughout the cell cycle but predominately in the G2 and S phases (Lukas et al., 2003). Chk1 has also been shown to be involved in the regulation of PTEN (Martin and Ouchi, 2008). Martin *et al* demonstrated that a lack of Chk1 results in lower levels of PTEN phosphorylation and also of total PTEN. Activation of both Chk1 and PTEN occurs in response to a series of phosphorylation events. Firstly Chk1 is phosphorylated at Ser137 by ATR

which results in the phosphorylation of PTEN by CK2 at Thr383. This phosphorylation cascade is essential for recovery of the cell cycle after stalled DNA replication (Martin and Ouchi, 2008)

PTEN loss has been shown to result in the cytoplasmic sequestration of ubiquitinated Chk1 and to impair Chk1-mediated checkpoint activation (Puc et al., 2005). Puc *et al* demonstrated that cells deficient in PTEN have a partially impaired checkpoint response when subjected to IR. In this study genomic instability appeared to be driven by a lack of PTEN, and subsequent enhanced PI3K pathway signalling, resulting in impaired Chk1 function through phosphorylation, ubiquitination and decreased nuclear localisation (Puc et al., 2005). In another study by Puc *et al* PTEN loss was shown to inhibit Chk1 and result in the accumulation of DNA DSBs (Puc and Parsons, 2005).

Taken together the evidence suggests that PTEN and Chk1 activity are closely linked. Loss of PTEN is associated with a loss of cell cycle checkpoint control that appears to be mediated by impaired Chk1 function. A failure of cell cycle checkpoints results in less time for DNA damage repair and may explain the accumulation of DNA DSBs observed in the study by Puc *et al* (Puc and Parsons, 2005).

1.8.5 PTEN and p53

PTEN and p53 have been shown to interact closely (Freeman et al., 2003). p53 can upregulate transcription of *PTEN* and *TSC2*, a downstream target in the PI3K signalling cascade (Stambolic et al., 2001, Inoki et al., 2003). This negative regulation of PI3K signalling is enhanced by p53-mediated repression

of the gene encoding the catalytic subunit of PI3K (Singh et al., 2002). Additionally, a co-operative tumour suppressor role has been identified between p53 and PTEN (Chang et al., 2008). This may be due to the fact that PTEN can modify p53 protein levels and also transcriptional activity by regulating the binding of p53 to DNA (Freeman et al., 2003).

A p53-binding element has been identified directly upstream of the *PTEN* gene and is required for p53 inducible transactivation by *PTEN*. Constitutive expression of PTEN is controlled by an additional p53-independent element. In cell lines that express wild-type p53, induction of p53 resulted in increased levels of *PTEN* mRNA, whereas similar results are not seen in p53 mutant cell lines (Stambolic et al., 1998).

Conversely, PTEN has been shown to modify p53 function. A study by Zhang *et al* has shown that the way in which p53 is expressed in response to DNA damage influences the outcome for the cell. If p53 is expressed in sustained pulses, then p53 acts as a prosurvival factor, promoting cell cycle arrest and DNA repair. This early phase response involves the induction of Wildtype p53-induced Phosphatase 1 (Wip1) which plays a key role in regulating p53 levels via the ATM-p53-Wip1 and p53-Mdm2 loops. ATM is activated in response to DNA damage and this in turn activates p53. Mdm2 is a key regulator of p53 and targets p53 for degradation. p53 and Mdm2, however, form a negative feedback loop whereby Mdm2 is a transcriptional target for p53. Wip1 acts to stabilise Mdm2 and thus target p53 for degradation. This cycling of feedback loops results in the pulsatile production of p53. If, however, PTEN is present, full activation of p53 occurs resulting in p53-

dependent apoptosis. This late phase response occurs via the p53-PTEN-AKT-Mdm2 loop. Higher levels of PTEN result in decreased levels of AKT and MDM2 and therefore increased levels of p53 (Zhang et al., 2011). A decrease in cellular AKT (Song et al., 2011) or an increase in p53 (Chang et al., 2008) levels can result in G1/S cell cycle arrest. AKT levels also modulate the G2/M cell cycle checkpoint. Cells lacking PTEN, or in which AKT is subject to an activating mutation, can lose control of the G2/M checkpoint allowing the accumulation of DNA damage (Liang and Slingerland, 2003).

1.8.6 PTEN and ATM

Recently, ATM has been shown to mediate phosphorylation of PTEN, which in turn promotes nuclear translocation. Activation of ATM in response to topotecan and cisplatin has been shown to result in phosphorylation of PTEN at serine 113 and to regulate nuclear translocation. In response to topotecan, nuclear PTEN localisation results in autophagy via activation of the p-JUN-SESN2/AMPK pathway (Chen et al., 2015).

1.8.7 PTEN expression in glioma

PTEN mutations have been found to occur more frequently in high grade glioma as compared to low grade tumours. Rasheed *et al* analysed 42 adult GBMs and thirteen adult anaplastic astrocytomas. They found that thirteen GBMs and 3 anaplastic astrocytomas contained *PTEN* mutations. Of the 21 low-grade gliomas studied none contained *PTEN* mutations. Interestingly, none of the 22 childhood gliomas of any grade demonstrated mutations of

PTEN (Rasheed et al., 1997). Another study has demonstrated that *PTEN* expression was significantly lower in high grade glioma as compared to lower grade lesions and that prognosis was better for patients who expressed higher levels of *PTEN*. In addition, immunostaining of GBMs found that approximately two-thirds had little or no expression of PTEN. Also of note, in patients with high PTEN expression, roughly two-thirds demonstrated noticeable heterogeneity in PTEN expression with regions with absent staining. Based on the results from both real-time polymerase chain reaction (PCR) and immunohistochemical analysis the authors of this study estimate that 80% of GBMs and 40% of lower grade gliomas have low expression of PTEN (Sano et al., 1999).

Despite *PTEN* mutations being exceptionally rare in low grade gliomas, levels of PTEN protein expression have been found to be lower than in normal brain. Levels of AKT phosphorylation and activity have also been shown to be raised in low grade gliomas. This finding is potentially explained by the work of Weincke *et al* who demonstrate that the *PTEN* promoter is frequently methylated in low grade gliomas and secondary GBMs. PTEN promoter methylation was found to be less common in *de novo* GBMs and was not observed in normal brain (Wiencke et al., 2007).

1.8.8 PTEN and synthetic lethality

Several studies have shown that PTEN mutant cells can be targeted with a synthetic lethality strategy by using PARP inhibitors (Mendes-Pereira et al., 2009, Minami et al., 2012). Menedes-Pereira *et al* have demonstrated that human tumour cells deficient in PTEN have defective HR, which in turn results in increased sensitivity to PARP inhibitors as compared to PTEN proficient tumour cells. $PTEN^{-/-}$ status was found to be associated with reduced expression of RAD51 in isogenically matched wild type and $PTEN^{-/-}$ colorectal and endometroid adenocarcinoma cell lines. In addition, these PTEN deficient cell lines were less able to form RAD51 foci in response to damage to DNA and also demonstrated a down-regulation of DSB repair by HR as assessed by HR reporter assay. PTEN deficient cell lines were found to be more sensitive to both PARP inhibitors and cisplatin (an agent that also targets HR). Interestingly, $PTEN^{-/-}$ cell lines did not demonstrate increased sensitivity to paclitaxel, a drug that targets microtubules. This was assumed to provide further evidence that the selective lethality seen with PARP inhibitors and cisplatin in PTEN deficient cells is due to HR deficiency rather than general chemosensitivity. Heterozygosity for PTEN did not reduce RAD51 levels by the same extent seen in $PTEN^{-/-}$ cell lines and only resulted in a small decrease in HR; damage induced RAD51 foci were still observed and PARP sensitivity did not occur (Mendes-Pereira et al., 2009).

Leading on from the work by Ashworth and colleagues, Minami *et al* demonstrated a synergistic effect from the combination of the PARP inhibitor olaparib with cisplatin in PTEN deficient lung cancer cell lines. In this study

PTEN deficiency was also associated with reduced levels of nuclear RAD51 and foci of RPA. In addition, phosphorylated Chk1 and MRE11 levels were also decreased, implying a link between PTEN inactivation and suppression of DNA repair (Minami et al., 2012).

McEllin *et al* studied both genetically defined primary murine astrocytes and human glioma cell lines and found that PTEN deficiency was associated with sensitivity to alkylating agent chemotherapy and PARP inhibition. PTEN deficiency was also found to be associated with a decrease in HR. Although no changes in RAD51 protein or mRNA levels were demonstrated in mouse astrocytes on PTEN loss, decreased transcript levels of the RAD51 homologs RAD51B, RAD51C and RAD51D were observed. It is proposed that, as these complexes are involved in RAD51 filament formation, decreased levels of these RAD51 homologs may be responsible for the diminished HR present in association with PTEN loss (McEllin et al., 2010).

Abbotts *et al* have shown that PTEN deficient melanoma cell lines can be targeted with inhibitors of APE1, a key component of the BER pathway. PTEN deficient melanoma cell lines were shown to have lower levels of expression of genes involved in DNA DSB repair as compared to PTEN proficient melanoma cell lines. Synthetic lethality was demonstrated in PTEN deficient melanoma cell lines treated with APE1 inhibitors. PTEN deficient melanoma cell lines were more sensitive to ATM inhibition, accumulated more DNA DSBs and had a higher rate of apoptosis than their PTEN proficient counterparts (Abbotts R et al., 2013).

However, the association between PTEN deficiency and sensitivity to inhibition of BER has not been observed by all groups. Fraser *et al* found no association between PTEN deficiency and loss of RAD51 function in prostate cancer. Using prostate cancer cell lines with varying p53 and PTEN expression they found that reduced PTEN expression was not associated with decreased RAD51 levels. In addition, no association was seen between PTEN expression and sensitivity to PARP inhibitors in these prostate cancer cell lines, with only mild PARP inhibitor sensitivity seen in PTEN deficient cells. PTEN deficient cells were found to be sensitive to IR, UV, hydrogen peroxide, mitomycin C and Methyl methanesulphonate (MMS). The same cell lines were not sensitive to campthothecin, cisplatin or paclitaxel. In this study MRE11 expression was associated with PARP inhibitor sensitivity. In human prostate cancer tumours loss of PTEN expression was associated with increased, and not decreased, expression of RAD51 (Fraser et al., 2012). Similarly, Gupta *et al* found no association between RAD51 expression and DNA damage repair kinetics post irradiation in PTEN deficient cells (Gupta et al., 2009). The work by Bassi *et al* may explain the variable associations seen between PTEN and RAD51 status. In this study it was shown that cells deficient in PTEN, cells lacking nuclear PTEN (i.e. K254R mutants) and cells lacking all PTEN phosphatase activity exhibited deficient DNA repair, as evidenced by defective resolution of γH2AX and BRCA1 foci and a failure to recruit RAD51 to sites of damaged DNA. The actual amount of RAD51 mRNA or protein was not associated with recruitment of RAD51 to DNA damage sites (Bassi et al., 2013).

It is likely that PTEN interactions are complex and may differ between cell types (Fraser et al., 2012). Potential explanations for the discrepancies observed have been proposed and are the subject of on-going studies. PTEN deficiency may be associated with DSB deficiency via a route other than RAD51, or PTEN itself may be a marker of another DNA repair defect. Loss of PTEN has been shown to be highly associated with BRCA1 mutations. Analysis of a small cohort of 34 BRCA1 mutated breast cancer patients found that over 80% had loss of PTEN by immunohistochemistry (Phuah et al., 2012). BRCA1, as discussed earlier, is an essential protein in HR and therefore PTEN may be a marker of DSB repair loss by association with BRCA1. PTEN loss has also been shown to inhibit CHK1, which plays an essential role in the regulation of cell cycle progression and maintenance of genomic integrity. CHK1 inhibition results in the accumulation of DNA DSBs. PTEN loss has been shown to result in the cytoplasmic sequestration of ubiquitinated CHK1 and impairs CHK1-mediated checkpoint activation (Puc et al., 2005).

Mereniuk *et al* have investigated a synthetic lethality relationship between polynucleotide kinase/phosphatase (PNKP) and PTEN. PNKP is an enzyme involved in DNA repair that has been shown to interact with DNA Polβ and XRCC1. They have found that PTEN deficiency, in association with PNKP inhibition, results in hypersensitivity to ionizing radiation. The synthetic lethality relationship observed appeared to be associated with the cytoplasmic lipid phosphatase activity of PTEN (Mereniuk TR et al., 2013).

These studies highlight the potential application of targeting PTEN deficiency in other tumours and also using alternative inhibitors of DNA repair or the

DDR. As discussed earlier, PTEN loss is observed in a significant proportion of human gliomas, therefore this tumour is an ideal candidate in which to explore PTEN and synthetic lethality. Clearly, however, further work is required to fully understand the mechanism behind PTEN deficiency and synthetic lethality.

1.8.9 ATM inhibition as a synthetic lethality strategy in PTEN-deficient tumours

As described in Section 1.6.3, ATM inhibition has been studied as a potential combination agent to be used with conventional treatment strategies. Over recent years the potential of exploiting a synthetic lethality approach with ATM inhibitors has been investigated. ATM inhibition has been found to be a potential treatment approach in individuals with deficiency of the Fanconi anaemia (FA)/BRCA pathway (Kennedy et al., 2007, Albarakati et al., 2015). McCabe *et al* screened isogenic ATM wild-type and deficient cell line systems for synthetic lethal interactions using an siRNA library. Using this approach PTEN was found to have a synthetically lethal relationship with ATM, alongside eight other genes including components of the FA/BRCA pathway, genes involved in cell cycle progression and the G1/S checkpoint plus genes involved in immunity and apoptosis. PTEN knock-down by siRNA in ATM isogenic cells resulted in reduced survival in ATM deficient cells as compared to those that express wild-type ATM. Similarly, in isogenic HCT116 and PC3 cell line models both PTEN deficient cell lines demonstrated increased sensitivity to both KU55933 and siRNA mediated knock-down of ATM. In HCT116 and PC3 isogenic models increased levels of reactive oxygen species

(ROS) were observed in association with PTEN deficiency. The level of ROS was further increased with the addition of ATM inhibition in PTEN deficiency as compared to the PTEN wild-type system. McCabe *et al* proceeded to demonstrate efficacy of ATM inhibition with KU66019 in a PTEN Tetracycline inducible PC3 model in a subcutaneous xenograft setting. Statistically significant slowing of tumour growth was demonstrated in PTEN-deficient tumours as compared to wild-type controls in response to ATM inhibition (McCabe et al., 2015).

1.9 Conclusion

DNA damage can be caused by both endogenous and exogenous sources. If this damage is not repaired it may be passed on to daughter cells or, if severe, result in cell death. Cells therefore have an armoury of repair mechanisms to attempt to counteract this problem. Different types of DNA damage can occur which are repaired by different DNA repair pathways. Overlap between these pathways does exist and can provide a 'back-up' pathway should the first pathway fail. Deficiency of DNA repair is central to both cancer development and response to cancer treatment. If DNA repair is ineffective then this can result in the accumulation of mutations within the DNA and lead to tumour development. Conversely, active DNA repair in the setting of cancer treatment can result in treatment resistance, as both chemotherapy and radiotherapy act by damaging cellular DNA. It is therefore unsurprising that DNA repair status has been found to have both prognostic and predictive significance in a variety of tumour types. The significance of DNA repair status has also been shown to

differ between cancer types, with high levels of a particular DNA repair gene associated with poor survival in one tumour type whereas in another cancer type it is low levels that are associated with an adverse outcome.

MGMT expression has been extensively studied in glioma with low levels shown to be associated with a better outcome and improved response to both chemotherapy and radiotherapy. The influence of other DNA repair genes on outcome in glioma has only been studied to a limited extent. This thesis is the first study to explore the prognostic significance of over 150 DNA repair genes in high grade glioma and to explore further the expression of the most important DNA repair genes at the protein level.

Individuals with a diagnosis of GBM have a poor prognosis, even with intensive treatment with surgery, radiotherapy and chemotherapy. New treatment approaches are required and targeting DNA repair may represent an exciting option. Previous studies have investigated the potential to enhance the effects of both chemotherapy and radiotherapy by adding in an inhibitor of DNA repair. Another option for DNA repair modulation involves the concept of synthetic lethality, which has been extensively studied in BRCA deficient breast and ovarian cancers treated with PARP inhibitors. PTEN has recently been identified as a synthetic lethality target in tumours and has been proposed to be associated with defective DSB repair and to be involved in genomic integrity. Synthetically lethal interactions have been demonstrated between PTEN loss and inhibition of components of the BER pathway. PTEN is commonly mutated in high grade glioma and therefore this introduces the potential to target PTEN deficient glioma with a synthetic lethality approach.

In this thesis the ability to induce synthetic lethality in PTEN deficient GBM cell lines by the addition of an ATM inhibitor will be explored.

1.10 Aims of study

1. To investigate the prognostic significance of a panel of DNA repair genes in adult GBM.

2. To develop a DNA repair gene prognostic index for use in adult GBM incorporating the most significant DNA repair genes identified in Aim 1.

3. To investigate the prognostic significance of DNA repair protein expression for the key DNA repair genes included in the prognostic index.

4. To explore the potential of targeting PTEN-proficient and PTEN-deficient GBM cell lines with an ATM inhibitor to induce synthetic lethality.

5. To create stable shRNA knock-down models of PTEN deficiency to substantiate the findings from the GBM cell line studies in Aim 4.

CHAPTER 2

Materials and Methods

2 Materials and Methods

2.1 DNA repair gene expression profiling in adult glioblastoma

Two publically available adult glioblastoma datasets were identified. One dataset was used as a Test dataset and the second as a Validation dataset.

2.1.1 Data acquisition for the Test dataset

The Test dataset (accession number E-GEOD-13041) was obtained from the Array Express on-line resource (http://www.ebi.ac.uk/arrayexpress) and is a publically available gene expression dataset of patients with a diagnosis of GBM (Kolesnikov et al., 2015). The dataset contains microarray gene profiling data for 267 patients using three different Affymetrix platforms. One-hundred and ninety-one GBM patients were included in the data analysis for the Test dataset, all of whom were profiled using the Affymetrix U133A array.

2.1.2 Data acquisition for the Validation dataset

The Validation dataset was downloaded from 'The Cancer Genome Atlas' (TCGA) on-line resource (http://cancergenome.nih.gov/) (Weinstein et al., 2013) and included patients with a diagnosis of GBM and gene expression assessed using the HT_HG_U133A Affymetrix array. This will be referred to as the TCGA dataset. Five hundred and forty-eight files were identified in this

dataset. Duplicate cases were removed alongside cases with missing survival data or that were identified as not meeting the original study criteria. A total of 508 cases were included in the subsequent analyses.

2.1.3 Probeset selection

Probes that represent the 188 DNA repair genes listed in Appendix 1 were identified using the NetAffx™ Analysis Center on-line resource (www.affymetrix.com) (Liu et al., 2003). A total of 157 DNA repair genes, represented by 248 probes, were present in both the test and TCGA datasets. Expression data was extracted for all 248 probes in the both datasets. Baseline demographic data was also obtained, including: age, gender and performance status. Treatment and survival data (if available) was also analysed. Initial data processing was performed in Microsoft Excel 2010 (Microsoft, Redmond, USA).

2.1.4 Patient demographics for the Test dataset

The median age of patients in the Test dataset was 54 years (range 18-86 years). 73/191 (38.2%) patients were female. Patients were followed up for a median of 385 days (range 7-3353 days) and at the end of follow-up 176/191 (92.1%) had died. Treatment data was not available for Test dataset patients.

2.1.5 Patient demographics for the TCGA dataset

In the TCGA dataset the median age of patients was 59 years (range 10-89 years). 200/508 (39.4%) patients were female. Karnofsky performance status

(KPS) data was available for 381 patients; median KPS was 80 (range 20-100). 416/508 (81.9%) patients had died after a median follow up time of 353 days (range 2-3880 days). Three hundred and fifty-one (69.1%) patients were treated with chemotherapy; 294 (57.9%) received temozolomide. Baseline demographic and treatment data for the TCGA dataset is shown in Table 2.1.

Table 2.1. Baseline demographic and treatment data for the TCGA dataset

Clinicopathological variable	**Number (%) (n=508)**
Gender	
Male	308 (60.6%)
Female	200 (39.4%)
Karnofsky performance status	
100	54 (10.6%)
80-90	231 (45.5%)
60-70	83 (16.3%)
40-50	11 (2.2%)
0-30	2 (0.4%)
Unknown	127 (25.0%)
Survival Status	
Alive	92 (18.1%)
Deceased	416 (81.9%)
Radiotherapy given	
Yes	373 (73.4%)
No	135 (26.6%)
Chemotherapy given	
Yes	351 (69.1%)
No	157 (30.9%)
Type of chemotherapy given[#]:	
Temozolomide	294 (57.9%)
*PCV**	26 (5.1%)
BCNU (Carmustine)	17 (3.3%)
Bevacizumab	73 (14.4%)
Targeted agent	39 (7.7%)
Gliadel wafer	30 (5.9%)
Other	99 (19.5%)

*Procarbazine, lomustine and vincristine. #Total >351 as some patients received more than one type of chemotherapy.

2.1.6 Survival analysis based on DNA repair gene expression level

Xtile (Version 3.6.1, Yale University, USA) was used to dichotomise (high/low) levels of DNA gene expression prior to Kaplan Meier survival analysis. Kaplan Meier survival curves were constructed in SPSS (Version 20, Chicago, USA) for 248 probes (in both the Test and TCGA datasets) and a log rank score calculated. The Benjamini and Hochberg False Discovery Rate calculation (BH FDR) (Benjamini, 1995) was applied to account for multiple comparisons.

Cox multivariate regression models were constructed for each dataset including probes significant (after BH FDR correction) in both datasets (n=14). Non-significant probes after the first round of analysis were removed and the analysis re-run. This was repeated until only significant probes remained (six probes in the Test dataset and seven probes in the TCGA dataset). The models for the two datasets were compared and the analysis re-run with the five probes significant in both datasets.

2.1.7 DNA repair gene prognostic index

Prognostic indices were calculated for each dataset (PI_1 = prognostic index for the Test dataset and PI_2 = prognostic index for the TCGA dataset) using the following equation:

Σ *(DNA repair gene expression level* β value)*

where DNA repair gene expression level is represented as 0 (low) or 1 (high) and the β value is the standardised regression coefficient obtained from the

final multivariate model described in Section 2.1.6. A combined prognostic index (PI_3) was also calculated using the mean β value for each gene from the two datasets. A prognostic index score was calculated for each patient, firstly using PI_1 in the Test and then the TCGA dataset. For each dataset, patients were divided into three to four prognostic groups based on their prognostic score and Kaplan Meier survival curves were constructed. The log rank test was applied to assess the survival difference between groups. This was repeated for PI_2 and PI_3. The combined prognostic index (PI_3) performed best and separated patients into three statistically significant prognostic groups in both the Test and TCGA datasets.

Using PI_3, curves were constructed to predict 1, 2 and 3 year survival in GBM patients in both the Test and TCGA datasets. Firstly, Kaplan Meier survival life tables were analysed to determine the percentage of patients alive at 1, 2 and 3 years. The percentage survival at 1 year (y axis) was plotted against the median prognostic score for patients within each of the three prognostic groups (x axis) and a 2^{nd} order polynomial curve fitted to the data. This process was repeated for 2 and 3 year survival.

2.2 Artificial neural network (ANN) analysis

Artificial neural network (ANN) analysis was used to identify genes that interact with *APE1, MGMT*, *NBN*, *PMS2* and *PTEN* in the TCGA dataset. This work was performed in collaboration with Miss Devika Aggarwal and Professor Graham Ball (Department of Biosciences, Nottingham Trent University, Nottingham, UK) who have published extensively on the use of this methodology to analyse complex datasets in a variety of biological settings (Lancashire et al., 2010, Lancashire et al., 2009).

2.2.1 Top 200 APE1, MGMT, NBN, PMS2 and PTEN interactors

Firstly an ANN algorithm was applied to identify genes that are the most able to predict *APE1*, *MGMT*, *NBN*, *PMS2* and *PTEN* expression. The probes selected to represent each gene were those used in the prognostic index (*PTEN* 204054_at, *APE1* 210027_s_at, *NBN* 202905_x_at, *MGMT* 204880_at and *PMS2* 209805_at) and a separate network analysis was performed for each gene. The expression level for the gene of interest was dichotomised into low and high expression based on the median expression value. A total of 22,277 probes were screened to identify those best able to predict *APE1, MGMT*, *NBN*, *PMS2* and *PTEN* expression.

The technique used was a non-linear, ANN modelling based, data mining approach. This algorithm uses a multi-layer perception approach consisting of an input node, a hidden layer and an output node. The input node represents the gene expression data for the 22,277 probes studied. The hidden layer

consists of two hidden nodes that use a sigmoidal transfer function to map a large amount of input data to a small output sample. The output node represented the expression level of the gene under investigated and was coded as 0 if a case was low *APE1*, *MGMT*, *NBN*, *PMS2* and *PTEN* expression (<median) and 1 if high *APE1*, *MGMT*, *NBN*, *PMS2* and *PTEN* expression (>median).

Application of the ANN algorithm connects each node in a layer with all the nodes in the following layer with initially a randomised weight. This information is fed forward through the network to generate an output using the sigmoidal transfer function. This information is then propagated back and the model weights for the next round of processing are modified based on the information gained from the previous epoch; hence the methodology is described as being a feed forward back propagation algorithm which employs a supervised learning strategy. The aim of this process to correctly predict the expression of the gene of interest for each of the input genes. For each input gene the predicted expression of the output gene for each individual patient was compared against the actual expression of the output gene. The predicted expression of the output gene (0 or 1) was then subtracted from the actual expression (0 or 1) to give a value between -1 and +1, known as the root mean square (RMS) value. A mean square error (MSE) was then calculated, as the sum of the RMS values for all the cases for a particular input gene divided by the number of cases. After repeated rounds of analysis the MSE decreased to give a minimum possible error. The algorithm was set to terminate when either the MSE had reached 0.01 or 100 epochs had been performed without a

reduction in the MSE, or alternatively a total of 300 epochs had completed. The speed at which this process is completed is influenced by both the learning rate and the speed of the back propagation (momentum), which were set at 0.5 and 0.1 respectively, based on the results of previous studies.

If this methodology was applied to the entire dataset then there is a risk of over-fitting the data. To prevent this a Monte-Carlo cross validation strategy was employed prior to ANN analysis, whereby the samples were randomly separated in to three groups; 60% for training, 20% for testing and 20% for validation of model performance. Initially the feed forward back propagation methodology was applied to the training set, with periodic evaluation in the test set. With successive rounds of analysis the MSE in both the training and test sets converged towards the minimum error. Using this methodology a point would be reached whereby the MSE would increase due to overfitting of the training set and therefore a model was produced that utilised the iteration that resulted in the lowest MSE in the test set. A total of 50 rounds of training were performed for each model. Subsequently this model was tested in the unseen validation set and subjected to 10 rounds of testing to ensure that the model was generalizable to future cases.

A final output model was generated represented by the average MSE in which each individual input gene could predict the output gene expression level (greater than or less than the median). Each input gene was ranked based on the MSE in ascending order. This enabled the top 200 genes that predict *APE1*, *MGMT*, *NBN*, *PMS2* and *PTEN* expression to be identified.

2.2.2 Gene-gene interactions for APE1, MGMT, NBN, PMS2 and PTEN

An interaction algorithm was applied to the top 200 genes that predict *APE1*, *MGMT*, *NBN*, *PMS2* and *PTEN* expression to determine the top 100 interactions that occur between these genes. Of the 508 patients sampled, 507 patients were grouped together as a training set, with the remaining sample serving as the test set. This is referred to as an ensemble sample classification method within a leave-one-out cross-validation scheme. The "Empirical Bayes moderated t-statistic" (Smyth, 2004) was used to independently select differentially expressed genes in each of the 507 training sets. A procedure known as external cross validation was used, whereby the training datasets were used to train a machine learning model. This model was evaluated based on its ability to classify the remaining test sample. The prediction results of four algorithms (Support Vector Machine, Random Forest, kNN and Prediction Analysis for Microarrays, with all parameters being optimised by using a grid search within a nested cross-validation) (Tibshirani et al., 2002) were combined to classify the test sample. This helped to compensate for any bias or variance introduced by the use of these machine learning methods. Genes were ranked based on the cross-validation results. The frequency of occurrence of a particular gene in a list of significantly differentially expressed genes across the range of cross-validation cycles was recorded. Genes that occur more frequently in this list have higher scores. All steps in this analysis were performed using the microarray analysis software available at www.arraymining.net.

This model predicted a weighted link (direction and magnitude) between each of the top 100 gene probe markers associated with *APE1*, *MGMT*, *NBN*, *PMS2* and *PTEN* expression. The 100 strongest pairwise interactions were then scaled between 1 and 15 and visualised as a network map using Cytoscape (Version 3.0.1, Cytoscape Consortium, San Diego, USA) (Smoot et al., 2011).

A flow chart summarising the steps in developing the prognostic index and ANN analysis are shown in Figure 2.1.

The top three hubs within each network were identified and Kaplan Meier survival analysis performed. Kaplan Meier survival analysis was also performed for any other DNA repair associated genes present within the network.

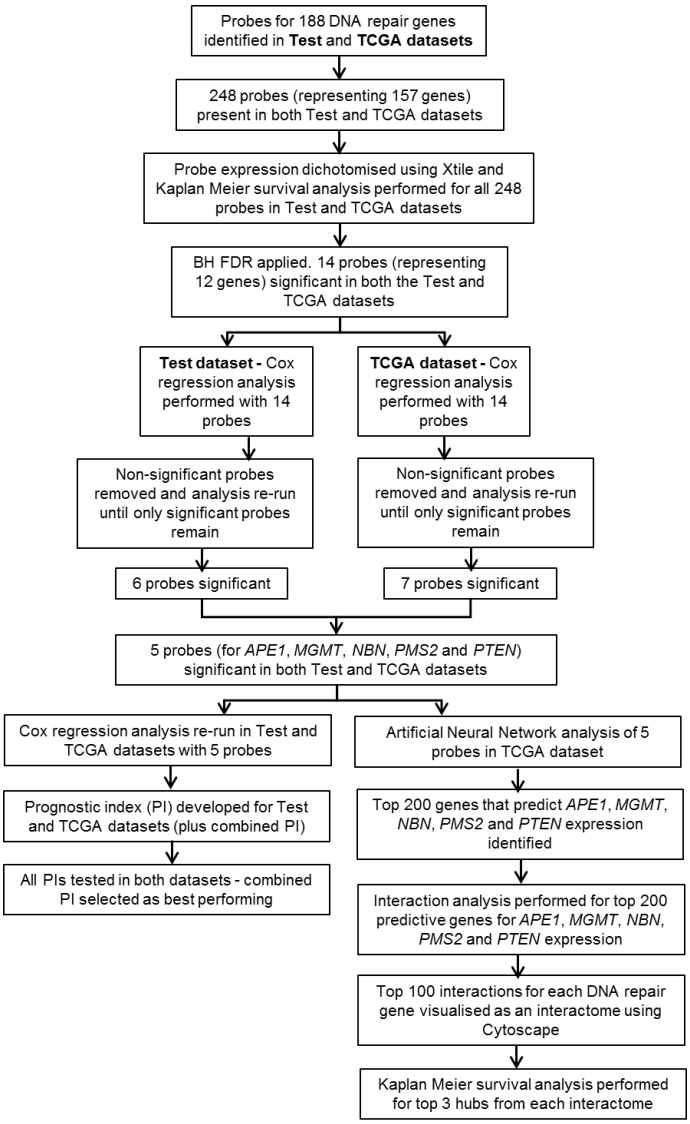


Figure 2.1. Summary of the methods used to develop the prognostic index and for Artificial Neural Network (ANN) analysis

Firstly the most significant DNA repair genes that influence prognosis in both the Test and TCGA datasets were identified. The results from the two datasets were compared and, after correction for multiple comparisons and multivariate analysis, five DNA repair genes were identified that were consistently associated with survival in both datasets. A prognostic index was then developed to incorporate these five genes. Artificial neural analysis was performed in the TCGA dataset for each of these five DNA repair genes to identify other genes with which they interact.

2.3 DNA repair gene expression profiling by immunohistochemistry

Investigation of the protein expression of ATM, APE1, NBN, PTEN and PMS2 in high grade glioma was carried out on paraffin-embedded tumour sections from 61 high grade adult glioma patients treated at Nottingham University Hospitals (NUH) between 2005 and 2011 (referred to as the Nottingham Cohort).

2.3.1 Immunohistochemistry technique

Immunohistochemical staining was performed using the Leica Novolink max polymer detection system (Leica Microsystems, Milton Keynes, UK), in combination with Shandon Sequenza chambers (Thermo Fisher Scientific, Loughborough, UK) as per the manufacturer's instructions (Leica Microsystems). Slides were heated to 60°C for 10 minutes and allowed to cool. Paraffin wax was removed with xylene and the tissue rehydrated by passage through decreasing concentrations of alcohol using the Leica Autostainer XL (Leica Microsystems) (Program 1). Slides were placed in citrate buffer (pH 6) and heated to 95°C (20 minutes) for antigen retrieval. After cooling, slides were loaded into cover-plates in a Sequenza rack (Thermo Fisher Scientific). Slides were then incubated with peroxidase block, followed by protein block, each for 5 minutes. Peroxidase block was used to neutralise endogenous peroxidases and protein block to reduce non-specific binding of the primary antibody and the polymer. Slides were then exposed to primary antibody at the chosen concentration for 60 minutes (overnight for ATM).

Post-primary (Rabbit anti Mouse IgG, used to detect mouse antibodies) and polymer block (used to detect rabbit immunoglobulins including the post primary and primary rabbit antibody bound to tissues) were applied for 30 minutes each. DAB (3,3' – Diaminobenzidine) working solution was added (1:20, chromogen:DAB substrate) for 5 minutes to produce a peroxidase reaction resulting in the formation of a brown precipitate at the reaction site. Between all incubation steps two Tris-buffered saline (TBS) 5 minute washes were performed. Haematoxylin was applied as a counterstain for 6 minutes and slides washed with distilled water to prevent further staining. 100µl of each reagent/antibody was used per slide for each step. Slides were dehydrated in alcohol and passed through a further xylene wash using the Leica Autostainer XL (Program 2). Sections were mounted under coverslips using DPX (mixture of distyrene, plasticizer, and xylene) and allowed to dry overnight. When staining the whole cohort for an individual marker, sections were processed in batches of approximately twenty slides.

2.3.2 Primary antibodies

Primary antibodies used for immunohistochemistry (IHC) analysis are shown in Table 2.2, along with the final concentration used. Prior to use for IHC all antibodies were evaluated for specificity by Western blot analysis as described in Section 2.4.4.

Table 2.2. Antibodies used for IHC analysis

Antigen	Antibody	Clone	Source	Catalog #	Dilution
APE1	Rabbit pAb		Novus	NB100-101	1:300
ATM	Rabbit mAb	Y170	Abcam	Ab32420	1:100
NBN	Rabbit pAb		Sigma	HPA001429	1:150
PMS2	Mouse mAb	A16-4	BD Pharmingen	556415	1:200
PTEN	Rabbit mAb	D4.3	Cell signalling	9188	1:150

pAb=polyclonal antibody. mAb=monoclonal antibody.

2.3.3 Antibody optimisation

Optimal concentration and conditions for staining were ascertained for each antibody using standard IHC technique, as described in Section 2.3.1. The 'Human Protein Atlas' on-line resource (www.proteinatlas.org) was used to identify tissues that are known to express the protein of interest and therefore would be suitable for antibody optimisation. The tissues and antibody concentrations used for optimisation are shown in Table 2.3.

Table 2.3. List of tissues used for optimisation and antibody concentrations

Antibody	Tissue	Concentration
APE1	Glioma (full-face section)	1:100, 1;150, 1:200, 1:300, 1:400, 1:500
ATM*	Glioma (full-face section) Breast TMA	1:75, 1:100, 1:125, 1:150, 1:200
PTEN	Gastric cancer TMA Ovarian TMA	1:100, 1:300, 1:500 1:150, 1:200
NBN	Glioma (full-face section)	1:150, 1:200, 1:300
PMS2	Colorectal cancer (full-face section) Glioma (full-face section)	1:100, 1:150, 1:200, 1:300, 1:500 1:100, 1:150, 1:200

*ATM has previously been optimised in breast tissue (Abdel-Fatah et al., 2014a). TMA=tissue microarray.

All sections were reviewed by an expert histopathologist (IS or TMA) after each round of optimisation to decide if further optimisation was required or to identify the optimal concentration for staining the cohort of glioma sections.

2.3.4 Inclusion of both positive and negative control sections

A negative control section was included which was exposed to antibody diluent only, with the omission of primary antibody. An absence of staining demonstrates that the staining observed in the presence of primary antibody is due to the antibody itself and not due to other reagents used.

Positive control sections were also included. In each run a full-face section (or TMA) was included that was known to stain positive for the marker of interest (antibody specific positive control). As the sections were stained in multiple batches the same positive control section was included in each run to enable consistency to be evaluated.

Beta 2 microglobulin (B_2M) is widely expressed in a variety of tissues, including breast cancer. A breast cancer TMA section was therefore also included as a positive control section to demonstrate that the actual technique of IHC had been successful.

2.3.5 Histopathological scoring of glioma sections

Scoring of all glioma sections was performed with an expert histopathologist (TMA). Whole field inspection of the core was undertaken and the intensity of nuclear staining grouped as follows: 0 = no staining, 1 = weak staining, 2 = moderate staining and 3 = strong staining. The percentage of each staining category was estimated, the values summed and a H score calculated (ranging from 0-300). The presence of any cytoplasmic staining was also recorded.

2.3.6 Patient demographics

Forty-three (70.5%) high grade patients were male and 19 patients (31.1%) were alive at the end of the study. Eighteen patients (29.5%) had been diagnosed with a glioma prior to inclusion in the study; 4 (6.6%) had already received radiotherapy and 3 (4.9%) had previously received chemotherapy. Median age at trial histology was 54 years (range 22-81 years). In total 52 (85.2%) patients were known to have received radiotherapy during the course of their illness and 39 (63.9%) were known to have received chemotherapy. The median number of lines of chemotherapy given was one (range 0-4). Baseline demographic and treatment data for the Nottingham Cohort is shown in Table 2.4.

2.3.7 Ethical approval

The immunohistochemistry study has been approved by the Leicestershire, Northamptonshire and Rutland Regional Ethics Committee (Reference number 08/H0406/102).

Table 2.4. Baseline demographic and treatment data for the Nottingham Cohort

Clinicopathological variable	Number (%)
Gender	
Male	39 (63.9)
Female	22 (36.1)
Survival Status	
Alive	19 (31.1)
Deceased	42 (68.9)
Previous diagnosis of glioma	
None	43 (70.5)
Grade 2	8 (13.1)
Grade 3	1 (1.6)
Grade 3-4	1 (1.6)
Grade 4	1 (1.6)
Low grade on imaging	7 (11.5)
Trial pathology	
Grade 3 (n=20)	
Anaplastic astrocytoma	12 (19.7)
Anaplastic oligoastrocytoma	4 (6.6)
Anaplastic oligodenodroglioma	4 4 (6.6)
Grade 4 (n=41)	
Glioblastoma	41 (67.2)
Surgery	
Biopsy	12 (19.7)
Debulking	49 (80.3)
Radiotherapy given	
Yes	52 (85.2)
No	2 (3.3)
Unknown	7 (11.5)
Chemotherapy given	
Yes	16 (26.2)
No	39 (63.9)
Unknown	6 (9.8)

2.3.8 Statistical analysis of IHC data

Baseline demographic data was collected, in addition to treatment and survival information, from the NUH patient computer records system (NotIS). Data analysis was performed in SPSS (Version 20, Chicago, USA). Categorical variables are expressed as a number and percentage and continuous variables as a median and range. Xtile (Version 3.6.1, Yale University, USA) was used to dichotomise expression of APE1 (low ≤160, high >160), NBN (low ≤85, high >85), nuclear PTEN (low ≤50, high >50), cytoplasmic PTEN (low ≤10, high >10) and PMS2 (low ≤120, high >120). Kaplan Meier survival curves were constructed for each marker and the log rank test used to determine the survival difference between groups. Chi squared tests (with or without Yates' continuity correction, as appropriate) were used to assess the association between two categorical variables. The Mann Whitney U test was used to compare continuous, non-normally distributed variables between two groups. Cox multivariate analysis was performed to determine independent predictors of survival. Statistical significance was defined as $p \leq 0.05$.

2.4 Cell line studies

2.4.1 Cell lines and culture requirements

A172, U87MG, LN18 and LN229 human glioblastoma cell lines were purchased from American Type Culture Collection (ATCC, Manassas, USA) and were grown as per the ATCC recommendations. The U87MG cell line was grown in Eagle's Minimum Essential Medium (EMEM with Earle's salts, non-essential amino acids and sodium bicarbonate) (Sigma, Dorset, UK) supplemented with 5ml L-glutamine (200mM) (Sigma) and 10% foetal bovine serum (FBS) (Sigma). The A172 cell line was maintained in Dulbecco's Modified Eagle Medium (DMEM with 4500mg/L glucose, L-glutamine, sodium pyruvate) with the addition of 10% FBS. LN229 and LN18 were grown in DMEM media, as above, with the addition of 5% FBS. All media was supplemented with 5ml of 1% penicillin/streptomycin (10,000 units penicillin and 10mg streptomycin/mL) (Sigma). All cell lines were maintained in a humidified incubator at 37°C with 5% carbon dioxide and grown as an adherent culture.

A frozen stock of all cell lines was acquired at an early passage. One million cells were harvested and re-suspended in 1ml of freezing medium. Freezing vials were transferred immediately to a Mr. Frosty™ freezing container (Thermo Fisher Scientific) and held at -80°C for at least 24 hours. For long-term storage, cell lines were stored in liquid nitrogen. Freezing medium for all cell lines (with the exception of LN18) consisted of 5% dimethyl sulfoxide (DMSO) (Sigma) and 95% standard growth media. Freezing medium for the

LN18 cell line consisted of 5% DMSO and 95% DMEM (supplemented with 10% FBS).

All cell lines were subcultured two to three times per week. Firstly, media was removed from the flask. Cells were then washed with Dulbecco's phosphate buffered saline (DPBS) (Sigma) to remove any dead cells or cell debris. Trypsin-EDTA solution (Sigma) was added and incubated at 37°C for three to four minutes until the adherent monolayer had detached. Media was added at a ratio of 1:4 (trypsin-EDTA:media) to neutralise the trypsin. The suspension was centrifuged at 2000 rpm for five minutes to pellet the cells. The cell pellet was re-suspended in normal culture medium and re-seeded into tissue culture flasks at a dilution between 1:5 to 1:10.

Logarithmic phase cells were used for all cell line-based experiments (i.e. 70-80% confluent). All experiments were done in at least triplicate unless otherwise stated. Cell lines were not used beyond passage 50.

2.4.2 DNA repair inhibitors and other compounds

The ATM inhibitor KU55933 [2-(4-Morpholinyl)-6-(1-thianthrenyl)-4H-pyran-4-one] was purchased from Tocris Bioscience (Bristol, UK) (Figure 2.2).

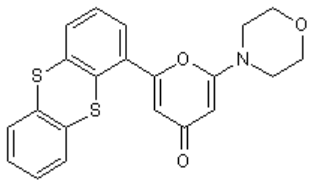


Figure 2.2. Chemical structure of KU55933

KU55933 was dissolved in DMSO and stored in aliquots at -20°C. For all inhibitor-based assays, DMSO was included as a vehicle control at the same (or highest) concentration used when drug was administered. Concentration ranges for inhibitory compounds used were based on previous work by Dr Nada Albarakti as part of her PhD studies (Albarakati et al., 2015). Methyl methanesulphonate (MMS) (Sigma) was used a positive control and stored at 4°C.

2.4.3 Cell line characterisation studies

2.4.3.1 Polymerase chain reaction

The Polymerase chain reaction (PCR) can be used to amplify specific sequences within a cDNA template. In this study, end-point PCR has been used to determine expression of single genes. Real-time PCR has been used to quantify differences in DNA repair gene expression between cell lines using the RT^2 DNA Profiler PCR array (#PAHS-042Z, Qiagen, Manchester, UK). Both techniques require the extraction of RNA from the cell line of interest and the production of cDNA. This cDNA is then mixed with sequence specific oligonucleotides and a heat-stable DNA polymerase prior to undergoing thermal cycling. This results in the generation of a PCR product.

In end-point PCR, detection and quantification of the PCR product is performed by gel electrophoresis, after completion of all the cycles of PCR. In real-time quantitative PCR, the amount of PCR product is measured after each cycle using fluorescent reporter dyes, such as SYBR® Green. The amount of

fluorescent signal generated by the reporter dye is directly proportional to the number of PCR product amplicons generated.

2.4.3.2 RNA extraction

$3x10^6$ cells were harvested and washed in DPBS. RNA extraction was performed using the RNeasy Mini Kit (Qiagen). 350µl of Buffer RLT and 350µl of 70% ethanol were added to the cell pellet. The mixture was homogenised and transferred to a RNeasy Mini spin column. The tube was spun at 10,000 rpm for 15 seconds and the flow-through removed. Subsequently, on column DNase digestion was performed using the RNase-Free DNase Set (Qiagen). Firstly, 350µl of Buffer RW1 was added to the mix, the sample was then centrifuged as before and the flow-through removed. Next, 10µl of DNase I stock solution was the added to 70µl of Buffer RDD to make a DNase incubation mix. The DNase incubation mix was transferred to an RNeasy column membrane and incubated at room temperature for 15 minutes. Subsequently, 350µl of Buffer RW1 was added to the DNase incubation mix, the sample centrifuged, and flow-through removed. To wash the RNA, 500µl of Buffer RPE was added to the membrane, the sample centrifuged and then this step was repeated. 50µl of RNase-free water was added to the column membrane and the sample centrifuged to elute the RNA. RNA purity and concentration was assessed using a microvolume spectrophotometer (NanoDrop, Wilmington, USA). Samples were stored at -80°C prior to cDNA synthesis.

2.4.3.3 cDNA synthesis

cDNA synthesis was performed using the RT^2 First Strand Kit (Qiagen). 10µl of genomic DNA elimination mix was incubated at 42°C for 5 minutes and then held at 4°C for at least 1 minute. The composition of the elimination mix is shown in Table 2.5.

Table 2.5. Composition of genomic DNA elimination mix.

Reagent	Amount
RNA	0.5µg
Buffer GE	2µl
RNase-free water	Made up to 10µl

10µl of reverse transcription mix (4µl 5X Buffer BC3, 1µl Control P2, 2µl RE3 Reverse Transcriptase Mix and 3µl RNase-free water) was added to the DNA elimination mix and incubated at 42°C for 15 minutes. The reaction was stopped by a 5 minute incubation at 95°C. 91µl of RNase-free water was added to the new mix. All incubation steps were performed using a Techne TC-312 thermal cycler (Stone, UK). Samples were stored at -80°C prior to PCR analysis.

2.4.3.4 Generation of no reverse transcriptase controls

No reverse transcriptase controls were made from RNA of each cell line to assess for genomic DNA contamination. The protocol was followed as per 2.4.3.3 except RE3 Reverse Transcriptase Mix was replaced with 2µl RNase/DNase free water.

2.4.3.5 End-point PCR analysis

2.4.3.5.1 Primer selection

All primers were purchased from Qiagen and reconstituted in 1.1ml RNase/DNase free water. Primers were stored at -20°C. Table 2.6 shows details of all primers used.

Table 2.6. List of QuantiTect® primers used for end-point PCR

Gene symbol	Catalogue number
ATM	QT00061593
PTEN	QT01676969
Beta actin	QT01680476

2.4.3.5.2 Endpoint PCR protocol

cDNA for each cell line was diluted to 100ng/µl in RNase/DNase free water. A PCR Mastermix was made (as shown in Table 2.7) for each primer.

Table 2.7. Components of PCR Mastermix

Reagent	Volume*
5x GoTaq green buffer	10µl
10mM dNTP mix	1µl
QuantiTect® primer mix	5µl
GoTaq DNA polymerase	0.25µl
RNase/DNase free water	31.75µl

*Volumes are for a single PCR reaction. All components were purchased from Promega (Southampton, UK) except QuantiTect® primers (Qiagen).

For each sample to be run, 2µl of 100ng/µl cDNA was added to 48µl of the relevant Mastermix. No template controls were included for each primer and consisted of 2µl DNase/RNase free water (in place of cDNA) added to 48µl mastermix. This enables the identification of any cross-contamination of reagents or surfaces. A no reverse transcriptase control was included for each

cell line and consisted of 2µl of no reverse transcriptase solution (2.4.3.4), in place of cDNA, added to 48µl of mastermix.

Samples were transferred to a Techne TC-312 thermal cycler for PCR and initially exposed to 95°C for 30 seconds, followed by 35 cycles of: 95°C for 30 seconds, 60°C for 30 seconds and 72°C for 90 seconds.

2.4.3.5.3 Electrophoresis of PCR products

Electrophoresis of the PCR products was performed. 2% agarose gels were made using 4g agarose (Ultrapure, Invitrogen) and 200ml 1x TAE (Tris base, acetic acid and EDTA) solution per gel. Agarose/TAE solution was heated until boiling in a microwave and 6µl ethidium bromide (100mg/ml stock, Sigma) added. Gels were poured and allowed to solidify at room temperature. Gels were loaded into an electrophoresis tank filled with 1x TAE solution and 20µl of PCR product loaded into each well. 0.5µg low molecular weight DNA ladder (New England Biolabs, Hitchin, UK) was loaded per gel. Gels were run at 100V for 60 minutes prior to visualisation using a UV transilluminator (BioRad).

2.4.3.6 Real-time PCR using RT^2 Profiler PCR Array

25µl of PCR components mix (1350µl 2x RT^2 SYBR Green Mastermix, 102µl cDNA synthesis reaction and 1248µl RNase-free water) was added to each well of the 96 well RT^2 DNA Profiler PCR array. Plates were sealed using optical adhesive film and centrifuged at 1000*g* prior to real-time PCR on the

Applied Biosystems 7500 FAST cycler (Foster City, USA). All components for PCR analysis were purchased from Qiagen.

2.4.3.6.1 Data analysis

Data analysis was performed using the Qiagen web-based data analysis centre (http://pcrdataanalysis.sabiosciences.com/pcr/arrayanalysis.php). Data was pre-prepared using a Microsoft Excel template prior to uploading to the data analysis centre. PCR array reproducibility and sample quality data was scrutinised for each cell line. Samples were required to pass checks for PCR array reproducibility, reverse transcriptase (RT) efficiency and genomic DNA contamination in order to be included in subsequent analysis. A summary of the quality assurance criteria are shown in Table 2.8.

Table 2.8. Summary of quality assurance criteria used for real-time PCR analysis

Quality control check	Criteria	Outcome
PCR reproducibility	AVG PPC C_t is 20±2 and no 2 arrays have PPC C_t >2 away from one another	Pass
RT efficiency	ΔC_t (AVG RTC – AVG PPC) ≤5	Pass
Genomic DNA contamination	C_t (GDC) ≥35	Pass

AVG=average. C_t=threshold cycle. GDC=genomic DNA control. PCR=polymerase chain reaction. PPC=positive PCR control. RT=reverse transcription. RTC=reverse transcription control.

Housekeeping gene selection was performed by manual review of the C_t values for all five housekeeping genes across all cell lines and experimental replicates. An ideal housekeeping gene demonstrates small changes in C_t value across sample groups. Expression of all genes on the array was normalised against the selected housekeeper. A fold change was calculated using the formula:

2^(-Delta Delta Ct). This is the normalised gene expression in the Test sample 2^(-Delta Ct) divided by the normalised expression in the Control sample 2^(-Delta Ct). Differences in gene expression between two cell lines were expressed as a fold regulation, rather than a fold change, as this has greater biological meaning. Fold change values greater than 1 represent a positive- or up-regulation, and are equal to the fold regulation. Fold change values less than 1 represent a negative- or down-regulation, and are equal to the negative inverse of the fold change. Data was visualised by scatterplots comparing the control cell line against the test samples. Genes that demonstrate either a two-fold increase or two-fold decrease in expression were identified.

2.4.4 Western blot analysis

2.4.4.1 Cell lysate preparation

Two million cells were harvested and lysed in RIPA buffer (88µL) with the addition of protease cocktail inhibitor (10µL, 1:10 dilution) and phosphatase inhibitors 2 & 3 (1µL each), and stored at -20°C. All lysate mix components were purchased from Sigma.

2.4.4.2 Sample preparation

Lysates were centrifuged at 13000 rpm for 20 minutes at 4°C and the supernatant collected. Samples for analysis were prepared by the addition of 5µl cell lysate, 2.5µl sample buffer (NuPAGE® LDS Sample Buffer), 1µl reducing agent (NuPAGE® Reducing Agent) and 6.5µl dH_20. Samples were heated at 70°C for 10 minutes.

2.4.4.3 Western blot analysis

NuPAGE® 4-12% Bis-Tris Mini Gels were loaded into a Xcell SureLock® Mini-Cell with 1X MOPS SDS Running Buffer (Thermo Fisher Scientific). 500µl NuPAGE® Antioxidant was added to 200ml 1X running buffer in the upper tank chamber. Samples and protein standards were loaded and the gel run at 125V for 90 minutes. At least one protein standard was included on each gel. Protein standards used include: Precision Plus Protein™ All Blue Pre-stained Protein Standard (BioRad, Hemel Hempstead, UK), MagicMark™ XP Western Protein Standard (Life technologies, Paisley, UK) or HiMark™ Pre-stained Protein Standard.

Filter paper, nitrocellulose membranes (GE Healthcare Life Sciences, Little Chalfont, UK) and sponges were soaked in transfer buffer (100ml methanol, 50ml NuPAGE® Transfer Buffer (20X), and 1ml NuPAGE® Antioxidant, made up to 1L with cold dH_20). Cassettes were assembled containing the pre-soaked sponges, filter paper, membrane and gel. The cassette was filled with transfer buffer and the outer tank with dH_20. Unless otherwise stated all western blot equipment and reagents were purchased from Thermo Fisher Scientific.

Proteins were electro-blotted to membranes at 25V for 90 minutes. Membranes were removed and placed in 5% milk solution (2.5g fat-free milk powder in 50ml PBS-T [1ml Tween 20 in 1L PBS]) for 60 minutes at room temperature. Primary antibody was added (diluted in 2ml 5% milk) and incubated at 4°C overnight. The primary antibodies used in this study are shown in Table 2.9, alongside the optimal concentration.

Table 2.9. List of antibodies used for western blot analysis

Antigen	Molecular weight	Antibody	Source	Catalog #	Dilution
ATM	350 kDa	Rabbit mAb	Abcam	ab32420	1:1000
APE1	37 kDa	Rabbit pAb	Novus	NB100-101	1:1000
DNA-PKcs	460 kDa	Rabbit mAb	Abcam	ab32566	1:1000
NBN	95 kDa	Rabbit pAb	Sigma	HPA001429	1:500
p53	53 kDa	Mouse mAb	Leica Biosystems	NCL-L-p53-D07	1:500
PI3K (p85 subunit)	85 kDa	Mouse mAb	Abcam	ab86714	1:1000
PMS2	110 kDa	Mouse mAb	BD Pharmingen	556415	1:500
PTEN	55 kDa	Rabbit pAb	Cell signalling	9188	1:500

mAb=monoclonal antibody; pAb=polyclonal antibody.

Membranes were washed with PBS-T (x3 10 minutes) prior to incubation with antibody against a housekeeper gene diluted in 5% milk solution for 1 hour at room temperature. Housekeeper genes were selected to vary in molecular weight to the target antigen and to have been raised in a different species to the primary antibody to aid detection (Table 2.10).

Table 2.10. List of housekeeper antibodies used for Western blot analysis

Antigen	Molecular weight	Antibody	Source	Catalog #	Dilution
Beta actin	42 kDa	Mouse	Sigma	A5441	1:10000
Beta tubulin	50 kDa	Rabbit	Abcam	ab6046	1:2000
GAPDH	37 kDa	Rabbit	Abcam	ab9485	1:2500

Following further PBS-T washes (x3 10 minutes) membranes were incubated in the dark with 1:10000 infrared dye-labelled secondary antibody solution (IRDye 800CW Donkey Anti-Rabbit IgG and IRDye 680CW Donkey Anti-

Mouse IgG, Li-Cor, Cambridge, UK) for 1 hour. After washing (x2 10 minutes with PBS-T, x1 10 minutes with PBS) membranes were scanned at 700 and 800nm on a Li-Cor Odyssey Fc Infrared Imager to determine protein expression.

2.4.4.3.1 Statistical analysis

Quantification of protein expression was performed using the Image Studio Lite version 3.1 (Li-Cor). For each cell line the signal intensity for the protein of interest was normalised against expression of the housekeeper. This normalised expression value for each cell line was then expressed relative to a control cell line. Normalisation was performed in Microsoft Excel 2010. Graphical representation and one way ANOVA assessment for significant differences in expression between cell lines was performed in GraphPad Prism 6 (GraphPad, La Jolla, USA).

2.4.5 Cell survival assays

2.4.5.1 Plating efficiency

80% confluent U87MG, A172, LN229 and LN18 cells were seeded into either 6 well plates or T25 tissue culture flasks at densities of 100-5000 cells per well/flask. Due to a failure of the cells to grow at low density the following protocols were tested:

- 100, 200, 300, 400, 500, and 600 cells per well in 6 well plates
- 1000 and 2000 cells per T25 tissue culture flasks
- 1000, 2000 and 5000 cells per well in 6 well plates

Plates or flasks were incubated at 37°C and 5% CO_2 for 14 days. After this time, culture vessels were removed, media discarded and wells washed with PBS. Cells were stained with 1ml of crystal violet (10% [w/v] in 70% ethanol) (Sigma) for 5 minutes. Crystal violet was removed; the plates washed with tap water and left to dry overnight. Colonies of at least 50 cells were counted. Plating efficiency (PE) was calculated using the following equation: PE = (no. of colonies formed/ no. cells seeded) x100. All experiments were performed in duplicate.

2.4.5.2 AQueous non-radioactive cell proliferation assay (MTS assay)

Cells were seeded at a density of 1000 cells per well (in 100µL volume) in 96 well plates and allowed to adhere overnight. Inhibitory compounds were added at a range of concentrations, after 16 hours, in a volume of 20µL and plates incubated for a total of 7 days. A DMSO vehicle control (0.1% final

concentration) was also included for each cell line. On Day 7, 20µL of CellTiter 96® AQ_{ueous} One Solution Reagent (Promega) was added to each well and incubated for 2 hours. Absorbance was measured at 490 nm using a FLUOstar OPTIMA microplate reader (BMG Labtech, Aylesbury, UK). The absorbance, relative to the DMSO control, was calculated for each drug concentration. All experiments were performed in triplicate at least three times, unless otherwise stated. Data analysis was performed in Microsoft Excel 2010 and GraphPad Prism 6. Two way ANOVA was used to assess for statistically significant differences in survival between cell lines for the various drug concentrations.

2.4.5.3 Growth assay with the addition of KU55933

100,000 cells were seeded into T25 tissue culture flasks and incubated overnight. After 16 hours KU55933 was added to a final concentration of 10µM (or 0.1% DMSO as a vehicle control). After 48 hours cells were harvested by trypsinisation and counted using a haemocytometer. The number of cells per ml were determined. This process was repeated every 24 hours to a maximum of 144 hours. Data was visualised in GraphPad Prism 6 (GraphPad, La Jolla, USA). This experiment was performed once in duplicate.

2.4.6 Functional consequences of DNA repair inhibitor treatment in glioma cell lines

2.4.6.1 Neutral COMET assay

150,000 cells per well were seeded in 6 well tissue culture plates and incubated at 37°C and 5% CO_2 overnight. KU55933 was added at a concentration of 10µM (or 0.1% DMSO vehicle control) after 16 hours. 100µM MMS was included as a positive control. Cells were harvested by trypinisation at 6, 24 or 48 hours post drug administration and re-suspended in 250µl PBS. 12.5µl of the cell suspension was combined with 125µl of molten agarose (Trevigen, Gaithersburg, USA) and 75µl of the mixture was pipetted rapidly onto pre-warmed COMET slides (Trevigen). Slides were incubated at 4°C for 30 minutes prior to immersion in cold lysis buffer (Trevigen) for 1 hour. Slides were then transferred to pre-chilled neutral electrophoresis buffer (60.57g TRIS base and 204.12g Sodium acetate dissolved in 500ml deionised water, diluted 1:10, pH9) for 30 minutes. Electrophoresis was performed in neutral electrophoresis buffer at 20V and 4°C for 1 hour. Excess buffer was then removed and slides immersed in DNA precipitation buffer (6.7ml 7.5M ammonium acetate, 43.3ml 95% ethanol) for 30 minutes at room temperature. Slides were fixed in 70% ethanol for 20 minutes at room temperature and then stored at 4°C overnight. Prior to visualisation using an Olympus BX40 microscope, 50µl SYBR® gold (Diluted 1:10000 in PBS, Trevigen) was added to each sample. 50 comets were scored per sample using Comet Assay III/IV image analysis software (Perceptive Instruments, Bury St Edmonds, UK). Graphical representation and statistical analysis was performed in GraphPad Prism 6 (GraphPad, La Jolla, USA). Two way ANOVA was used to determine

whether statistically significant differences exist in mean tail moment between untreated and treated cells for each cell line. All experiments were done in duplicate three times.

2.4.6.2 Gamma H2AX and cell cycle analysis by flow cytometry

150,000 cells were seeded in 6 well plates and allowed to adhere overnight. KU55933 was added after 16 hours at a concentration of 10µM (or 0.1% DMSO vehicle control). Cells were harvested by trypsinisation and centrifuged at 1000 rpm for 5 minutes at 24 hours post drug exposure. Cell pellets were then fixed in 1ml 70% ice cold ethanol and stored at 4°C until required. On further processing suspensions were centrifuged at 1000 rpm for 5 minutes and the supernatant removed. The Flow Cytometry H2AX Phosphorylation Assay Kit (Merck Millipore, Watford, UK) was used for further processing of samples.

Cell pellets were washed twice in PBS and gently resuspended in 50µL 1X Permeabilisation Solution. Samples were then incubated with 3.5µL of fluorescein isothiocyanate (FITC) anti-phospho-Histone H2A.X (ser139) for 20 minutes. Excess FITC labelled antibody was removed by washing with 200µL 1X Wash Solution and samples spun at 1000 rpm for 5 minutes. The supernatant was removed and the cell pellet re-suspended in 100µL of 100µg/ml RNAse (Qiagen) in PBS. After 15 minutes incubation, 400µL propidium iodide (PI) (2.5µg/ml, Merck Millipore) was added and the samples incubated for a further 15 minutes in FACS tubes. Samples were analysed on a Cytomics F500 flow cytometer (Beckman Coulter, High Wycombe, UK) with

emission data for PI collected using a 620nm bandpass filter (FL3) and for FITC-anti-phospho-Histone H2A.X with a 575nm bandpass filter (FL1). At least 10,000 events from each sample were analysed. All experiments were done in duplicate three times. An unstained control was included for each cell line, alongside PI only and FITC anti-phospho-Histone H2A.X only samples, and a FITC IgG isotype control. 100μM MMS was included as a positive control.

Weasel (Victoria, Australia) flow cytometry analysis software was used for data analysis. Graphical representation and statistical analysis was performed in GraphPad Prism 6. Two-way ANOVA was used to determine whether statistically significant differences exist in γH2AX levels between untreated and treated cells for each cell line. The percentage of cells in each phase of the cell cycle was determined for each cell line in both the untreated and treated state. Two way ANOVA was used to determine whether statistically significant differences exist in the distribution of cells throughout the cell cycle after treatment with KU55933.

2.4.6.3 Apoptosis detection by FITC-Annexin V flow cytometry

150,000 cells were seeded into 6 well plates and allowed to adhere overnight. KU55933 was added at a concentration of 10μM after 16 hours (or 0.1% DMSO vehicle control). 24 hours after drug administration, media was collected from the wells. Cells were then harvested by trypsinisation and the cell suspension added to the media removed from the flask. This suspension was spun at 1000 rpm for 5 minutes and the cell pellet washed twice in ice cold

PBS. All reagents used were from the FITC-Annexin V Apoptosis Detection Kit I (BD Pharmingen, San Jose, USA) unless otherwise stated. Cell pellets were resuspended in 100µl Annexin V binding buffer, transferred to FACS tubes and incubated for 10 minutes. 2.5µl propidium iodide (PI) and 2.5µl Annexin V were added to each tube and samples incubated in the dark for 15 minutes. Samples were analysed on a Cytomics F500 flow cytometer (Beckman Coulter) with emission data collected using a 575nm bandpass filter (FL1) for annexin V-FITC and a 620nm bandpass filter (FL3) for PI. At least 10,000 cells from each sample were analysed. Untreated controls were included for each cell line and each time point. An unstained, PI only and Annexin V only control were also included for each cell line. All experiments were done in duplicate three times.

Weasel flow cytometry analysis software was used for data analysis. The percentage of apoptotic cells was ascertained (FITC-Annexin V positive, PI positive and FITC-Annexin V positive, PI negative) and compared for untreated and treated samples by two-way ANOVA. Graphical representation and statistical analysis was performed in GraphPad Prism 6.

2.4.7 Production of inducible PTEN knockdown glioma cell lines by lentiviral transduction

2.4.7.1 Cell lines and culture requirements

LN18 and LN229 PTEN proficient glioma cell lines were grown in DMEM with 4.5g per litre glucose (Sigma) and supplemented with 10% Tet System Approved FBS (Clontech, Mountain View, USA). Cells were grown in a monolayer and in a humidified incubator with 5% CO_2 at 37ºC.

2.4.7.2 Determination of puromycin sensitivity

Sensitivity of the LN18 and LN229 cell lines to puromycin was determined using an alamarBlue® Cell Viability Assay (Thermo Fisher Scientific) and was performed by Mrs Pamela Collier. LN18 and LN229 cell lines were seeded in replicates of 6 in 96 well plates at a density of 3000 cells per well, and left overnight to adhere. Media containing Puromycin at various concentrations (0.125 to 32µg/µl) was added to the plate after 24 hours. Samples were incubated for 72 hours prior to the addition of alamarBlue® reagent (10% of the sample volume). Plates were then incubated for a further hour at 37°C. Fluorescence was read on a microplate reader (FlexStation384) and plotted against Puromycin concentration in GraphPad Prism 6.

2.4.7.3 shRNA plasmid preparation and recombinant lentiviral production using Lenti-X system

shRNA plasmid preparation, restriction digests and recombinant lentiviral production were performed by Dr Rachel Abbotts as part of her PhD project and are described in detail in (Abbotts, 2013). Lentiviral-mediated gene

knockdown was performed using TRIPZ Inducible Lentiviral shRNA (Thermo Fisher Scientific) (PTEN clone IDs V2THS_92317, V2THS_231477 and V3THS_213158). Clones for use in this current study were selected based on their ability to induce successful PTEN knockdown in the previous work by Dr Rachel Abbotts.

In summary, shRNAs were supplied as *Escherichia coli* frozen glycerol stock which were inoculated into culture and incubated at 37°C in a shaking incubator overnight. Cultures were harvested and plasmid purification performed using the Qiagen Plasmid Maxi Kit and the QIAGEN-tip 500. DNA concentration was determined using a microvolume spectrophotometer (NanoDrop). 2µg plasmid DNA was used for restriction digest to confirm plasmid amplification.

Recombinant lentiviral production was performed by Mrs Pamela Collier (Division of Pre-Clinical Oncology, School of Medical and Health Sciences, University of Nottingham) using the Lenti-X-high-titer lentiviral packaging system (Clontech) for Dr Rachel Abbotts. In brief, for each plasmid two reaction tubes were prepared, tube 1 (DNA solution, containing plasmid DNA) and tube 2 (polymer solution). The contents of both tubes were mixed and incubated at room temperature for 10 minutes. The DNA-polymer mix was added to 80-90% confluent Lenti-X 293T cells and cultures incubated at 37°C. After 48 hours supernatants were harvested and virus production confirmed using Lenti-X GoStix. Virus aliquots were stored at -80°C until required.

2.4.7.4 Target cell transduction using lentiviral vectors

Transduction of LN18 and LN229 target cell lines was performed by Mrs Pamela Collier. Target cell lines were plated into 6 well tissue culture plates at a density of $3x10^5$ cells per well in 2ml of culture medium. The following day viral supernatant was added to the target cell culture to obtain a 'multiplicity of infection' (MOI) of 1. An MOI of 1 implies one lentiviral particle per target cell and assumes the doubling time of target cells overnight results in a density of $6x10^5$ cells per well. In order to reduce charge repulsion between the virus and cell membrane, polybrene (Sigma) was added at a final concentration of 8mg/ml. After 48 hours incubation at 37°C, virus-containing media was replaced with fresh culture media. Once cells reached 70-80% confluence they were sub-cultured into T25 tissue culture flasks. After 48 hours, puromycin selection was initiated at a concentration of 2µg/ml, as determined by the puromycin sensitivity assay (see section 2.4.7.2). For initial selection cells were grown in puromycin for 1 week and then maintained in puromycin for 48 hours during each one week period.

2.4.7.5 Doxycycline induction of target gene knockdown

The TRIPZ vector contains a tetracycline-inducible promotor (TRE) that drives expression of both the shRNAmir and a TurboRFP reporter. The presence of the TurboRFP reporter allows monitoring of shRNA expression under a fluorescence microscope, a process that was performed on at least a weekly basis. Cells were grown in standard growth media with the addition of 0.5µg/ml doxycycline (Clontech) to induce shRNA expression and subsequent knockdown of PTEN. Doxycycline was replaced every 72 hours. For each

cell line and construct, a flask was maintained without doxycycline to act as control.

2.4.7.6 Confirmation of target gene knockdown

Cell lysates were made at both day 7 and day 14 post doxycycline administration and Western blots performed as in Section 2.4.4. PTEN signal was normalised against the beta actin signal and then expressed relative to non-doxycycline induced samples for each construct. End-point PCR (see Section 2.4.3.5) was also performed for the construct showing the greatest level of PTEN protein knockdown in both cell lines at fourteen days post doxycycline administration. In addition, ATM expression was determined by Western blot and end-point PCR analysis in both cell lines, with and without doxycycline for induction of *PTEN* knockdown.

2.4.7.7 MTS assay to confirm sensitivity of PTEN knockdown glioma cell lines to ATM inhibitor treatment

The standard MTS assay protocol was followed as described in Section 2.4.5.2, with the following modifications. For each cell line two 96 well tissue culture plates were seeded. Prior to the addition of inhibitory compound, media was removed and replaced with 120µl fresh culture media (with or without 0.5µg/ml doxycycline) plus KU55933 at the desired concentration. For each cell line one plate received additional doxycycline (0.5µg/ml) after 72 hours; the second plate received the same volume of culture media. As in the normal culture, setting doxycycline was replaced every 48-72 hours; these two protocols were performed to determine whether the addition of extra

doxycycline influenced the results of the MTS assay. All experiments were done in triplicate three times.

2.4.7.8 MTS growth assay to determine the effect of PTEN re-expression on growth after doxycycline removal

Cells were seeded at a density of 1000 cells per well in 96 well tissue culture plates and the standard MTS protocol followed. For each cell line the following were seeded: 1) cells never exposed to doxycycline, 2) cells previously exposed to doxycycline, and 3) cells exposed to (and to be maintained in) doxycycline. Absorbance was read every 48 hours to a maximum of 192 hours. Doxycycline (0.5µg/ml) was added every 48 hours to cells to be maintained in doxycycline.

In addition, lysates were made every 48 hours after doxycycline withdrawal (Group 2 cells) for Western blot analysis to ascertain the timescale for PTEN re-expression.

2.4.7.9 MTS assay to determine the influence of doxycycline on non-transduced LN18 and LN229 cell lines

Cells were seeded at a density of 1000 cells per well in 96 well tissue culture plates and the standard MTS protocol followed. For each cell line both non-transduced and transduced (doxycycline exposed) cells were seeded. Absorbance was recorded every 48 hours to a maximum of 192 hours. Doxycycline (0.5µg/ml) was added to transduced (doxycycline exposed) cells every 48 hours.

CHAPTER 3

DNA repair gene expression profiling in glioblastoma

3 Results

3.1 Introduction

Patients with glioblastoma have a poor prognosis, with a median survival of less than 18 months (Stupp et al., 2005). In common with many other tumour sites, however, prognosis does vary between individuals. The ability to identify patients who have either a good or bad prognosis can both assist patients in planning for the future but may also influence disease management decisions. This strategy is commonly used in oncological practice, for example where patients with high risk disease features are offered adjuvant chemotherapy (Labianca et al., 2010, EBCTCG, 2005, Arriagada et al., 2010). In other situations, prognosis may be so poor that the appropriate management decision is to spare the patient intensive treatment regimens and focus on good palliative care. Over recent years, work has focused on molecular markers that may predict response to treatment, with the intention to tailor treatment to the individual patient (Debiec-Rychter et al., 2004, Shaw et al., 2013, Maemondo et al., 2010), for example the use of PARP inhibitors in BRCA deficient breast and ovarian cancers (Underhill et al., 2011).

This chapter explores the prognostic significance of the expression of 188 DNA repair genes in two large, publically available, gene expression datasets. The most important DNA repair genes capable of predicting prognosis in GBM are identified and a prognostic index has been developed to separate patients into three distinct prognostic groups. Furthermore, Artificial Neural Network

(ANN) analysis techniques have been employed to identify genes that interact with these key DNA repair genes and that may warrant further investigation in future studies.

3.1.1 Prognostic scores in high grade glioma incorporating clinical variables

Although prognosis for patients with GBM is poor, and the disease widely accepted to be universally fatal, some patients fair better than others. The Medical Research Council (MRC) prognostic index was first published in 1990 and was devised from, and validated in, two MRC trials that recruited patients between 1979 and 1986. The index incorporates information on four clinical variables: age, World Health Organisation (WHO) clinical performance status, extent of neurosurgery and a history of fits. The score breakdown for the MRC prognostic index is shown in Table 3.1. Patients with a higher prognostic index score have a worse prognosis than individuals with a lower score (Latif et al., 1998, MRC, 1990).

Table 3.1. Medical Research Council (MRC) prognostic index

Prognostic factor	Score
Age (years)	
<45	0
45-59	6
>60	12
WHO performance status	
0-1	1
2	4
3-4	8
Extent of neurosurgery	
Complete resection	0
Partial resection	4
Biopsy	8
History of fits	
>3 months	0
<3 months	5
None	10

3.1.2 IDH mutations, MGMT promoter methylation status and 1p19q deletions as prognostic and predictive markers in GBM

Two of the most widely studied molecular markers in glioma are mutations in the isocitrate dehydrogenase genes (*IDH1* and *IDH2*) and *MGMT* promoter methylation status. Point mutations within either *IDH1* or *IDH2* occur in approximately 80% of adult diffuse astrocytoma (grade 2) but are generally absent from the paediatric equivalent (Ichimura et al., 2009, Parsons et al., 2008, Jones et al., 2011). In adult low grade glioma, the presence of an *IDH* mutation has been shown to be associated with good prognosis and also with response to chemotherapy (Houillier et al., 2010). Mutations within the *IDH* genes are regarded as characteristic of secondary glioblastoma and represent a marker of both good prognosis and a good response to treatment (Hartmann et al., 2013).

MGMT promoter methylation status has been widely studied in glioma. Heigi *et al* found that *MGMT* promoter methylation status was an independent predictor of a favourable prognosis in GBM. The same study also found that in patients with a methylated *MGMT* promoter, who were treated with radiotherapy plus temozolomide, prognosis was significantly better than in those who received radiotherapy alone. In patients without *MGMT* promoter methylation, no significant difference in survival was seen in patients treated with chemo-radiotherapy versus those treated with radiotherapy alone. The study concludes that the benefit seen in patients treated with temozolomide is only seen in patients with a methylated *MGMT* promoter (Hegi et al., 2005). *MGMT* promoter status may also have a role to play in selecting treatment for elderly patients (>65 years) with GBM. In current practice these patients are usually treated with radiotherapy alone. Wick *et al* compared radiotherapy alone to temozolomide alone in elderly patients with GBM and found that in the cohort as a whole temozolomide was non-inferior to radiotherapy. However, if *MGMT* methylation status was taken into account, patients with a methylated promoter had a longer event free survival if treated with temozolomide (Wick et al., 2012).

Total loss of 1p19q is a characteristic feature of oligodendroglioma. Over 80% of oligodendroglial tumours have been shown to have loss of alleles from 19q and nearly 70% have allelic deletions on 1p; typically these two deletions are found in combination (Reifenberger et al., 1994). The presence of the 1p19q deletion has been shown to be associated with improved survival. Median survival is 12-15 years for patients with oligodendroglioma and more than 7

years for patients with anaplastic oligodendoglioma with a 1p19q co-deletion, significantly higher than in patients without the deletion (Cairncross et al., 2006, Kaloshi et al., 2008, van den Bent et al., 2006). Patients with a low grade glioma and 1p19q deletion have a better prognosis and response to temozolomide than those without the deletion (Kaloshi et al., 2007). Similarly, in high grade glioma, the presence of a 1p19q deletion is associated with a more favourable prognosis and response to both PCV (procarbazine, lomustine and vincristine) chemotherapy (Cairncross et al., 1998) and radiotherapy (Bauman et al., 2000).

3.1.3 Prognostic scores in high grade glioma incorporating clinical and molecular variables

More recently, prognostic scores have been described that incorporate both clinical and molecular markers. Sub-group analysis of a randomised trial investigating the addition of temozolomide to radiotherapy identified factors associated with favourable prognosis in adult patients with GBM which include: more extensive tumour resection, Mini-Mental Status Examination (MMSE) score >27, younger age, better performance status and *MGMT* promoter methylation (Gorlia et al., 2008).

Bertolini *et al* at describe a prognostic score that incorporates patient age (<50 years versus ≥50 years), Eastern Cooperative Oncology Group (ECOG) performance status (0 versus ≥1), extent of surgery (gross tumour resection versus partial resection/biopsy) and *MGMT* promoter methylation status. Favourable prognostic markers include younger age, lower performance status, more extensive surgery and methylated *MGMT* promoter and all score one

point. Unfavourable prognostic markers score zero. Based on their score, patients were separated into 3 groups, with group 1 having the highest score (≥3) and the best prognosis (Bertolini et al., 2012).

3.1.4 Prognostic scores in high grade glioma based on molecular markers

Other studies have used global gene expression patterns to stratify high grade glioma patients into prognostic groups. Phillips *et al* described three molecular signatures in high grade glioma patients which are referred to as: Proneural, Proliferative and Mesenchymal. The Proneural subtype was found to be associated with grade 3 gliomas and carried a better prognosis than either the Proliferative or Mesenchymal subtypes. A transition from the Proneural to Mesenchymal subtype was commonly seen on disease recurrence (Phillips et al., 2006). Sturm *et al* studied patients with GBM and identified six distinct methylation clusters, including two distinct clusters characterised by specific mutations in the *H3F3A* gene (which encodes the histone protein H3.3) and one characterised by *IDH1* mutational status. In addition, the two *H3F3A* mutations appear to be associated with the differential expression of the transcription factors *OLIG1*, *OLIG2* and *FOXG1*. Moreover, expression pattern of these transcription factors is associated with tumour location, with low expression of *FOXG1* associated with midline tumours (Sturm et al., 2012).

3.1.5 DNA repair gene prognostic and predictive scores

Effective repair of DNA damage is essential for genomic stability and therefore deregulated DNA repair may promote genomic instability and tumour development (Torgovnick and Schumacher, 2015). The extent to which DNA repair pathways are deregulated in several tumour types has been described recently, including the use of this information to develop DNA repair prognostic scores. Kassambarra *et al* identified five DNA repair genes associated with a favourable prognosis and seventeen DNA repair genes associated with a poor prognosis in patients with previously untreated multiple myeloma. A DNA repair gene prognostic score was developed that incorporated these 22 DNA repair genes and that was able to predict both event free and overall survival (Kassambara et al., 2014). Kang *et al* developed a DNA repair gene score to predict response to platinum-based chemotherapy in patients with serous ovarian cystadenocarcinoma. Twenty-three genes that are known to be associated with the repair of platinum-induced DNA damage were included in the score. The score was found to be significantly associated with both survival and response to treatment (Kang et al., 2012). Abdel-Fatah *et al* studied a panel of DNA genes in oestrogen receptor negative and triple negative breast cancers and found XRCC1, Polβ, FEN1 and BRCA1 were independently associated with survival. A DNA repair gene prognostic index was developed incorporating the expression levels of these four genes and was able to separate patients into two distinct prognostic groups. The combination of prognostic index score and lymph node stage significantly improved prediction of breast cancer specific survival, as compared to either the prognostic score or lymph node status alone (Abdel-Fatah et al., 2015).

3.1.6 Array express and The Cancer Genome Atlas (TCGA) datasets

In this current study two publically available gene expression datasets have been analysed to identify DNA repair genes that are associated with survival in GBM. The first of these datasets was E-GEOD-13041 and was downloaded from www.ebi.ac.uk/arrayexpress (Kolesnikov et al., 2015). This was used as a Test dataset and comprised gene expression data collected from 191 patients using the Affymetrix U133A array. ArrayExpress is an on-line databank of functional genomics data held by the European Bioinformatics Institute (EBI). This site collects data from both microarray and high-throughput sequencing studies. Data is either obtained by direct submission or imported weekly in a systematic fashion from the National Centre for Biotechnology Information (NCBI) Gene Expression Omnibus database. High data quality is maintained by the application of the "Minimum Information About A Microarray Experiment" (MIAME) and "Minimum Information About a Sequencing Experiment" (MINSEQE) standards. Demographic data for the Test dataset is shown in Section 2.1.4.

The second, larger, Validation dataset was obtained from The Cancer Genome Atlas (TCGA) resource (http://cancergenome.nih.gov/) (Weinstein et al., 2013, Omberg et al., 2013). In this dataset, gene expression data was collected using the HT_HG_U133A Affymetrix array and after processing consisted of 508 glioblastoma patients. The Cancer Genome Atlas was established as a three year pilot project in 2006 with funding from the National Cancer Institute (NCI) and National Human Genome Research Institute (NHGRI). This pilot project proved that results could be pooled from multiple teams working on

different but related projects and that this was cost effective. By making the data freely available, the project enabled researchers worldwide to access the data, resulting in the validation of important discoveries. Following on from the success of the pilot study the National Institutes of Health pledged significant resources to enable the collection of data for more than 20 additional tumour types. The tumours selected for study by the TCGA were those with a poor prognosis or significant public health impact and those with availability of both human tumour and matched-normal tissue samples (Chin et al., 2011b, Chin et al., 2011a). Glioblastoma was one of the early priority areas and one of the first tumour sites to have data collected. Demographic data for the TCGA dataset is shown in Section 2.1.5.

In this current study gene expression data from these two datasets was used to identify DNA repair genes that are significantly and consistently associated with survival in GBM. A DNA repair gene prognostic index was then computed incorporating these DNA repair genes.

3.1.7 Artificial neural network analysis techniques

Once five key DNA repair genes were identified, ANN inference work was carried out in collaboration with Ms D. Aggarwal and Professor G. Ball (Department of Biosciences, Nottingham Trent University, Nottingham, UK). An ANN was constructed for each of the five key DNA repair genes included in the prognostic index. This was performed in the larger TCGA dataset.

Artificial neural networks are a form of machine learning derived from the discipline of artificial intelligence, with established pattern recognition

abilities. Artificial neural networks are able to deal with highly dimensional, complex, datasets, such as those derived from gene expression microarrays. Such techniques have been used to investigate concepts such as disease classification and the identification of biomarkers (Lancashire et al., 2009). In this study ANN analysis has been used to identify key genes that interact with each of the five DNA repair genes of interest. In the technique described here, probes were selected to represent each of the five key DNA repair genes; the probes selected were those which demonstrated the highest significance in the previous survival analysis. Artificial neural network analysis techniques were then used to determine the probes within the TCGA dataset which are best able to predict expression of each of the five DNA repair genes. The top 200 gene probes able to predict expression of the DNA repair gene were identified. An ANN based network inference algorithm was then applied to predict the magnitude and direction of the interactions between each of the 200 gene probes. The 100 strongest interactions were then identified and visualised in a network map using the Cytoscape programme (Smoot et al., 2011).

3.1.8 Study rationale

The ability of a cell to repair damage to its DNA has been shown to influence both the development of cancer and also how that tumour responds to various treatment modalities (Torgovnick and Schumacher, 2015). In this Chapter, two large publically available gene expression datasets have been screened for a panel of DNA repair genes to identify key DNA repair genes that influence survival in adult GBM. This data has been used to develop a DNA repair gene prognostic index. Further exploration of how these key DNA repair genes

interact within the wider genome has been performed using ANN techniques, enabling the identification of additional genes that may be involved in the development of high grade glioma.

3.2 Results

3.2.1 Exploring DNA repair gene expression in the Test and TCGA datasets

A list of the 188 DNA repair genes screened is shown in Appendix 1. Probes representing these genes were identified using NetAffx™ Analysis Center (https://www.affymetrix.com/analysis/index.affx). Two-hundred and forty eight probes, representing 157 genes, were present in both datasets. mRNA expression levels of each probe were dichotomised into low and high expression using Xtile.

3.2.1.1 DNA repair gene expression is associated with poor prognosis in adult glioblastoma

Kaplan Meier survival analysis was performed for the 248 DNA repair gene probes in both the Test dataset and the TCGA dataset (Appendix 2). After correction for multiple comparisons, using the Benjamini Hochberg False Discovery Rate calculation (BH FDR) (Benjamini, 1995), fourteen DNA repair gene probes (for 12 DNA repair genes) were found to be significantly associated with survival in both the Test and TCGA datasets (Table 3.2).

Table 3.2. DNA repair genes significantly associated with survival in both datasets after BH FDR correction

Gene	**Probe**	**Level associated with worse survival**	***p* value (Test dataset)**	***p* value (TCGA dataset)**
APE1	210027_s_at	Low	0.003	0.000018
PARP2	204752_x_at	Low	0.014	0.014
ERCC6	207347_at	Low	0.006	0.010
RAD21	200607_s_at	Low	0.001	0.004
PTEN	204054_at	Low	0.004	0.001
NBN	202905_x_at	Low	0.000001	0.001
	202907_x_at	Low	0.001	0.000125
	217299_s_at	Low	0.000006	0.006
MGMT	204880_at	High	0.003	0.001
BRCA1	214727_at	High	0.009	0.002
PMS2	209805_at	High	0.007	0.011
PARP3	209940_at	High	0.002	0.004
DDB2	203409_at	High	0.000097	0.00005
RAD23B	201222_s_at	High	0.000254	0.006

BH FDR=Benjamini Hochberg False Discovery Rate calculation.

The twelve significant DNA repair genes on univariate analysis represent most of the major DNA repair pathways (Figure 3.1). Half of the significant DNA repair genes are involved in either the base excision or nucleotide excision repair pathways.

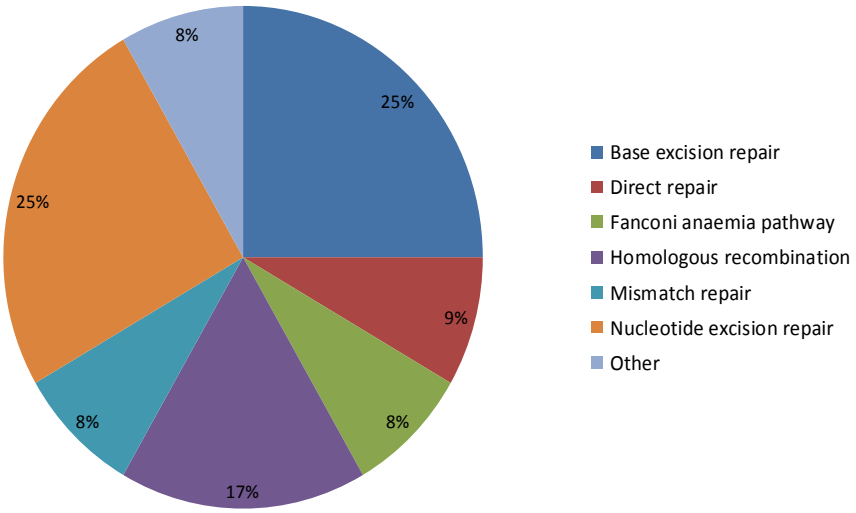


Figure 3.1. DNA repair pathways represented by the 12 genes significant in both the Test and TCGA datasets.

This pie chart shows that the 12 DNA repair genes that are significantly associated with survival in GBM, in both the Test and TCGA datasets, represent the majority of DNA repair pathways. The base excision repair and nucleotide excision repair pathways are the most widely represented.

3.2.1.2 Five DNA repair genes are independent predictors of survival in adult glioblastoma

Cox multivariate regression analysis was performed in both the Test and TCGA datasets to identify the DNA repair genes most significantly associated with survival independent of other DNA repair genes. All fourteen significant probes from Kaplan Meier analysis were included in the model. On multivariate cox regression analysis in the Test dataset six DNA repair gene probes were independently associated with survival; *APE1* (p=0.000810), *Rad23B* (p=0.000167), *PMS2* (p=0.000190), *NBN* (202905_x_at p=0.000846), *MGMT* (p=0.001326) and *PTEN* (p=0.001108). In the TCGA dataset seven DNA repair gene probes were independently associated with survival including *APE1* (p=0.000128), *PMS2* (p=0.012998), *NBN* [202905_x_at p=0.000025 and

202907_s_at p=0.003634), *BRCA2* (p=0.000188), *MGMT* (p=0.002090) and *PTEN* (p=0.001221). Five DNA repair gene probes were significant on multivariate analysis in both the Test and TCGA datasets. Multivariate analysis was performed including only these five probes in both datasets. The final multivariate models for both the Test and TCGA datasets are shown in Table 3.3. The Kaplan Meier survival curves for these five DNA repair genes in both the Test and TCGA datasets are shown in Figure 3.2 and Figure 3.3, respectively.

Table 3.3. Multivariate analysis results for the Test and TCGA datasets.

Gene	**Test dataset**		**TCGA dataset**	
(Probe)	HR (95% CI)	*p* value	HR (95% CI)	*p* value
APE1 (210027_s_at)	0.62 (0.45-0.85)	0.002616	0.57 (0.42-0.77)	0.000311
PMS2 (209805_at)	1.84 (1.32-2.55)	0.000299	1.47 (1.17-1.86)	0.000970
NBN (202905_x_at)	0.48 (0.34-0.69)	0.000053	0.60 (0.48-0.77)	0.000045
MGMT (204880_at)	1.55 (1.09-2.19)	0.013562	1.41 (1.11-1.79)	0.004329
PTEN (204054_at)	0.62 (0.45-0.84)	0.001994	0.67 (0.52-0.86)	0.001575

HR=hazard ratio; CI=Confidence interval.

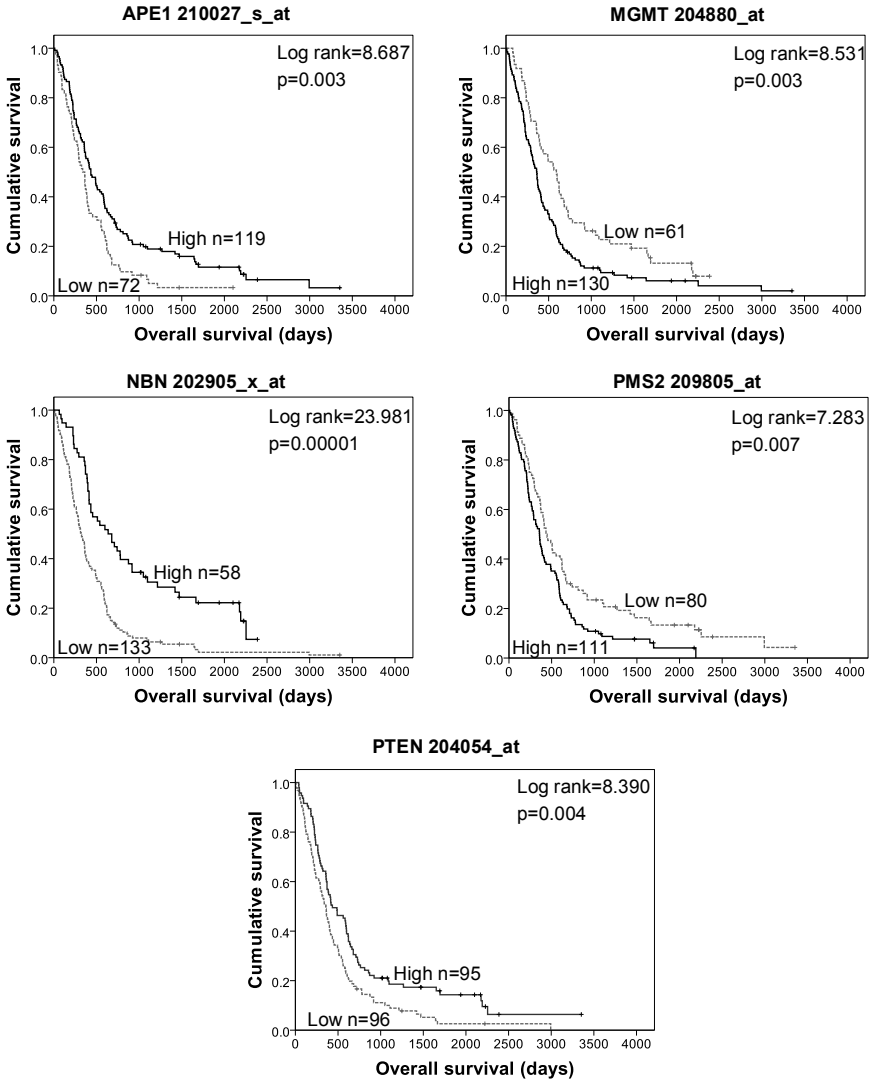


Figure 3.2. Kaplan Meier survival curves for the five DNA repair genes included in the prognostic index in the Test dataset.

Low APE1 (p=-0.003), low NBN (p=0.00001), low PTEN (p=0.004), high MGMT (p=0.003) and high PMS2 (p=0.007) expression are significantly associated with poor overall survival in GBM in the Test dataset. Xtile software was used to determine biologically relevant cut-off levels for classification of expression into high and low groups.

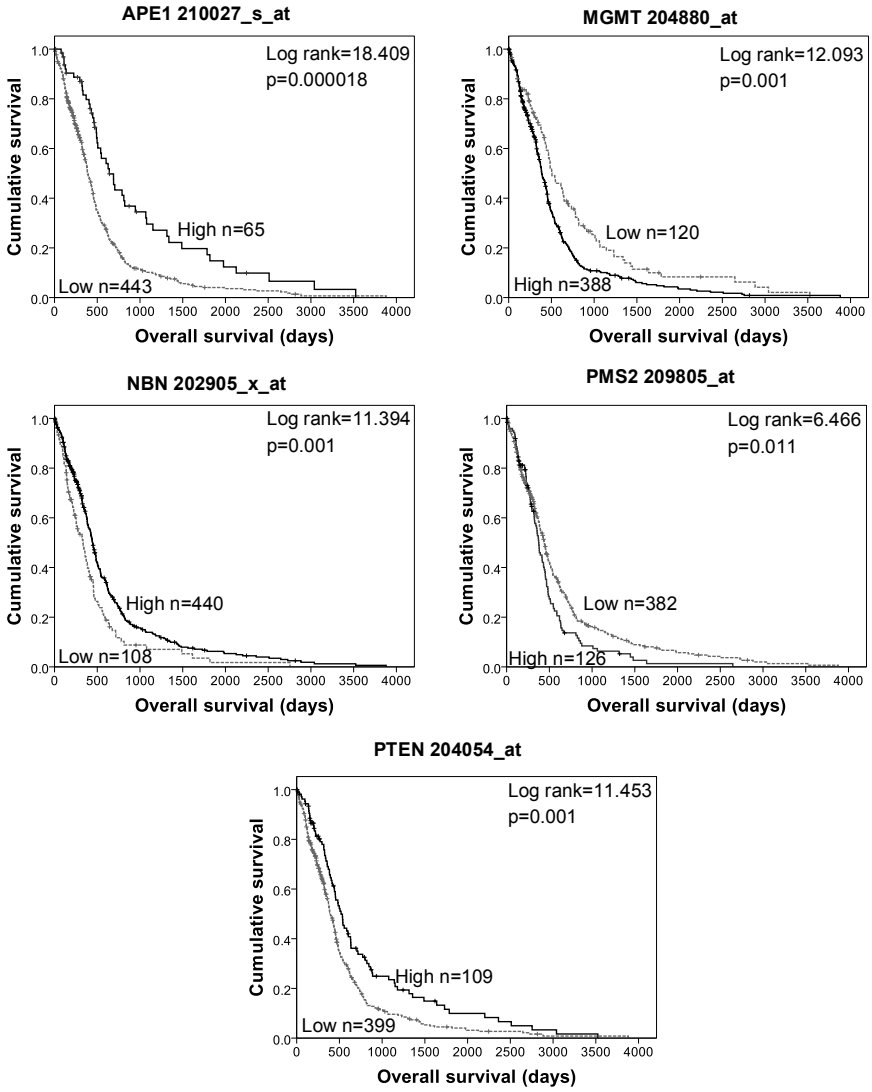


Figure 3.3. Kaplan Meier survival curves for the five DNA repair genes included in the prognostic index in the TCGA dataset.

Low APE1 (p=-0.0000018), low NBN (p=0.001), low PTEN (p=0.001), high MGMT (p=0.001) and high PMS2 (p=0.011) expression are significantly associated with poor overall survival in GBM in the Test dataset. Xtile software was used to determine biologically relevant cut-off levels for classification of expression into high and low groups.

3.2.1.3 Development of a DNA repair gene prognostic index in adult glioblastoma

After multivariate analysis, *APE1*, *MGMT*, *NBN*, *PMS2* and *PTEN* were found to be the DNA repair genes most significantly associated with survival in adult GBM. As discussed in Section 3.1.5 previous studies have developed DNA repair gene prognostic scores for several tumour types (Kassambara et al., 2014, Kang et al., 2012, Abdel-Fatah et al., 2015), but not GBM. A DNA repair gene prognostic index was therefore developed incorporating *APE1*, *MGMT*, *NBN*, *PMS2* and *PTEN* mRNA expression levels; the methodology for this is described in more detail in Section 2.1.7. Initially a prognostic index was developed in the Test dataset (PI_1) and the TCGA dataset (PI_2) separately. A combined prognostic index (PI_3) was generated using the mean β value for each gene from the two datasets. The formulae for the three prognostic indices are shown in Table 3.4. Expression of each DNA repair gene is classified as either 0 (low) or 1 (high) and multiplied by the β value obtained from the cox multivariate analysis (or the mean β value for PI_3). This is performed for each of the five DNA repair genes and the values summed to give a final prognostic score.

Table 3.4. Formulae for the three Prognostic Indices developed in the Test and TCGA datasets.

Prognostic index	**Formula**
PI_1	(*APE1**-0.478) + (*PMS2**0.607) - (*NBN**0.725) + (*MGMT**0.437) - (*PTEN**0.484)
PI_2	(*APE1**-0.569) + (*PMS2**0.389) - (*NBN**0.514) + (*MGMT**0.345) - (*PTEN**0.393)
PI_3	(*APE1**-0.524) + (*PMS2**0.498) - (*NBN**0.620) + (*MGMT**0.391) - (*PTEN**0.439)

PI_1=Prognostic Index developed in the Test dataset. PI_2=Prognostic Index developed in the TCGA dataset. PI_3=Combined Prognostic Index which is an average of PI_1 and PI_2.

All three prognostic indices were tested in both the Test and TCGA datasets. For each patient in both datasets the prognostic index formula was applied and a prognostic score calculated. For example, using PI_1 a prognostic score was calculated for all patients in the Test dataset. Based on these prognostic scores, patients in the Test dataset were divided into 3 or 4 groups. Kaplan Meier survival analysis was performed and a log rank score calculated. This was performed using different prognostic score cut-offs. The process was then repeated using PI_2 and PI_3 in the Test dataset and all 3 prognostic indices in the TCGA dataset. Log rank scores were then compared and the prognostic index, and associated cut-offs, best able to separate patients into prognostic groups in both the Test and TCGA datasets was selected.

3.2.1.4 A DNA repair gene prognostic index can separate adult patients with GBM into three distinct prognostic groups

In both the Test and TCGA datasets (Figure 3.4) PI_3 can separate GBM patients into three distinct prognostic groups using Kaplan Meier methodology. In both datasets, patients in Prognostic Group 1 have a significantly better survival than patients in Prognostic Group 2. Similarly, patients in Prognostic Group 2 have a significantly better survival than patients in Prognostic Group 3.

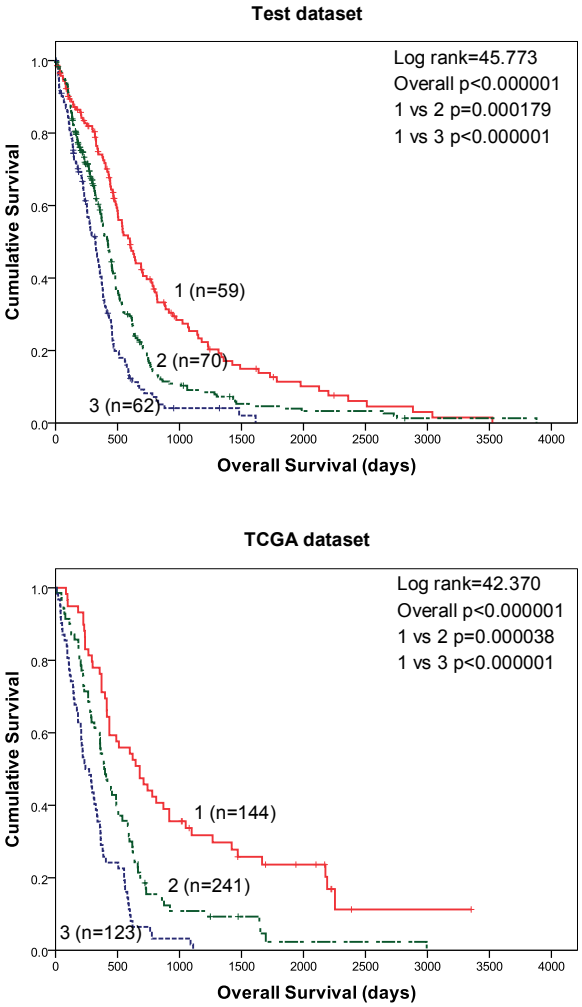


Figure 3.4. Kaplan Meier survival curves for the Combined Prognostic Index (PI_3) in both the Test and TCGA datasets.

Patients in Group 1 (red line) represent the best prognostic group. Group 2 patients (green line) are an intermediate prognostic group and Group 3 patients (blue line) are the group with the worst prognosis. In both the Test and TCGA datasets poorer overall survival is seen in patients in Prognostic Groups 2 and 3 as compared to Prognostic Group 1. Cut-offs for prognostic scores for each group were as follows: Group [1] -1.583 to -0.524, Group [2] -0.439 to -0.048 and Group [3] 0.000 to 0.889.

3.2.1.5 A DNA repair gene prognostic index can predict 1, 2 and 3 year survival in adult GBM

As discussed in the Methods section, using PI_3, survival curves were constructed for the both the Test and TCGA cohorts to predict survival at 1, 2 and 3 years. The median prognostic score for each of the three prognostic groups was plotted against the percentage survival for each of the three time points. Survival at 1, 2 and 3 years for each prognostic index group is shown in Figure 3.5. For all three time points, patients with a lower prognostic score had a greater chance of being alive than those with a higher score.

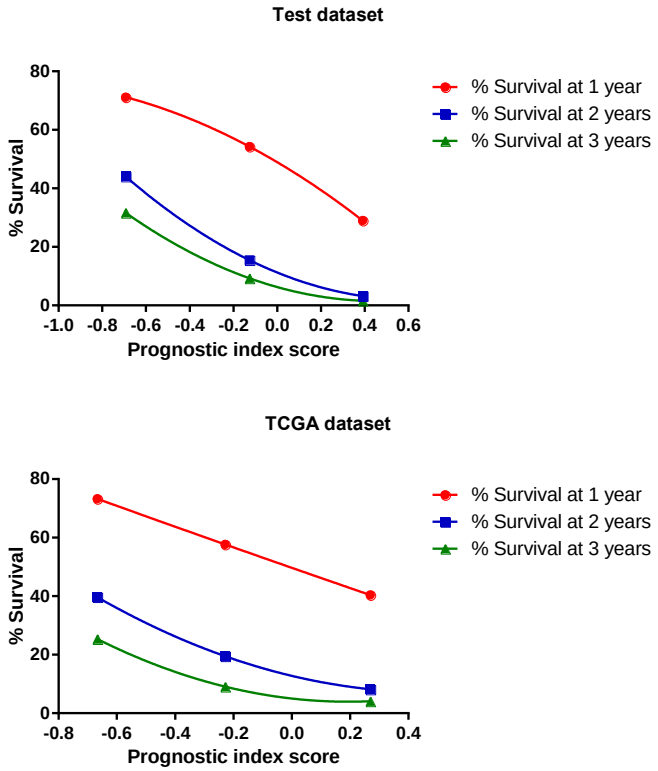


Figure 3.5. Graphical representation of 1, 2 and 3 year survival as predicted by the Prognostic Index score in the Test and TCGA datasets.

Mean Prognostic Index score for each of the 3 prognostic groups is plotted against the average survival of patients in that group at 1, 2 and 3 years. In both the Test and TCGA datasets lower survival is demonstrated in association with higher prognostic index scores for all three time points.

3.2.1.6 Predictive significance of APE1, MGMT, NBN, PMS2, and PTEN mRNA expression in GBM

APE1, *MGMT*, *NBN*, *PMS2*, and *PTEN* have been shown to have prognostic significance in adult GBM. In order to investigate the predictive significance of these genes in GBM, Kaplan Meier survival analysis was performed for each gene in patients who did, and in those that did not, have radiotherapy in the TCGA cohort. This analysis could not be performed in the Test dataset as treatment data was missing.

In total, 373 (73.4%) patients in the TCGA cohort received radiotherapy during the course of their illness. Median age of patients who were treated with radiotherapy was 57 years (range 10-86 years) and 141 (37.8%) were female. 294 (78.8%) had died by the end of the follow-up period and median survival was 448 days. In patients that received radiotherapy, low *APE1* (p=0.000086), low *NBN* (p=0.002), low *PTEN* (p=0.017), high *MGMT* (p=0.000192) and high *PMS2* (p=0.002) were associated with poor overall survival (Figure 3.6).

One-hundred and thirty-five patients were recorded as not having received radiotherapy. Median age of these patients was 66 years (range 14-89 years) and 59 (43.7%) were female. Median survival for patients not treated with radiotherapy was 142 days and 122 (90.4%) had died by the end of the follow-up period. In patients that did not receive radiotherapy only low *NBN* expression (p=0.048) was significantly associated with survival (Figure 3.7).

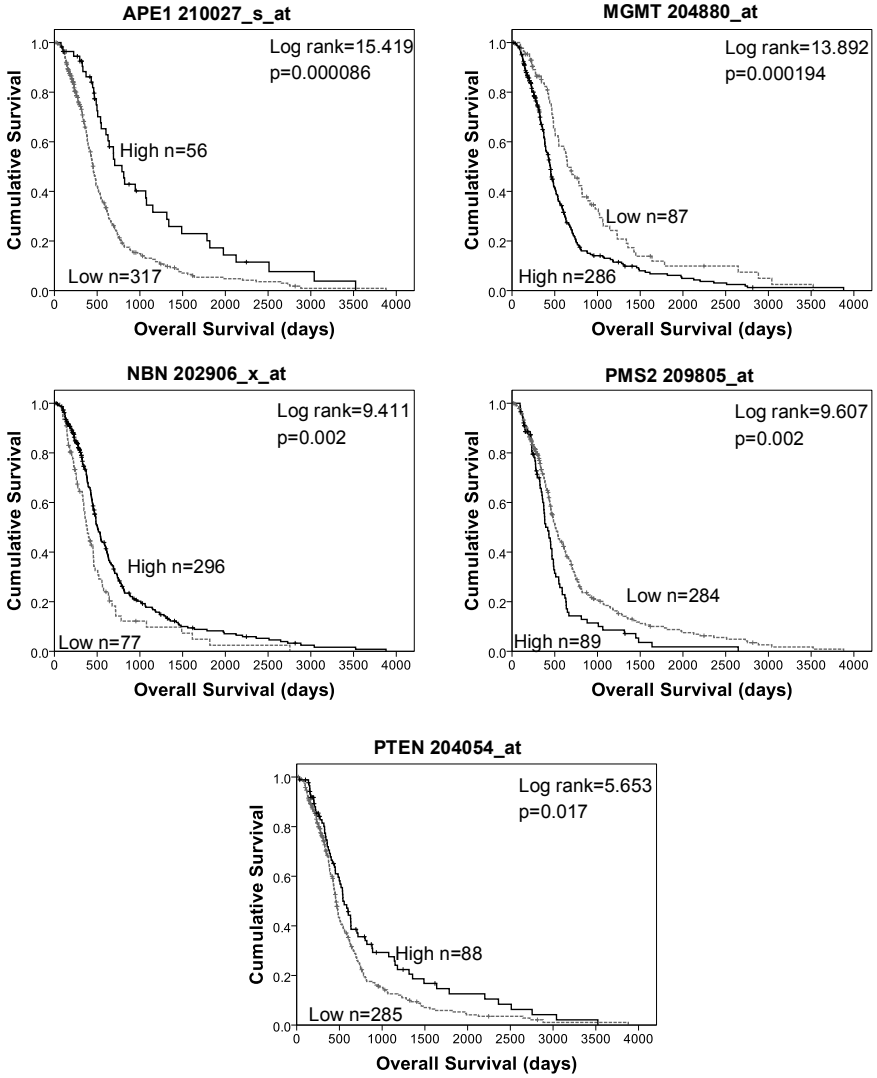


Figure 3.6. Kaplan Meier survival curves for the five DNA repair genes included in the prognostic index in radiotherapy treated patients in the TCGA dataset.

Low APE1 (p=-0.000086), low NBN (p=0.002), low PTEN (p=0.017), high MGMT (p=0.000194) and high PMS2 (p=0.002) expression are significantly associated with poor overall survival in GBM patients treated with radiotherapy in the TCGA dataset.

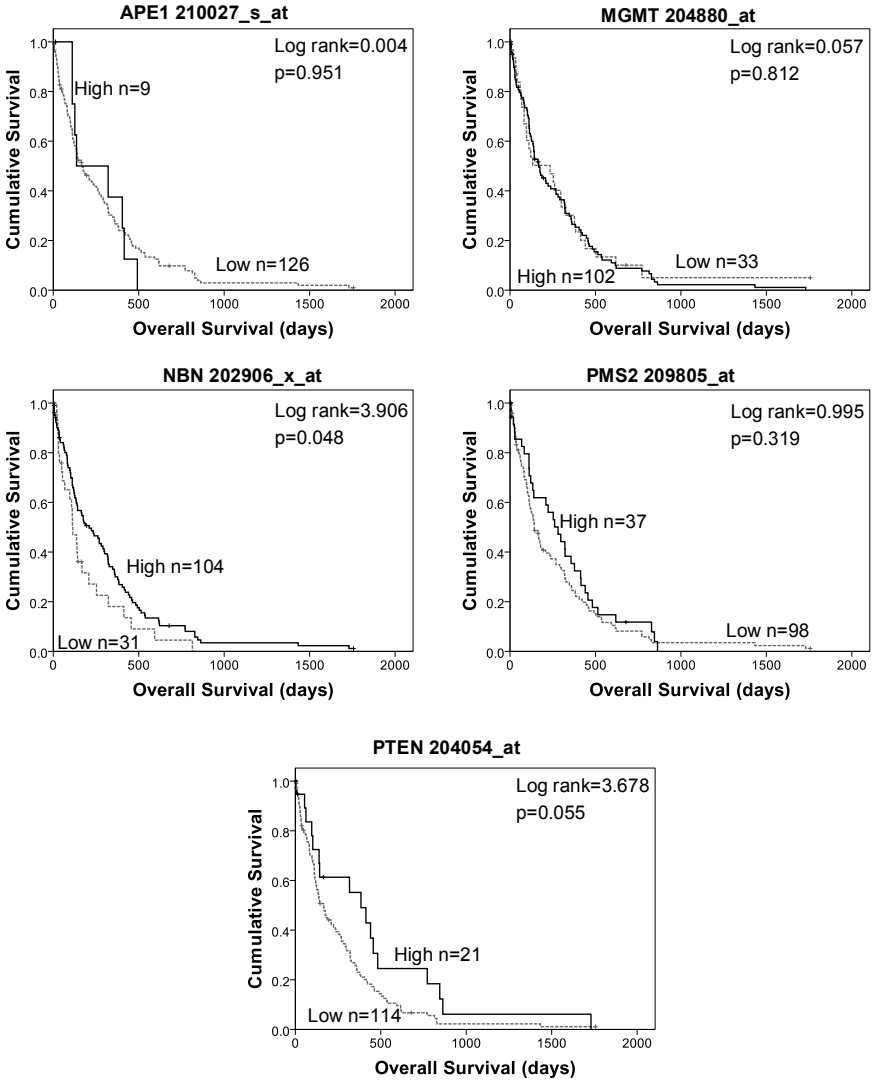


Figure 3.7. Kaplan Meier survival curves for the five DNA repair genes included in the prognostic index in patients not treated with radiotherapy in the TCGA dataset.

Only low NBN (p=0.048) expression is significantly associated with poor overall survival in GBM patients not treated with radiotherapy in the TCGA dataset.

Three-hundred and fifty-one patients in the TCGA dataset received chemotherapy during the course of their illness. Median age of chemotherapy treated patients was 58 years (range 10-89 years) and 129 (36.8%) were female. By the end of the follow-up period 277 (78.9%) patients had died and median survival was 453 days. In these chemotherapy treated patients, low *APE1* (p=0.000124), low *NBN* (p=0.001), low *PTEN* (p=0.017), high *MGMT* (p=0.000357) and high *PMS2* (p=0.001) were significantly associated with poor overall survival (Figure 3.8).

One-hundred and fifty seven patients were not recorded as receiving chemotherapy. Median age of non-chemotherapy treated patients was 65 years (range 14-88 years) and 86 (54.8%) were female. Median survival of patients not treated with chemotherapy was 233 days and 137 (87.3%) of these patients had died by the end of the follow-up period. In the patients not treated with chemotherapy only low *NBN* (p=0.005) and low *PTEN* (p=0.025) were associated with survival (Figure 3.9).

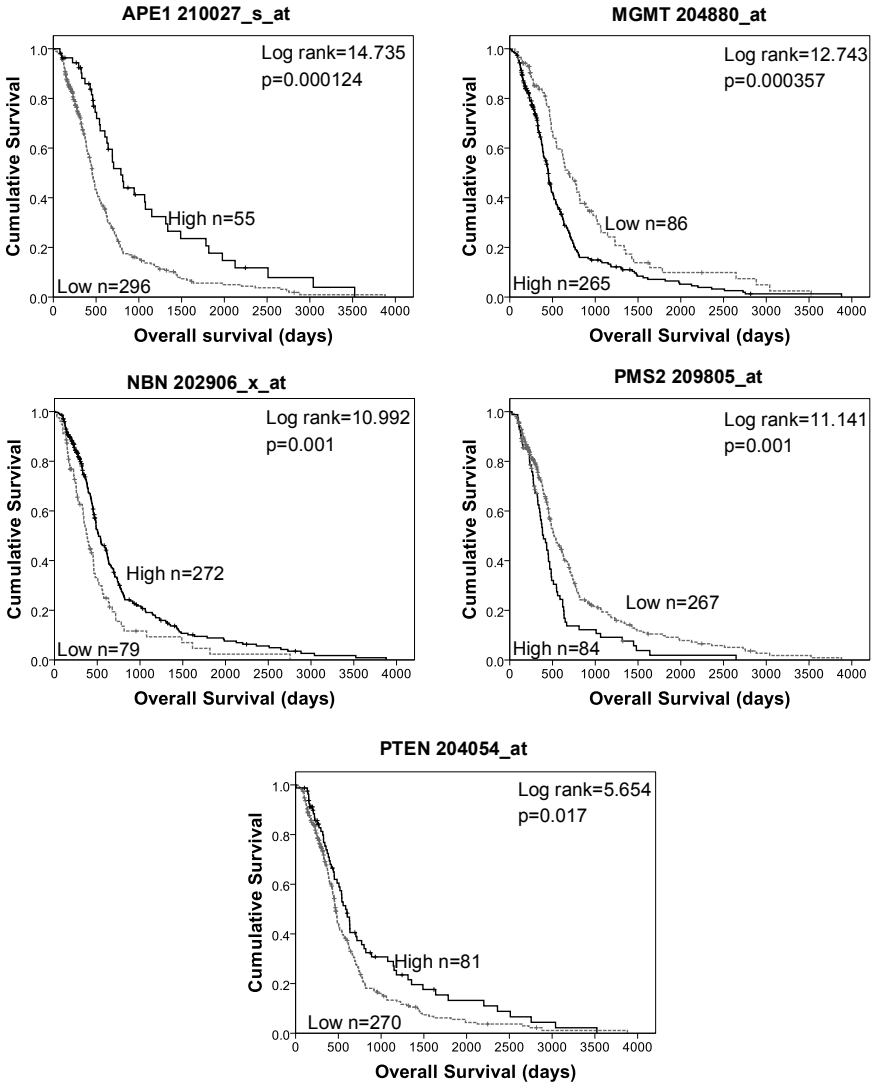


Figure 3.8. Kaplan Meier survival curves for the five DNA repair genes included in the prognostic index in chemotherapy treated patients in the TCGA dataset.

Low APE1 (p=-0.000124), low NBN (p=0.001), low PTEN (p=0.017), high MGMT (p=0.000357) and high PMS2 (p=0.001) expression are significantly associated with poor overall survival in GBM patients treated with chemotherapy in the TCGA dataset.

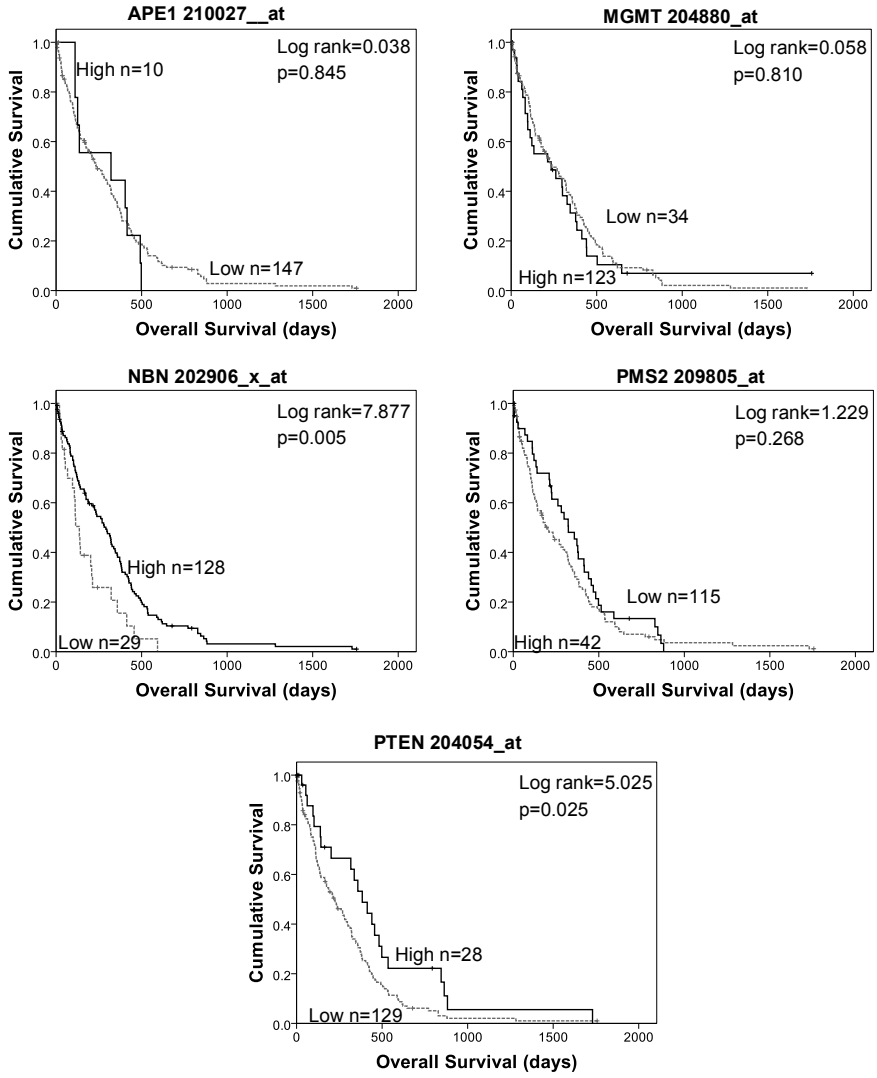


Figure 3.9. Kaplan Meier survival curves for the five DNA repair genes included in the prognostic index in patients not treated with chemotherapy in the TCGA dataset.

Only low NBN (p=0.005) and low PTEN (p=0.025) expression are significantly associated with poor overall survival in GBM patients not treated with chemotherapy in the TCGA dataset.

3.2.1.7 APE1, PMS2 and PTEN mRNA expression levels are associated with age in adult GBM

As discussed previously, age is known to be associated with survival in adult GBM, with older patients having a worse outcome. In both the Test (p=0.001) and TCGA (p<0.00001) datasets patients aged 60 years or older had a significantly worse outcome than patients less than 60 years (Figure 3.10).

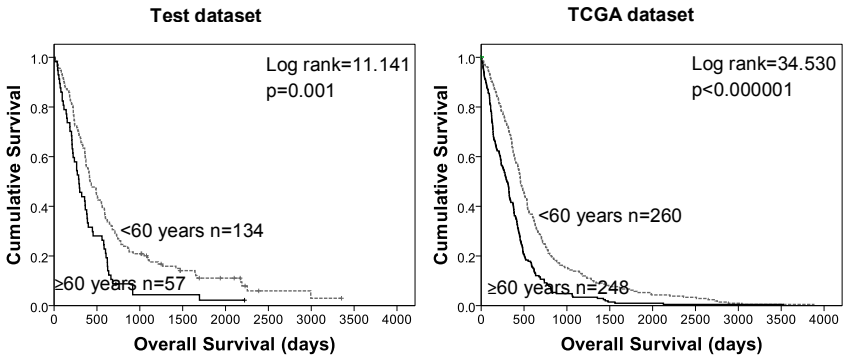


Figure 3.10. Kaplan Meier survival curves for age at diagnosis in the Test and TCGA datasets.

GBM patients aged <60 years at diagnosis have significantly better overall survival than patients aged ≥60 years at diagnosis in the both the Test and TCGA datasets.

An association between *APE1*, *MGMT*, *NBN*, *PMS2* and *PTEN* expression and age at diagnosis was assessed using the Mann-Whitney U test (Table 3.5). In the Test dataset low *APE1* (p=0.031), low *NBN* (p=0.007) and high *PMS2* (p=0.014) were associated with higher age at diagnosis. *MGMT* (p=0.540) and *PTEN* (p=0.426) expression were not associated with age at diagnosis. In the TCGA cohort low *APE1* (p<0.001), high *PMS2* (p=0.016) and low *PTEN* (p<0.001) were associated with higher age at diagnosis. *NBN* (p=0.218) and *MGMT* (p=0.372) expression were not associated with age at diagnosis.

Table 3.5. The association between APE1, PMS2, NBN, MGMT and PTEN expression and age in the Test and TCGA datasets.

Gene		**Test dataset**		**TCGA dataset**	
		Median age (years)	*p* value	Median age (years)	*p* value
APE1	Low	56.5	**0.031**	60	**<0.001**
	High	53		53	
PMS2	Low	52.5	**0.014**	59	**0.016**
	High	56		61	
NBN	Low	56	**0.007**	60	0.218
	High	49		59	
MGMT	Low	54	0.540	60.5	0.372
	High	54		59	
PTEN	Low	54	0.426	61	**<0.001**
	High	55		53	

Significant p values (≤0.05) are shown in bold.

3.2.1.8 Female gender is associated with PTEN mRNA expression

In the TCGA cohort, Chi squared analysis revealed that female gender is associated with low *PTEN* mRNA expression (p=0.026). No association was seen between *APE1* (p=0.665), *MGMT* (p=0.149), *NBN* (p=0.150) and *PMS2* (p=0.257) mRNA expression and gender. Gender data was not available for the Test dataset.

3.2.1.9 Low APE1 mRNA expression is associated with poor performance status

Chi squared analysis in the TCGA cohort demonstrated that low *APE1* mRNA expression (p=0.018) is associated with poor performance status (WHO performance status 2-4). No association was seen between *MGMT* (p=0.411), *NBN* (p=0.522), *PMS2* (p=0.880) and *PTEN* (p=0.220) mRNA levels and

WHO performance status. Performance status data was not available for the Test dataset.

3.2.1.10 Patients who received radiotherapy or chemotherapy had higher APE1 mRNA expression

On Chi Squared analysis in the TCGA dataset, patients with high *APE1* mRNA expression were more likely to have received radiotherapy (p=0.013). No association was seen between *MGMT* (p=0.793), *NBN* (p=0.572), *PMS2* (p=0.414) and *PTEN* (p=0.051) mRNA levels and the administration of radiotherapy treatment. Similarly, high *APE1* mRNA expression (p=0.004) was also associated with chemotherapy treatment. No association was seen between *MGMT* (p=0.485), *NBN* (p=0.304), *PMS2* (p=0.496) and *PTEN* (p=0.183) mRNA levels and the administration of chemotherapy. Patient treatment information was not available for the Test dataset.

3.2.1.11 APE1, MGMT, NBN and PMS2 mRNA expression levels are independent predictors of survival in GBM

Cox multivariate regression analysis was performed including *APE1*, *MGMT*, *NBN*, *PMS2* and *PTEN* mRNA expression levels and age, gender, WHO performance status and treatment delivered for the TCGA cohort. Treatment was recorded as having received radiotherapy or having received chemotherapy. Non-significant variables, including gender, chemotherapy treatment and PTEN mRNA expression, were removed after the first round of multivariate Cox regression analysis and the analysis re-run. After the second round of analysis all variables retained significance ($p \leq 0.05$) and the results are shown in Table 3.6. Older age (p=0.002), poor performance status

(p=0.00007), not receiving radiotherapy treatment (p<0.000001) and low *APE1* (p=0.007), low *NBN* (p=0.002), high *MGMT* (p=0.005) and high *PMS2* (p=0.005) mRNA expression were independently associated with poor overall survival.

Table 3.6. Multivariate analysis for DNA repair gene expression and clinical variables combined.

Variable	**TCGA dataset**	
	HR (95% CI)	*p* value
Age	0.678 (0.531-0.865)	0.002
Performance status	0.568 (0.430-0.750)	0.00007
Radiotherapy treatment	2.087 (1.569-2.777)	<0.000001
APE1 expression	1.611 (1.140-2.275)	0.007
MGMT expression	0.679 (0.518-0.890)	0.005
NBN expression	1.633 (1.206-2.211)	0.002
PMS2 expression	0.674 (0.513-0.886)	0.005

HR=hazard ratio; CI=confidence interval.

3.2.1.12 Associations are seen between expression levels of the five DNA repair genes

In the Test dataset, low *MGMT* expression was associated with high expression of *NBN* (p=0.012) and low *PMS2* expression was associated with low *PTEN* expression (p=0.046). The associations between the five DNA repair genes of interest in the Test dataset are shown in Appendix 3.

In the larger TCGA cohort, low *MGMT* expression was associated with high expression of *NBN* (p=0.003). Significant associations were also seen between *NBN* and both *PMS2* and *PTEN*. Low *NBN* was associated with low *PMS2* expression (p=0.001) and low *NBN* was also associated with high *PTEN* expression (p=0.004). In addition, low *PMS2* expression was associated with high *MGMT* expression (p=0.003). The associations between the five DNA repair genes of interest in the TCGA dataset are shown in Appendix 4.

3.2.2 Artificial neural network analysis for *APE1*, *MGMT*, *NBN*, *PMS2* and *PTEN*

In this study, ANN methodology has been used to identify key genes able to predict expression of *APE1*, *MGMT*, *NBN*, *PMS2* and *PTEN* in the TCGA dataset. A list of the top 200 probes able to predict expression of, for example, *APE1* was obtained. The gene represented by each of these probes was identified and the functional group to which they belong was ascertained using the Panther Classification System (Mi et al., 2013). Any genes that have a role in DNA repair were identified at this stage.

An interaction analysis was then run to establish how these top 200 probes interact with each other and to identify the top 100 interactions. These 100 interactions were visualised in a network map (or interactome) with gene probes replaced by gene names. Any non-specific probes were not included in the final interaction map. Self-loops (when a gene forms an interaction with itself) were also removed. In the interaction map, squares represent genes and are referred to as nodes. Nodes with ≥5 interactions are referred to as hubs and are circular. The three most influential hubs, as defined by the greatest number of interactions, are shaded yellow. Lines represent interactions and the arrow the direction of the interaction. Line width represents the strength of the interaction i.e. the wider the line the stronger the interaction. A blue line signifies an inhibitory interaction and a red line a stimulatory interaction. Genes that influence expression of other genes are referred to as input nodes, whereas genes that are influenced by the expression of other genes are called output nodes. The functional groups to which the genes in the interaction analysis fall are described; again the Panther Classification system was used.

Again, any genes with a potential role in DNA repair have been identified. In addition, for the three most influential hubs within each network, Kaplan Meier survival analysis was performed to assess whether expression of these genes is associated with survival.

3.2.2.1 Artificial neural network analysis for APE1

3.2.2.1.1 Top 200 probes that predict APE1 expression

The top 200 probes that predict *APE1* expression were identified; the top ten probes are shown in Table 3.7 alongside their biological function. Five genes were identified with a role in DNA repair and are summarized in Table 3.8, including their rank in the list of genes that predict *APE1* expression. *PARP2* is represented by three probes: 204752_x_at (ranked 3), 215773_x_at (ranked 4) and 214086_s_at (ranked 8).

Table 3.7. Top 10 probes that predict APE1 expression.

Probe	Gene symbol	Protein name	Function
217786_at	*PRMT5*	Protein arginine N-methyltransferase 5	Arginine methyltransferase activity. Negatively regulates cyclin E1 promoter activity and cellular proliferation.
204752_x_at 215773_x_at 214086_s_at	*PARP2*	Poly [ADP-ribose] polymerase 2	Involved in the base excision repair pathway – catalyses poly(ADP-ribosyl)ation of acceptor proteins involved in chromatin architecture and DNA metabolism.
213794_s_at	*NGDN*	Neuroguidin	Involved in the translational repression of cytoplasmic polyadenylation element (CPE)-containing mRNAs.
219816_s_at	*RBM23*	Probable RNA-binding protein 23	Probable RNA-binding protein. Potential role in pre-mRNA splicing process.
201684_s_at	*TOX4*	TOX high mobility group box family member 4	Component of the PTW/PP1 phosphatase complex. Role in control of chromatin structure and cell cycle progression.
200043_at	*ERH*	Enhancer of rudimentary homolog	Potential role in cell cycle.
208799_at	*PSMB5*	Proteasome subunit beta type-5	The proteasome is a multi-catalytic proteinase complex. May protect against oxidative damage.
211115_x_at	*GEMIN2*	Gem-associated protein 2	Role in RNA processing and splicing.

Table 3.8. DNA repair genes present in the top 200 APE1 predicting genes.

Rank	Gene symbol	Gene name	Function
3, 4, 8	*PARP2*	Poly (ADP-ribose) polymerase 2	Component of BER pathway and involved in poly(ADP-ribosyl)ation
22	*POLE2*	Polymerase (DNA directed), epsilon 2, accessory subunit	Involved in BER pathway and chromosomal DNA replication
28	*VRK*	Vaccinia related kinase 1	Serine/threonine kinase involved in Golgi disassembly during the cell cycle
35	*ALKBH1*	AlkB, Alkylation Repair Homolog 1 (E. Coli)	Dioxygenase that repairs alkylated single-stranded DNA and RNA containing 3-methylcytosine by oxidative demethylation
49	*MNAT1*	MNAT CDK-Activating Kinase Assembly Factor 1	Stabilises the CDK-activating kinase (CAK) and is involved in cell cycle regulation and RNA transcription by RNA polymerase II.

3.2.2.1.2 Interaction analysis for APE1

An interaction analysis was performed for the top 200 probes shown to predict *APE1* expression. Only one gene was identified that has a role in DNA repair which is *MNAT1*.

The interactome for *APE1* is shown in Figure 3.11. Eight nodes were identified within the interactome with ≥5 edges. The genes representing these nodes are shown in Table 3.9 with their biological function. The top three genes within the APE1 interactome are *FOXG1*, *TGOLN2* and *ACACB*.

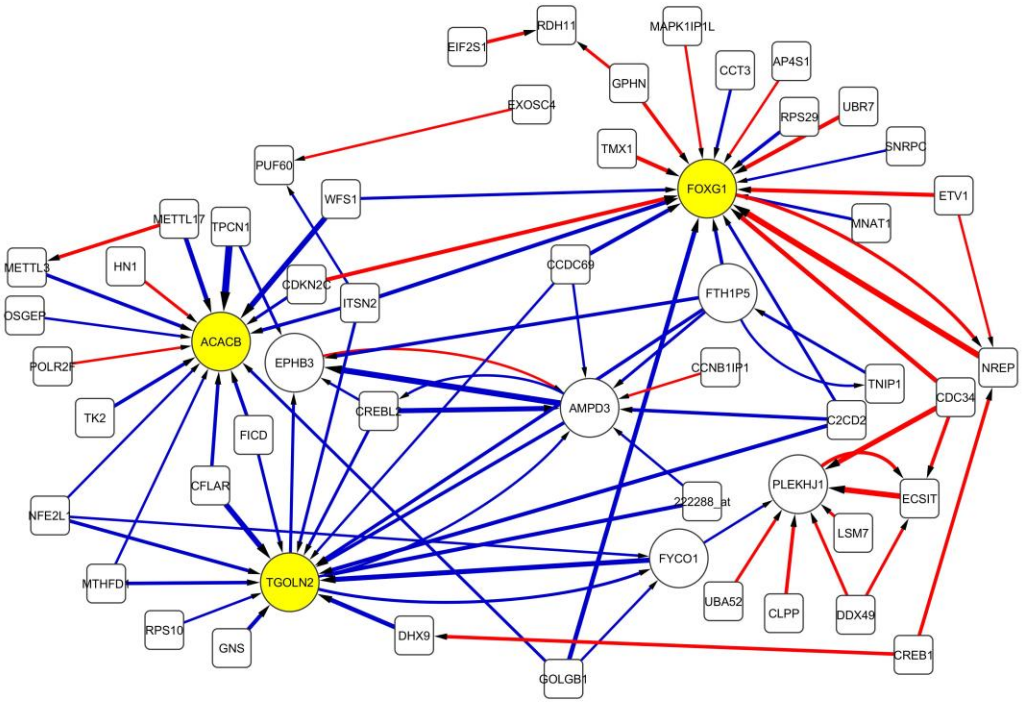


Figure 3.11. *APE1* interactome.

The interactome includes the top 100 *APE1* interactions. Red lines represent positive interactions and blue lines negative interactions. Wider lines signify a stronger interaction. Genes that interact with greater than 5 genes are referred to as hubs and are represented by a circle. The top 3 hubs (i.e. genes with the greatest number of interacting genes) are shown as a filled yellow circle; in this example *ACACB*, *FOXG1* and *TGOLN2*.

Table 3.9. Function of gene hubs within the APE1 interactome.

Gene symbol	Protein name	No. of edges	Function
FOXG1	Forkhead box G1	20	Transcription repression factor involved in brain development.
TGOLN2	Trans-Golgi network integral membrane protein 2	18	Potential role as regulator of membrane traffic to and from trans-Golgi network.
ACACB	Acetyl-CoA carboxylase beta	15	Involved in fatty acid biosynthesis.
AMPD3	Adenosine monophosphate deaminase 3	11	Plays critical role in energy metabolism.
PLEKHJ1	Pleckstrin homology domain containing, family J member 1	8	Potential role in phospholipid binding.
EPHB3	Ephrin type-B receptor 3	6	Receptor tyrosine kinase involved in contact-dependent bidirectional signalling into neighbouring cells. Involved in axon guidance during development.
FTH1P5	Ferritin, heavy polypeptide 1 pseudogene 5	6	Unknown function.
FYCO1	FYVE and coiled-coil domain containing 1	5	Potential role in microtubule plus end-directed vesicle transport.

3.2.2.1.3 FOXG1 and TGOLN2 expression are associated with survival in GBM

On Kaplan Meier survival analysis of the top three hubs within the *APE1* interactome, a significant association was seen between high *TGOLN2* (probe 20388_a_at p=0.016, probe 212043_at p=0.017) expression and poor survival

and high *FOXG1* expression and poor survival (p=0.002). *ACACB* expression was not associated with survival in this cohort (p=0.084) (Figure 3.12).

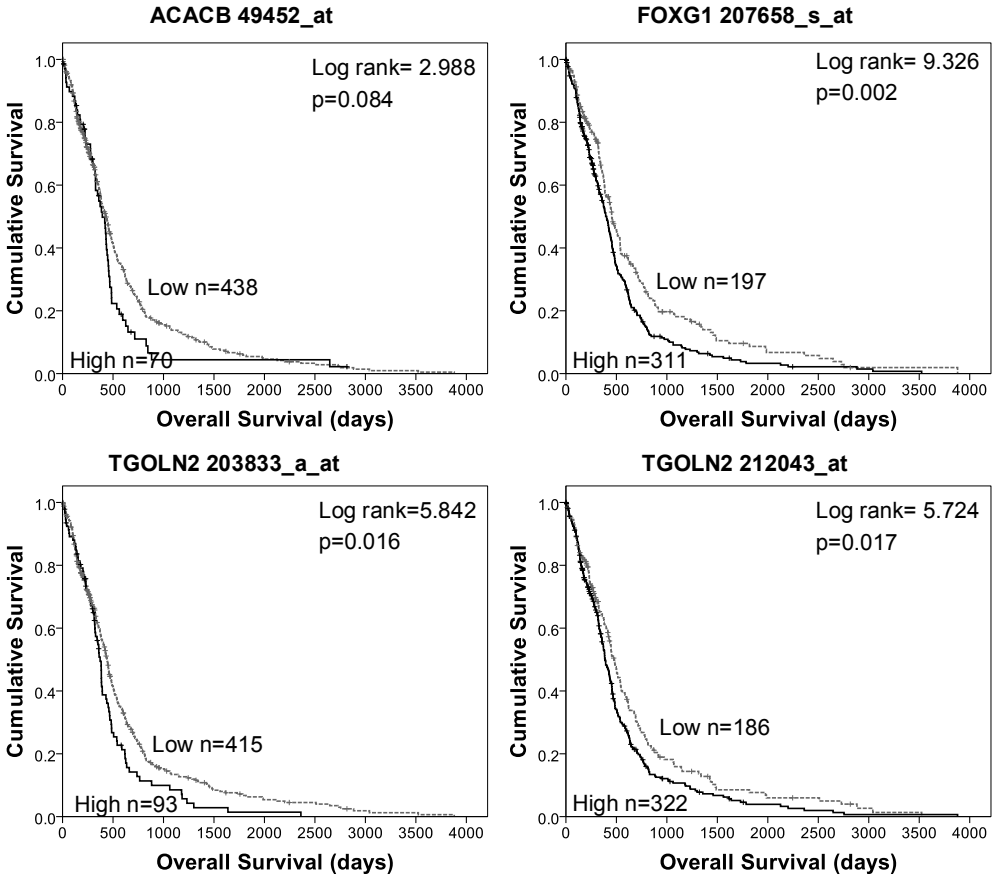


Figure 3.12. Kaplan Meier survival curves for the top 3 hubs within the APE1 interactome.

On survival analysis of the top 3 hubs (represented by 4 probes) in the APE1 interactome, high FOXG1 expression (p=0.002) and high TGOLN2 expression (203833_a_at p=0.016, 212043_at p=0.017) are associated with poor overall survival. Xtile software was used to determine biologically relevant cut-off levels for classification of expression into high and low groups.

3.2.2.2 Artificial neural network analysis for MGMT

3.2.2.2.1 Top 200 probes that predict MGMT expression

The top 200 probes that predict *MGMT* expression were identified; the top ten probes are shown in Table 3.10 alongside their biological function. One DNA repair gene was identified within the top 200 *MGMT* predicting genes which was *SMARCA4* and was ranked 125.

Table 3.10. Top 10 probes that predict MGMT expression.

Probe	Gene symbol	Protein name	Function
201925_s_at 201926_s_at	*CD55*	Complement decay-accelerating factor	Involved in activation and regulation of complement.
213572_s_at 212268_at	*SERPINB1*	Serpin peptidase inhibitor, clade B (ovalbumin), member 1	Negative regulator of endopeptidase activity.
218967_s_at	*PTER*	Phosphotriesterase related protein	Acts as a divalent metal cation with zinc ion binding capacity. Possesses hydrolase activity that acts on ester bonds.
201079_at	*SYNGR2*	Synaptogyrin 2	Potential role in regulating membrane traffic in non-neuronal cells.
210715_s_at	*SPINT2*	Kunitz-type protease inhibitor 2	Inhibitor of HGF activator, plasmin, plasma and tissue kallikrein, and factor Xia.
219298_at	*ECHDC3*	Enoyl-CoA hydratase domain-containing protein 3, mitochondrial	Plays a role in fatty acid biosynthesis.
221760_at	*MAN1A1*	Mannosyl-oligosaccharide 1,2-alpha-mannosidase IA	Role in maturation of Asn-linked oligosaccharides.
202720_at	*TES*	Testis derived transcript (3 LIM domains)	Negative regulator of cell growth and may act as a tumour suppressor. Acts as a scaffold protein and may play a role in cell adhesion, cell spreading and in the reorganization of the actin cytoskeleton.

3.2.2.2.2 Interaction analysis for MGMT

An interaction analysis was performed for the top 100 probes shown to predict *MGMT* expression. No genes were identified in the interactome that are involved in DNA repair.

The *MGMT* interactome is shown in Figure 3.13. Five nodes were identified within the interactome with ≥5 edges. The genes representing these nodes are shown in Table 3.11 with their biological functional group. The top three gene hubs within the *MGMT* interactome are *HPRT*, *DCLK2* and *TANC2*.

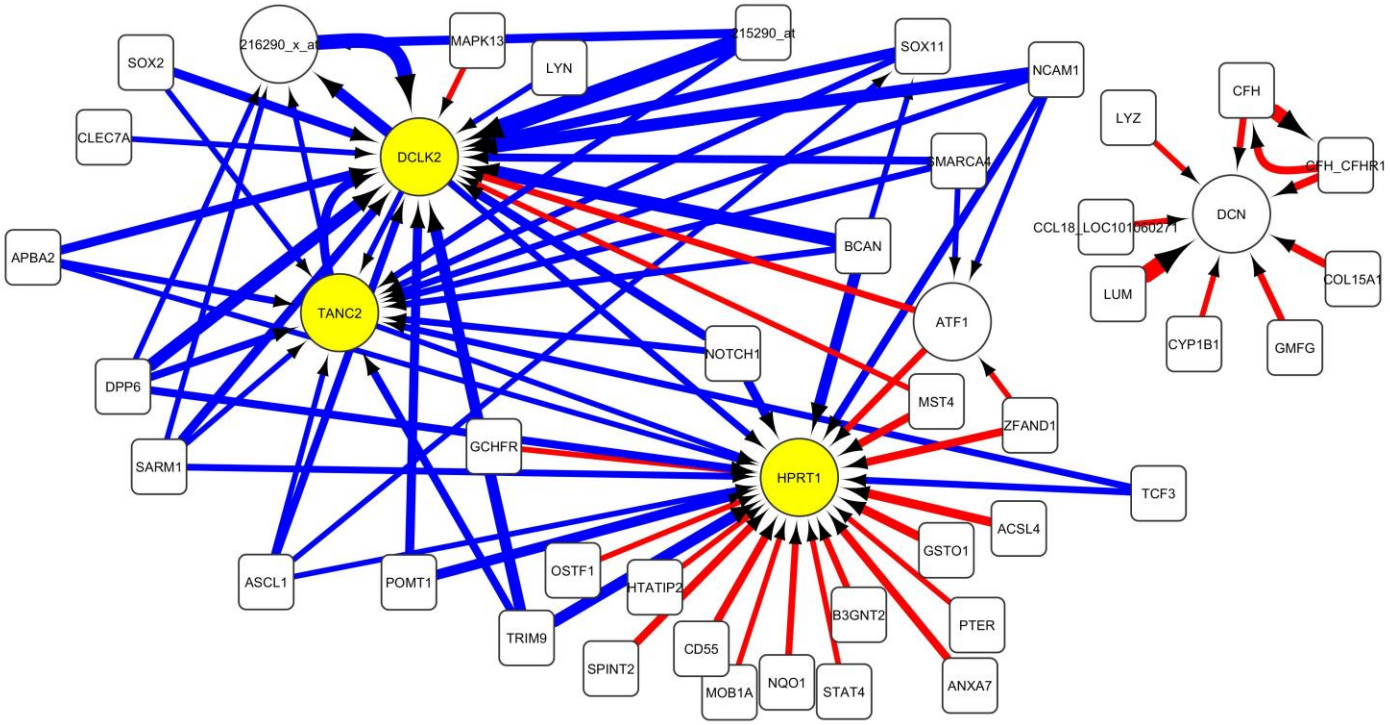


Figure 3.13. *MGMT* interactome.

The interactome includes the top 100 *MGMT* interactions. Red lines represent positive interactions and blue lines negative interactions. Wider lines signify a stronger interaction. Genes that interact with greater than 5 genes are referred to as hubs and are represented by a circle. The top 3 hubs (i.e. genes with the greatest number of interacting genes) are shown as a filled yellow circle; in this example *DCLK2*, *HPRT1* and *TANC2*.

Table 3.11. Function of gene hubs within the MGMT interactome.

Gene symbol	Protein name	No. of edges	Function
HPRT1	Hypoxanthine-guanine phosphoribosyltransferase 1 protein	28	Catalyses conversion of hypoxanthine to inosine monophosphate and guanine to guanosine monophosphate via transfer of the 5-phosphoribosyl group from 5-phosphoribosyl 1-pyrophosphate.
DCLK2	Serine/threonine-protein kinase DCLK2	22	Potential role in down-regulation of CRE-dependent gene activation.
TANC2	Protein TANC2	17	Involved in *in utero* embryonic development.
DCN	Decorin	6	Potential role in fibril formation.
ATF1	Cyclic AMP-dependent transcription factor ATF-1	5	Binds to the cAMP response element and mediates stimulation of CRE-reporter genes. Triggers cell proliferation and transformation.

3.2.2.2.3 HPRT expression predicts survival in GBM

On Kaplan Meier survival analysis for the three key hubs in the *MGMT* interactome, *HPRT* expression (p=0.018) was significantly associated with overall survival in the TCGA cohort. *TANC2* (p=0.135) and *DCLK2* expression (p=0.295) were not associated with survival (Figure 3.14).

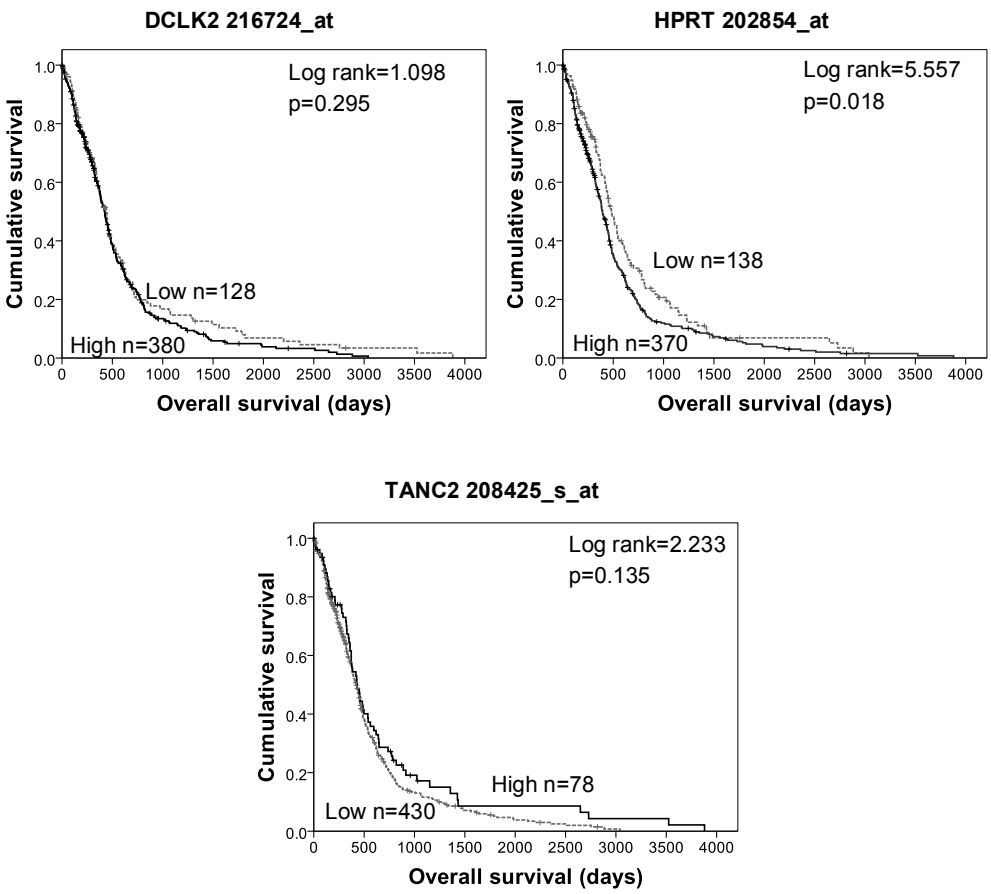


Figure 3.14. Kaplan Meier survival curves for the top 3 hubs within the MGMT interactome.

On survival analysis of the top 3 hubs in the MGMT interactome only high HPRT expression (p=0.018) is associated with poor overall survival. Xtile software was used to determine biologically relevant cut-off levels for classification of expression into high and low groups.

3.2.2.3 Artificial neural network analysis for NBN

3.2.2.3.1 Top 200 probes that predict NBN expression

The top 200 probes that predict *NBN* expression were identified; the top ten probes are shown in Table 3.12 alongside their biological function. Four genes were identified that are involved in DNA repair: *MBD4*, *SMC3*, *SMC5* and *CHD4*. The function of these genes and their rank in predicting *NBN* expression are summarised in Table 3.13.

Table 3.12. Top 10 probes that predict NBN expression.

Probe	Gene symbol	Protein name	Function
212332_at	*RBL2*	Retinoblastoma-like protein 2	Key regulator of entry into cell division. Potential tumour suppressor.
213328_at	*NEK1*	Serine/threonine-protein kinase Nek1	Implicated in control of meiosis. Involved in cilium assembly. Limits mitochondrial cell death in response to injury.
202065_s_at	*PPFIA1*	Protein tyrosine phosphatase, receptor type, f polypeptide (PTPRF), interacting protein (liprin), alpha 1	Potential role in cell-matrix interactions.
209996_x_at	*PCM1*	Pericentriolar material 1 protein	Involved in centromere assembly and function.
201728_s_at	*KIAA0100*	Protein KIAA0100	Possible role in membrane trafficking. Associated with in situ carcinoma and breast cancer.
210693_at	*SPPL2B*	Signal peptide peptidase-like 2B	Cleaves type II membrane signal peptides. Functions in ITM2B and TNF processing. Potential role in regulation of immunity.
219576_at	*MAP7D3*	MAP7 domain-containing protein 3	Involved in assembly and stability of microtubules.
211360_s_at	*ITPR2*	Inositol 1,4,5-trisphosphate receptor type 2	Receptor for inositol 1,4,5-trisphosphate, a second messenger that mediates the release of intracellular calcium.
214305_s_at	*SF3B1*	Splicing factor 3B subunit 1	Involved in mRNA processing.
202301_s_at	*RSRC2*	Arginine/serine-rich coiled-coil protein 2	Involved in mRNA splicing.

Table 3.13. DNA repair genes present in the top 200 NBN predicting genes.

Rank	Gene symbol	Protein name	Function
24	*SMC5*	Structural Maintenance Of Chromosomes 5	Core component of the SMC5-SMC6 complex, involved in the repair of DNA DSBs by homologous recombination
60	*CHD4*	Chromodomain Helicase DNA Binding Protein 4	Component of the histone deacetylase NuRD complex which is involved in the remodelling of chromatin by deacetylating histones
144	*MBD4*	Methyl-CpG Binding Domain Protein 4	Mismatch-specific DNA N-glycosylase
199	*SMC3*	Structural Maintenance Of Chromosomes 3	Central component of cohesion, a complex required for chromosome cohesion during cell cycle

3.2.2.3.2 Interaction analysis for NBN

An interaction analysis was performed to assess how the top 200 probes that predict *NBN* expression interact with each other. The top 100 probe-probe interactions were identified. Only one gene was identified that has a role in DNA repair which is *MBD4*.

The interactome for *NBN* is shown in Figure 3.15. Six genes are identified as hubs (genes that connect to ≥5 other genes). The genes representing these hubs are shown in Table 3.14 along with their biological function and the number of genes with which they interact. The three top hubs were *THRA*, *RFX4*, and *CD55*.

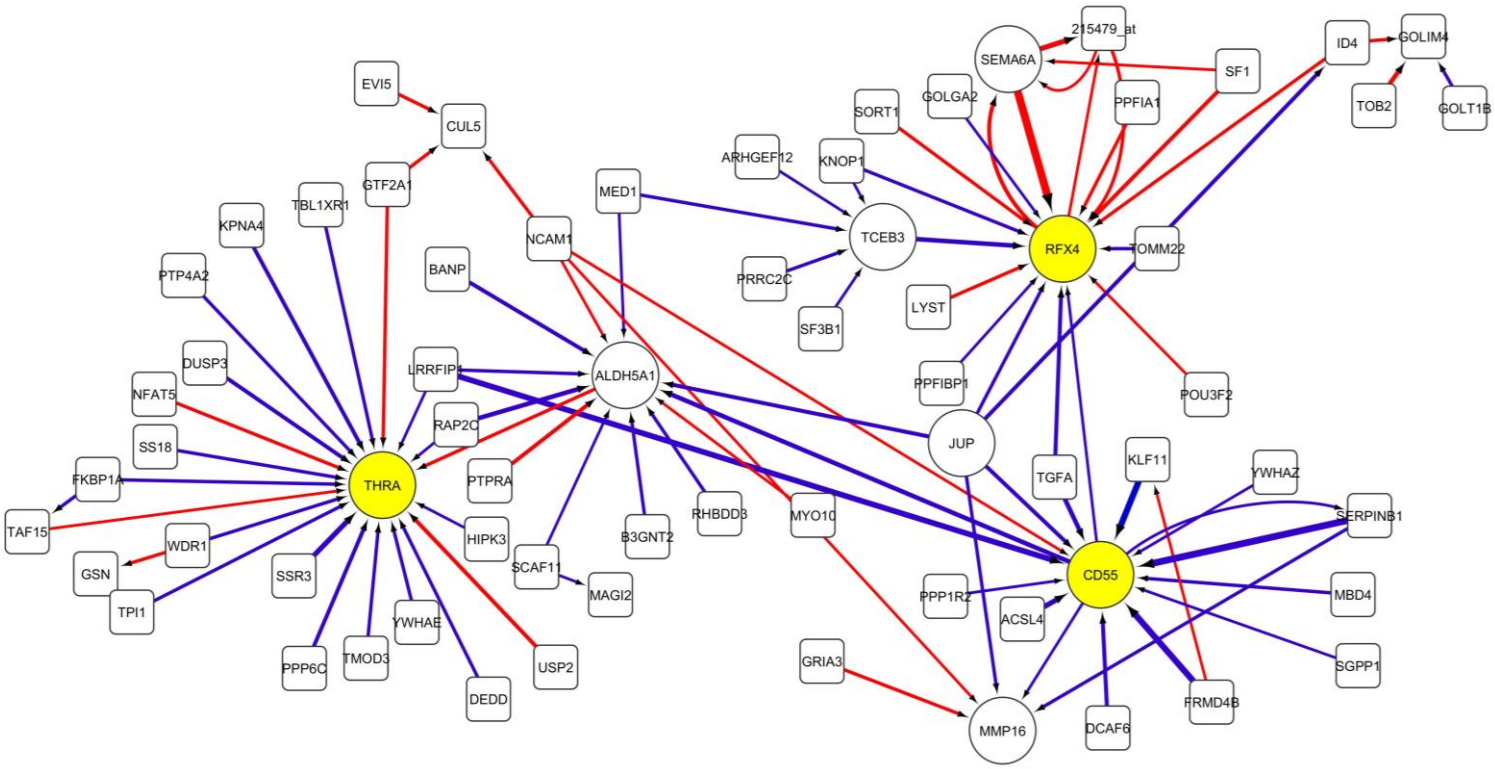


Figure 3.15. *NBN* interactome.

The interactome includes the top 100 *NBN* interactions. Red lines represent positive interactions and blue lines negative interactions. Wider lines signify a stronger interaction. Genes that interact with greater than 5 genes are referred to as hubs and are represented by a circle. The top 3 hubs (i.e. genes with the greatest number of interacting genes) are shown as a filled yellow circle; in this example *CD55*, *RFX4* and *THRA*.

Table 3.14. Function of gene hubs within the NBN interactome.

Gene symbol	Protein name	No. of edges	Function
THRA	Thyroid Hormone Receptor, Alpha	21	Nuclear hormone receptor for triiodothyronine.
RFX4	Transcription factor RFX4	18	May activate transcription by direct interaction with the X-box.
CD55	Complement decay-accelerating factor	17	Involved in activation and regulation of complement.
ALDH5A1	Succinate-semialdehyde dehydrogenase, mitochondrial	13	Catalyzes one step in the degradation of the inhibitory neurotransmitter gamma-aminobutyric acid.
TCEB3	Transcription Elongation Factor B (SIII), Polypeptide 3	6	Acts as general transcription elongation factor.
SEMA6A	Semaphorin-6A	5	Cell surface receptor for PLXNA2 that plays an important role in cell-cell signalling.

3.2.2.3.3 THRA expression is associated with survival in GBM

THRA expression was identified as the top hub in the *NBN* interactome and on Kaplan Meier survival analysis low *THRA* expression was significantly associated with poor overall survival in GBM (p=0.008). Neither *RFX4* (p=0.143) nor *CD55* (p=0.058) expression were associated with survival in the TCGA cohort (Figure 3.16).

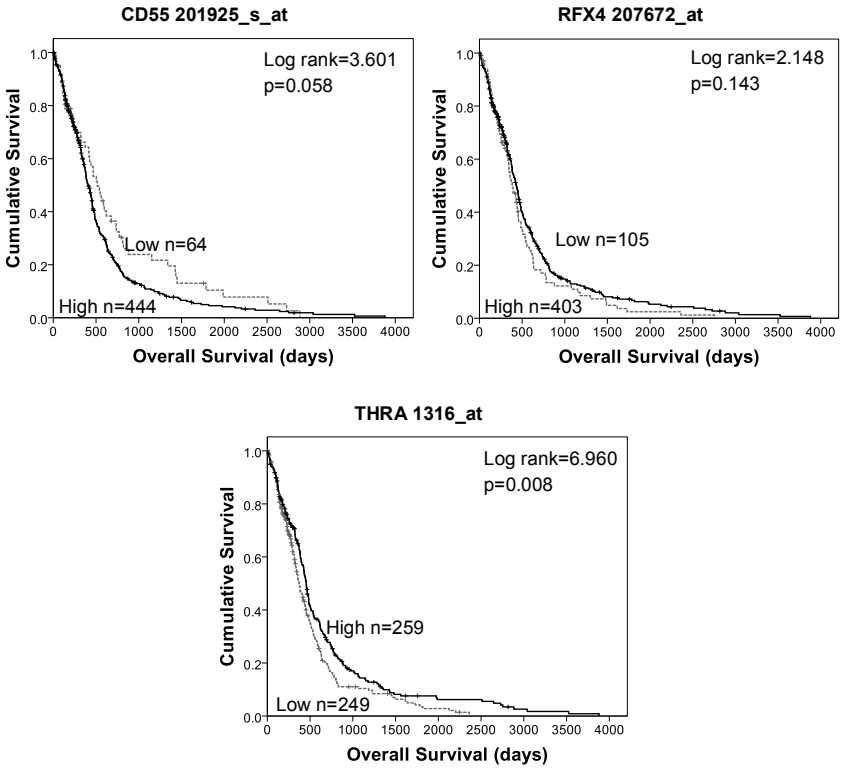


Figure 3.16. Kaplan Meier survival curves for the top 3 hubs within the NBN interactome.

On survival analysis of the top 3 hubs in the NBN interactome, only low THRA expression is associated with poor overall survival (p=0.008). Xtile software was used to determine biologically relevant cut-off levels for classification of expression into high and low groups.

3.2.2.4 Artificial neural network analysis for PMS2

3.2.2.4.1 Top 200 probes that predict PMS2 expression

The top 200 probes that predict *PMS2* expression were identified; the top ten probes are shown in Table 3.15 alongside their biological function. Three DNA repair genes were identified within the top 200 *PMS2* predicting genes which were: *PMS2P3*, *BLM* and *MSH5*. These genes are summarised in Table 3.16, including their function and rank within the top 200. *PMS2P3* is represented by 3 probes: 214473_x_at (ranked 15), 216111_x_at (ranked 28) and 216525_x_at (ranked 99).

Table 3.15. Top 10 probes that predict PMS2 expression.

Probe	Gene symbol	Protein name	Function
213097_s_at	*DNAJC2*	DnaJ (Hsp40) homolog, subfamily C, member 2	Acts as a chaperone protein for nascent polypeptide chains as they exit the ribosome.
219571_s_at	*ZNF12*	Zinc finger protein 12	Acts as a transcriptional repressor of activation protein I and serum response element-mediated transcriptional activity.
215412_x_at 214757_at (*PMS2L2*)	*LOC100132832 /PMS2L2/ PMS2P5*	Putative postmeiotic segregation increased 2 (PMS2) protein 2 & 5	Role in mismatch repair.
217754_at	*DDX56*	DEAD (Asp-Glu-Ala-Asp) box helicase 56	Possesses intrinsic ATPase activity. Potential role in processing of pre-ribosomal particles.
212178_s_at	*POM121 /POM121C*	Nuclear envelope pore membrane protein POM 121 (C)	Critical component of the nuclear pore complex.
208503_s_at 213018_at	*GATAD1*	GATA zinc finger domain containing protein 1	Component of chromatin complex. Recruited to chromatin sites with methylated 'Lys-4' of histone H3
218425_at	*RNF216*	E3 ubiquitin-protein ligase RNF216	Involved in protein ubiquitination, regulation of the antiviral response and regulation of interferon-beta production.
217736_s_at	*EIF2AK1*	Eukaryotic translation initiation factor 2-alpha kinase 1	Inhibits protein synthesis at the initiation level in response to stress.

Table 3.16. DNA repair genes present in the top 200 PMS2 predicting genes.

Rank	Gene symbol	Protein name	Function
15, 28, 99	*PMS2P3*	Putative postmeiotic segregation increased 2-like protein 3	Involved in mismatch repair and regulation of transcription.
60	*MSH5*	MutS protein homolog 5	Involved in meiotic recombination and mismatch DNA repair.
76	*BLM*	Bloom syndrome protein	Involved in DNA replication and double strand break repair.

3.2.2.4.2 Interaction analysis for PMS2

An interaction analysis was performed for the top 100 probes shown to predict *PMS2* expression. Only one gene was identified that has a role in DNA repair which is *BLM*.

The *PMS2* interactome is shown in Figure 3.17. Seven nodes were identified within the interactome with ≥5 edges. The genes representing these nodes are shown in Table 3.17 with their biological functional group. The top three hubs within the *PMS2* interactome are *STXBP6*, *PTER* and *ACSL4*.

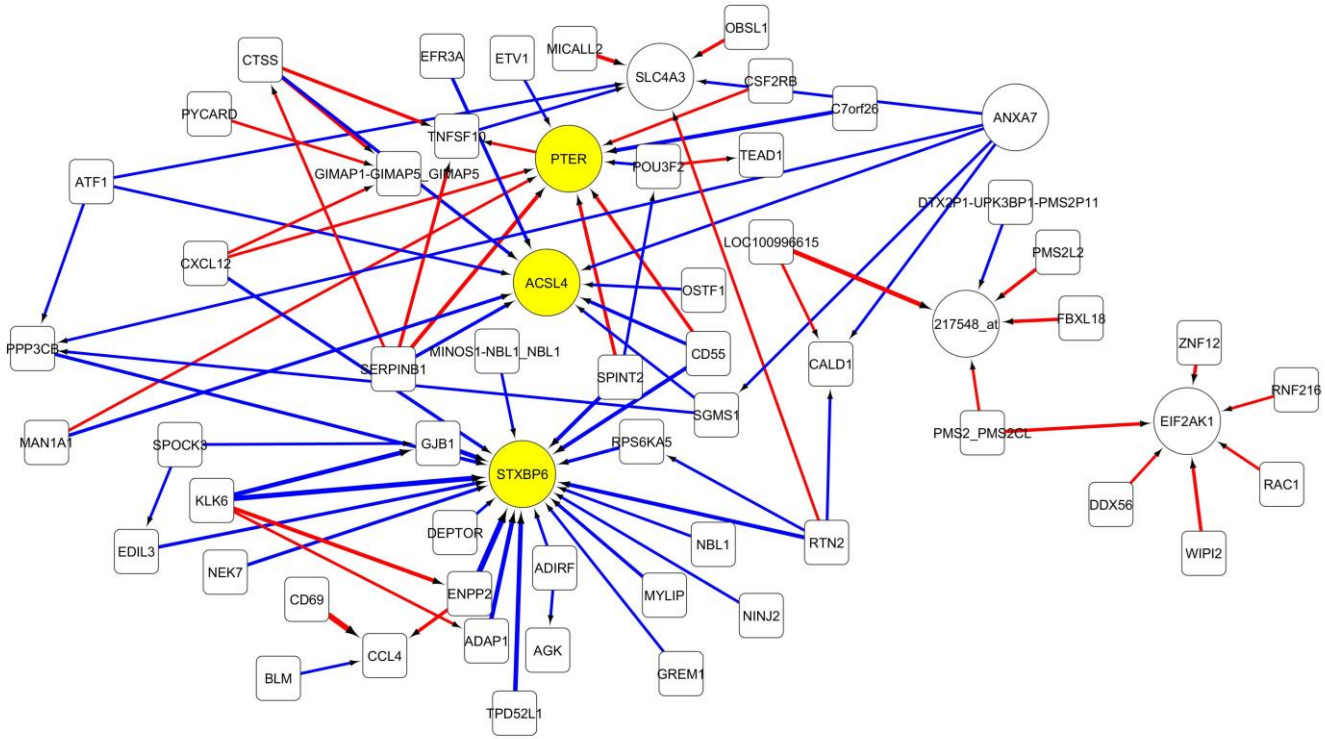


Figure 3.17. *PMS2* interactome.

The interactome includes the top 100 *PMS2* interactions. Red lines represent positive interactions and blue lines negative interactions. Wider lines signify a stronger interaction. Genes that interact with greater than 5 genes are referred to as hubs and are represented by a circle. The top 3 hubs (i.e. genes with the greatest number of interacting genes) are shown as a filled yellow circle; in this example *ACSL4*, *PTER* and *STXBP6*.

Table 3.17. Function of gene hubs within the PMS2 interactome.

Gene symbol	Protein name	No. of edges	Function
STXBP6	Syntaxin Binding Protein 6 (Amisyn)	20	Potential role in the modulation of functional SNARE complexes and exocytosis.
PTER	Phosphotriesterase Related protein	10	Acts as a divalent metal cation with zinc ion binding capacity. Possesses hydrolase activity that acts on ester bonds.
ACSL4	Long-chain-fatty-acid--CoA ligase 4	9	Involved in activation of long-chain fatty acids.
EIF2AK1	Eukaryotic Translation Initiation Factor 2-Alpha Kinase 1	6	Inhibits protein synthesis at the translation initiation level in response to stress.
SLC4A3	Anion exchange protein 3	6	Plasma membrane associated anion exchange protein. Involved in part of $Cl^-/HCO3^-$ exchange in cardiac monocytes.
217548_at	Unidentifiable probe	5	Unknown.
ANXA7	Annexin A7	5	Involved in calcium/phospholipid-binding promoting membrane fusion. Also involved in exocytosis.

3.2.2.4.3 PTER, ACSL4 and STXBP6 expression are not associated with survival in GBM

On Kaplan Meier survival analysis of the three most important hubs from the *PMS2* interactome neither *PTER* (p=0.052), *ACSL4* (p=0.341) or *STXBP6* (p=0.266) were associated with survival (Figure 3.18).

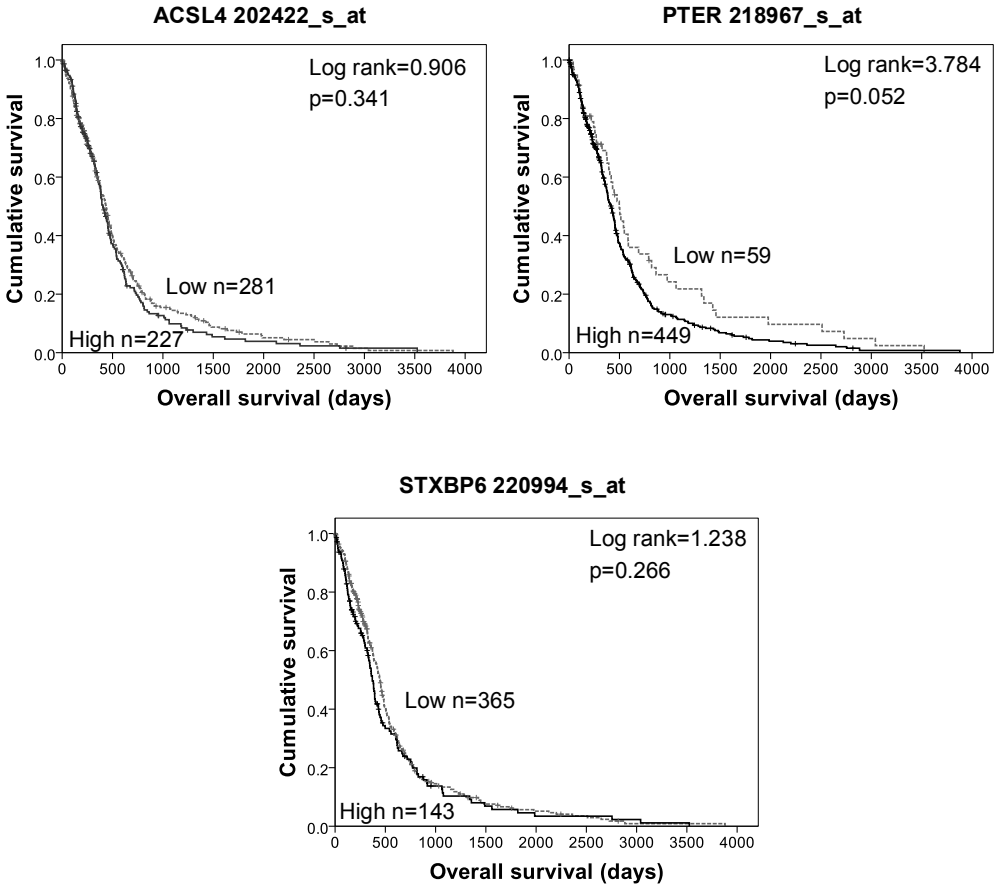


Figure 3.18. Kaplan Meier survival curves for the top 3 hubs within the PMS2 interactome.

On survival analysis of the top 3 hubs in the PMS2 interactome none of these genes are associated with overall survival. Xtile software was used to determine biologically relevant cut-off levels for classification of expression into high and low groups.

3.2.2.5 Artificial neural network analysis for PTEN

3.2.2.5.1 Top 200 probes able to predict PTEN expression

The top 200 probes that predict *PTEN* expression were identified; the top ten probes are shown in Table 3.18 alongside their biological function. Two genes were identified that have a role in DNA repair, *NIPBL* and *SMC3*. *NIPBL* was represented by two probes and ranked 34 (213918_s_at) and 65 (207108_s_at) in the list of genes able to predict *PTEN* expression and *SMC3* was ranked 125.

Table 3.18. Top 10 probes that predict PTEN expression.

Probe	Gene symbol	Protein name	Function
209864_at	*FRAT2*	GSK-3-binding protein FRAT2	Positive regulator of Wnt signalling pathway by stabilisation of beta-catenin.
#202730_s_at #202731_at	*MIR4680 / PDCD4*	MicroRNA 4680/ Programmed cell death protein 4	Inhibits translation. Modulates activity of MAP kinase pathway. Possible role in apoptosis.
*212264_s_at *212267_at	*WAPAL*	Wings apart-like protein homolog	Regulator of sister chromatid cohesion in mitosis.
219326_s_at	*B3GNT2*	Beta-1,3-N-acetyl glucosaminyltransferase	Involved in the biosynthesis of poly-N-acetyllactosamine chains.
214578_s_at	*ROCK1*	Rho-associated protein kinase 1	Key regulator of actin cytoskeleton and cell polarity.
217758_s_at	*TM9SF3*	Transmembrane 9 superfamily member 3	Not recorded
206875_s_at	*SLK*	STE20-like serine/threonine-protein kinase	Mediator of apoptosis and actin stress fibre dissolution
202777_at	*SHOC2*	Leucine-rich repeat protein SHOC-2	Regulatory subunit of protein phosphatase 1. Involved in activation of MAPK pathway.

*212264_s_at ranked 3 and 212267_at ranked 10. #202730_s_at ranked 2 and 202731_at ranked 4.

3.2.2.5.2 Interaction analysis for PTEN

An interaction analysis was performed to assess how the top 200 probes that predict *PTEN* expression interact with each other. The two DNA repair genes, *NIPBL* and *SMC3*, identified previously as key predictors of *PTEN* expression appear in the interactome.

The interactome for the top 100 interactions is shown in Figure 3.20. Four hubs can be seen within the interactome (genes that connect to ≥5 other genes). The four hubs include *CDH1*, *ACP5* and *HIF1AN*, with the fourth hub being a non-specific probe (215667_x_at), and has therefore been excluded from subsequent analyses. The function of the three hubs, alongside the number of interacting genes (number of edges), is shown in Table 3.19.

Table 3.19. Function of gene hubs within the PTEN interactome.

Gene symbol	Gene name	No. of edges	Function
CHD1	Chromodomain Helicase DNA Binding Protein 1	38	ATP-dependent chromatin-remodelling factor. Regulator of transcription. Modulates efficiency of pre-mRNA splicing.
ACP5	Acid Phosphatase 5, Tartrate Resistant	26	Catalyses the conversion of orthophosphoric monoester to alcohol and orthophosphate.
HIF1AN	Hypoxia Inducible Factor 1, Alpha Subunit Inhibitor	14	Oxygen sensor. In normoxic conditions prevents interaction of HIF1 with transcriptional co-activators. Negative regulator of NOTCH signalling pathway.

3.2.2.5.3 HIF1AN expression is associated with survival in GBM

On Kaplan Meier survival analysis of *HIF1AN*, *ACP5* and *CHD1* expression in the TCGA cohort only *HIF1AN* expression was associated with survival; patients with low *HIF1AN* expression had a worse prognosis than patients with high *HIF1AN* expression (p=0.024). Both *ACP5* (p=0.257) and *CHD1* expression (p=0.120) were not associated with overall survival (Figure 3.20).

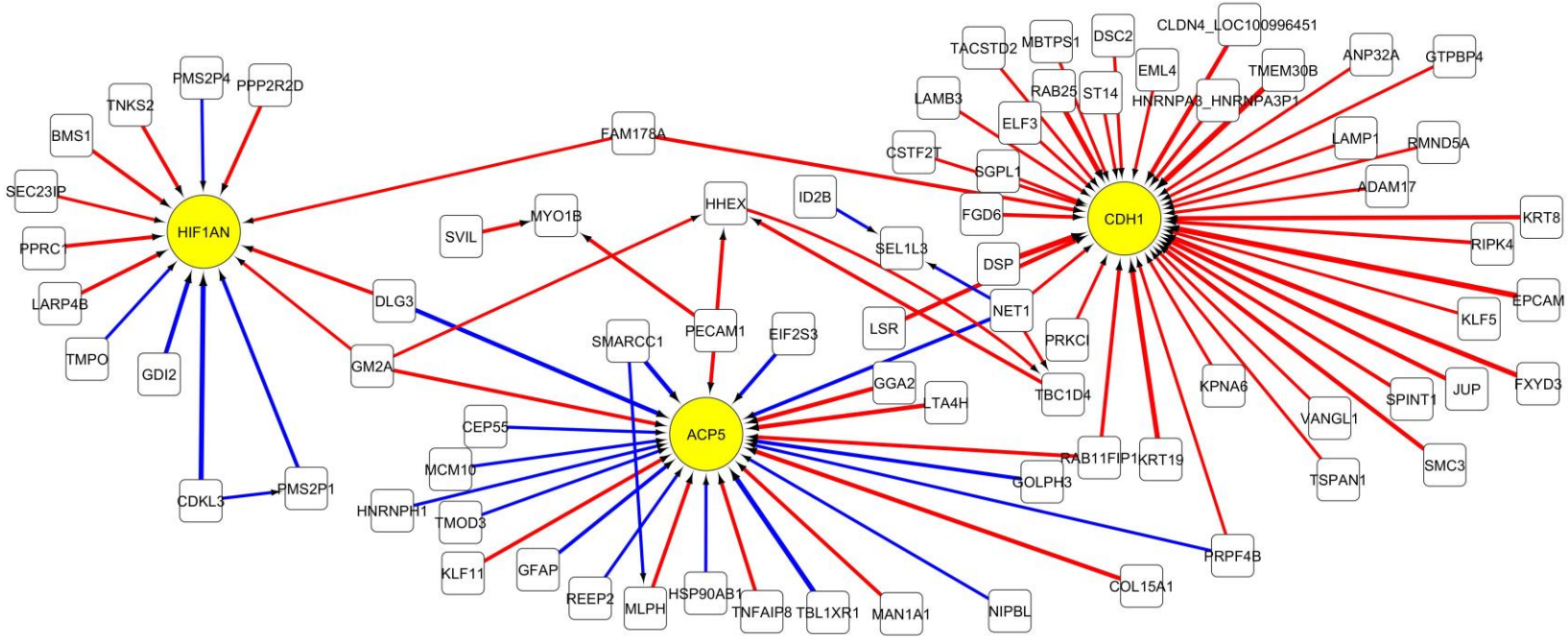


Figure 3.19. *PTEN* interactome.

The interactome includes the top 100 *PTEN* interactions. Red lines represent positive interactions and blue lines negative interactions. Wider lines signify a stronger interaction. Genes that interact with greater than 5 genes are referred to as hubs and are represented by a circle. The top 3 hubs (i.e. genes with the greatest number of interacting genes) are shown as a filled yellow circle; in this example ACP5, CDH1 and HIF1AN.

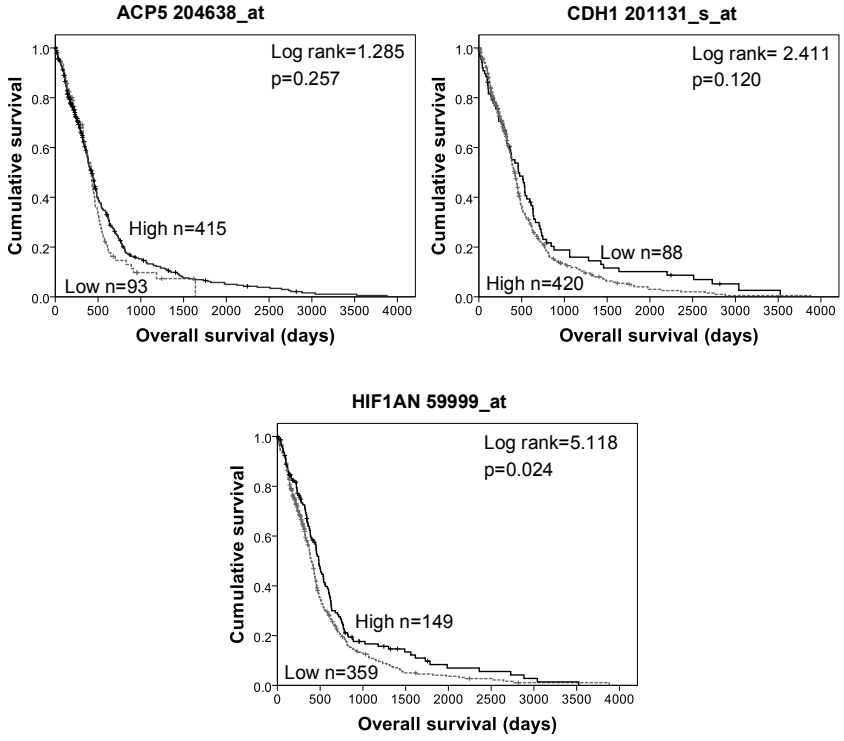


Figure 3.20. Kaplan Meier survival curves for the top 3 hubs within the PTEN interactome.

On survival analysis of the top 3 hubs in the PTEN interactome, only low HIF1AN expression is associated with poor overall survival (p=0.024). Xtile software was used to determine biologically relevant cut-off levels for classification of expression into high and low groups.

3.2.2.6 Multivariate analysis of APE1, MGMT, NBN, PMS2 and PTEN and the top 15 hubs from ANN analysis

A multivariate cox regression analysis was performed including *APE1*, *MGMT*, *NBN*, *PMS2* and *PTEN* plus the top 15 hubs from the five ANN analyses (Table 3.20). Of note *APE1*, *MGMT*, *NBN*, *PMS2* and *PTEN* retained significance plus *FOXG1*, *TOGLN2*, *DCLK2*, *THRA*, *RFX4*, *STXBP6* and *HPRT* were also found to be independent predictors of survival. Of the 20 genes studied, *FOXG1* ($p<0.000001$) was the most highly significant independent predictor of survival in GBM.

Table 3.20. Multivariate analysis for APE1, MGMT, NBN, PMS2 and PTEN expression plus the 15 gene hubs.

Gene	HR (95% CI)	*P* value
FOXG1	0.560 (0.449-0.699)	**<0.000001**
NBN	1.757 (1.317-2.344)	**0.000128**
APE1	1.888 (1.361-2.619)	**0.000141**
MGMT	0.650 (0.503-0.838)	**0.001**
PTEN	1.573 (1.206-2.051)	**0.001**
PMS2	0.726 (0.572-0.922)	**0.009**
RFX4	1.445 (1.091-1.913)	**0.010**
STXBP6	0.739 (0.580-0.942)	**0.014**
TOGLN2	0.715 (0.546-0.937)	**0.015**
DCLK2	0.726 (0.557-0.945)	**0.017**
HPRT1	0.752 (0.584-0.968)	**0.027**
THRA	1.252 (1.006-1.558)	**0.044**
ACSL4	1.242 (0.978-1.578)	0.076
HIF1AN	1.228 (0.976-1.546)	0.079
ACACB	0.819 (0.607-1.106)	0.192
PTER	1.247 (0.882-1.762)	0.211
CD55	0.841 (0.611-1.157)	0.286
CDH1	0.909 (0.695-1.188)	0.485
TANC2	0.920 (0.680-1.245)	0.590
ACP5	1.051 (0.800-1.381)	0.720

Significant p values (≤0.05) are shown in bold. HR = hazard ratio. CI = confidence interval.

3.2.2.7 Additional DNA repair genes were identified that are associated with survival

As a result of the interaction analysis for *APE1*, *NBN*, *MGMT*, *PTEN* and *PMS2*, five further DNA repair-associated genes were identified within the interaction networks. Kaplan Meier survival analyses were performed for each of these DNA repair-associated genes and are shown in Figure 3.21. High *BLM* (p=0.015), high *MBD4* (p=0.005), low *MNAT* (p=0.003) and low *SMC3* (p=0.000490) expression were associated with poor overall survival. *NIPBL* expression was not significantly associated with survival on Kaplan Meier analysis.

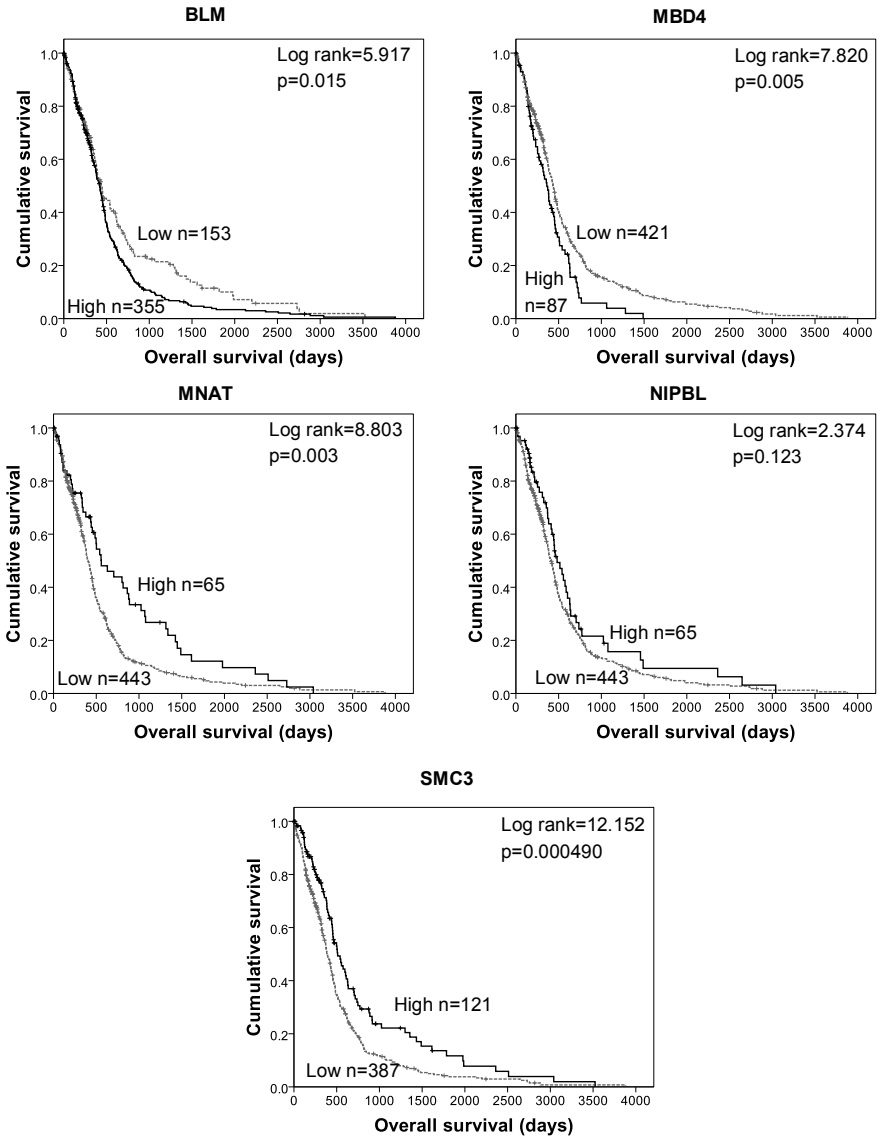


Figure 3.21. Kaplan Meier survival curves for the 5 additional DNA repair genes identified in the interaction analyses.

On survival analysis of the 5 additional DNA repair genes identified from the interaction analyses high BLM (p=0.015), high MBD4 (p=0.005), low MNAT (p=0.003) and low SMC3 (p=0.000490) expression are associated with poor overall survival. Xtile software was used to determine biologically relevant cut-off levels for classification of expression into high and low groups.

3.3 Discussion and conclusions

3.3.1 APE1, MGMT, NBN, PMS2 and PTEN are identified as the most important DNA repair genes for predicting prognosis in adult GBM

This is the first study to explore the relationship between a large panel of DNA repair genes and survival in GBM. Twelve key DNA repair genes have been identified that are associated with survival, representing most of the major DNA repair pathways. Of these twelve genes 25% are involved in the BER pathway and 25% in the NER pathway, suggesting that the functioning of these pathways are important in GBM. For six of these twelve DNA repair genes, high expression is associated with poor survival, including *MGMT*, *BRCA1*, *PMS2*, *PARP3*, *DDB2* and *RAD23B*. These may therefore represent genes that are associated with treatment resistance with higher levels of expression being associated with repair of damage induced by chemotherapy or radiotherapy (Croteau, 2013). Several of these DNA repair genes have been previously studied with regards to their association with malignancy and response to treatment. As described earlier, *MGMT* expression is probably the most extensively studied DNA repair gene in terms of glioma prognosis and response to treatment. A non-methylated *MGMT* promoter, and the associated higher levels of *MGMT* expression, have been shown to be associated with a worse outcome for patients treated with both chemotherapy and radiotherapy as compared to patients with a methylated promoter and lower levels of *MGMT* (Costa et al., 2010, Hegi et al., 2005, Rivera et al., 2010, Zhang et al., 2013). The role of *BRCA1* and treatment resistance has also been studied previously, predominately in breast cancer. In the pre-clinical and clinical settings, *BRCA1*

mutations have been found to be associated with increased sensitivity to chemotherapeutic agents. In patients with sporadic breast cancer, the association between BRCA1 expression and sensitivity to chemotherapy is less clear (Reviewed in (Kennedy et al., 2004)). High DDB2 expression, in combination with elevated levels of ERCC1, has previously been linked to chemo-resistance in a Phase II clinical trial of advanced resectable gastric cancer patients treated with doxcetaxel, cisplatin and S-1 chemotherapy (Hirakawa et al., 2013).

No studies have investigated PARP3 expression alone as a marker of prognosis in cancer. PARP3 has, however, previously been proposed to be a negative regulator of telomerase activity and therefore cells with higher levels of PARP3 would be expected to have lower levels of telomerase (Fernandez-Marcelo et al., 2014). Interestingly, tumours with lower levels of telomerase have generally been shown to have a better prognosis than tumours with higher telomerase levels (Le et al., 1998, Tatsumoto et al., 2000, Hiyama et al., 1995). This would appear to go against findings in the current study, where high PARP3 expression was found to be associated with poor overall survival.

In an earlier study, low RAD23B protein expression was shown to be associated with aggressive features in breast cancer (Linge et al., 2014) whereas in the current study high *RAD23B* expression was associated with poor overall survival. Similarly, in the current study, high *PMS2* expression was associated with poor overall survival whereas earlier studies found low expression of *PMS2* to be associated with recurrent glioma (Felsberg et al., 2011). This would suggest that further work is required to investigate the role

of PARP3, RAD23B and PMS2 expression in GBM. PMS2 will be discussed further in Chapter 4 where protein expression levels will be studied in high grade glioma patients.

Low levels of *APE1*, *PARP2*, *ERCC6*, *RAD21*, *PTEN* and *NBN* were found to be associated with poor survival. This suggests that these genes may play a role in gliomagenesis and promote an unstable, hypermutated phenotype resulting in a more aggressive tumour. The observation that low *PTEN* expression is associated with poor survival in GBM is in keeping with the results from earlier studies (Ermoian et al., 2002b). *ERCC6* mutations have been reported previously in glioma, with mutations/deletions of exon II found to be associated with increased glioma grade (Zhang, 2000). Defects in the *NBN* gene are known to result in Nijmengen breakage syndrome. This is a rare autosomal dominant condition that is characterised by microcephaly and a predisposition to cancer, including glioma (Group, 2000), supporting a role for *NBN* in gliomagenesis. Previous studies investigating APE1 expression have often shown high APE1 expression to be associated with poor survival (Wang et al., 2004, Al-Attar et al., 2010). However, in a large study of breast cancer patients, low APE1 protein expression was found to be associated with aggressive features, resistance to endocrine treatment and poor survival (Abdel-Fatah et al., 2014b) in keeping with the results in this current study. APE1, NBN and PTEN will be discussed in further detail in Chapter 4, where protein expression will be explored in a cohort of high grade glioma patients.

While PARP1 has been extensively studied in terms of its association with cancer, less is known about PARP2 (Yelamos et al., 2011). The finding that

low PARP2 expression is associated with poor prognosis in GBM is therefore a potential novel finding. Similarly, RAD21 expression has not previously been studied in the context of glioma. A study in colorectal cancer, however, found RAD21 expression to be associated with aggressive features and resistance to chemo-radiotherapy, particularly in tumours with *KRAS* mutations (Deb et al., 2014). Further investigation of both PARP1 and RAD21 may therefore be warranted.

Interestingly, when compared to the study by Kassambara *et al* in multiple myeloma, none of the 22 DNA repair genes identified in the myeloma DNA repair pathway score appear in the current analysis (Kassambara et al., 2014). When compared to the DNA repair gene score developed by Kang *et al* in ovarian cancer, only high *RAD23A* expression was found to be associated with poor survival in both studies. Low *NBN* expression was also found in patients with poor survival in both this current study and the work by Kang *et al*, but did not reach statistical significance in the ovarian study. It must be noted that the ovarian study only included DNA repair genes involved in pathways that are thought to be involved in the repair of platinum-induced lesions, as it was specifically designed to predict outcome in platinum-treated ovarian cancer (Kang et al., 2012). Comparison of the GBM DNA repair gene prognostic index developed in this current study with the score developed by Abdel-Fatah *et al* in breast cancer reveals no DNA repair genes in common between the two scores. Again it should be noted that the score developed in breast cancer patients only profiled a select group of DNA repair genes (Abdel-Fatah et al., 2015). The available evidence would therefore suggest that DNA repair status

is clearly important in determining outcome in various tumour types but that the exact pattern of DNA repair gene expression that is associated with survival varies with cancer site.

After multivariate analysis, five DNA repair genes were identified that are independently associated with survival; *APE1*, *MGMT*, *NBN*, *PMS2* and *PTEN*. High expression of *PMS2* and *MGMT* was found to be associated with poor survival, whereas low expression of *APE1*, *NBN* and *PTEN* was associated with poor survival. Using a previously reported technique (Lancashire et al., 2010), expression levels of these genes were incorporated into a prognostic index that could be applied to both datasets and separate patients into three prognostic groups. This is the first description in the literature of a DNA repair gene prognostic index for GBM. Further extrapolation of the data enables prediction of 1, 2 and 3 year survival for these GBM patients based on their prognostic index score, allowing the identification of individuals with a poor prognosis. This then poses the question as to how to use this information to influence clinical outcome for patients. It may be that poor prognosis patients could be spared intensive, aggressive conventional treatment regimens and offered good palliative care or that this group could be targeted with new novel agents. Clearly this is not a question that can be answered with the results of this study and further prospective, large scale clinical studies are required. This would enable further validation of the prognostic index in a new cohort and also the collection of prospective treatment data.

One limitation of this component of the study is the lack of treatment and demographic data available for the Test dataset. Also, although treatment data

is available for the TCGA cohort, the presentation of this data makes it difficult, and often impossible, to determine exactly what treatment was delivered and in what order, and also what the outcome was of the treatment. Therefore, for the purposes of this study, patients were either defined as having received or not received chemotherapy or radiotherapy. Many of these patients will have received both chemotherapy and radiotherapy, although not necessarily at the same time. This has hindered the ability of the study to investigate the predictive significance of these DNA repair genes. However, this aside, subgroup survival analysis for each of the DNA repair genes in patients who did or did not receive radiotherapy reveals that low *APE1*, low *PTEN*, high *MGMT* and high *PMS2* expression are only associated with poor overall survival in patients that received radiotherapy. This would suggest a potential predictive role for these genes in determining patients that may benefit from radiotherapy. Similarly, subgroup analysis in patients that did and did not receive chemotherapy found that low *APE1*, high *MGMT* and high *PMS2* were only associated with poor survival in chemotherapy treated patients. This, again, would suggest that expression of these genes may predict response to chemotherapy. High levels of *PMS2* and *MGMT* expression may result in repair of DNA damage induced by chemotherapeutic agents and radiation. Interestingly *NBN* expression was associated with poor survival across all treatment groups and therefore appears to have more of a prognostic rather than predictive role.

A future large clinical trial with prospective collection of treatment data would allow the development of a predictive index. A predictive index would have

the advantage of identifying patients who are likely to respond to a particular treatment and also those who would not benefit. New treatment strategies would need to be developed for the non-responding patients, which, in the longer-term, could mean newer targeted treatment approaches, and in the short-term spare patients from receiving potentially toxic, non-beneficial treatments.

Low *APE1* expression was found to be associated with poor performance status, older age and a lack of chemotherapy and radiotherapy treatment. It is therefore possible that the association between *APE1* and survival is purely as a consequence of its association with poor prognostic features. Similarly, in patients treated with chemotherapy and radiotherapy, the association between low *APE1* expression and poor overall survival may be due to these individuals being less able to tolerate intense treatment regimens and hence survival is in fact decreased by the treatment. High *PMS2* expression was also found to be associated with older age at diagnosis in both datasets and again it may be the fact that patients with high *PMS2* expression tend to be older that is influencing survival.

However, incorporating the five DNA repair genes into a multivariate analysis, alongside known prognosis factors, revealed that low *APE1*, low *NBN*, high *MGMT* and high *PMS2* are independent predictors of survival. Older age, poor performance status and a lack of radiotherapy treatment were also found to be associated with poor overall survival, as would be expected. Interestingly, *PTEN* status was not an independent predictor of survival. A subject of future studies may be to incorporate clinical variables into the DNA repair gene prognostic score to determine if this will improve its prognostic value.

3.3.2 Various quality assurance measures help provide confidence in the results obtained

Both the Test and TCGA datasets are derived from sources with high data quality standards. The TCGA GBM dataset has been extensively published in the literature (Devarakonda et al., 2013, Doucette et al., 2013, Ma and Ellis, 2013) and is likely to be the largest and most reliable source of gene expression data for this tumour site. Use of both a Test and Validation dataset is also cited in the literature (Lancashire et al., 2010) and is a method used to reduce the likelihood of associations being found by chance, especially when screening for a large number of genes. Application of a correction for multiple comparisons, in this study the Benjamini Hochberg False Discovery Rate correction, is another method to reduce the likelihood of chance associations being taken forward. Finally, ensuring a focused search strategy is employed, for example studying all DNA repair genes, helps ensure meaningful results are obtained.

Further indicators that validate the techniques used in this study include the inclusion of genes known to be associated with survival in the final prognostic index. For example, the identification of high *MGMT* as associated with worse overall survival is in keeping with the results of previous studies (Rivera et al., 2010, Hegi et al., 2005, Costa et al., 2010), and therefore provides evidence to support the methodology used. Low *PTEN* expression has also been previously shown to be associated with poor overall survival (Ermoian et al., 2002b), providing additional validation for the techniques used. Further scrutiny of the demographic factors of the patients in the two cohorts would suggest that these individuals are broadly representative of GBM patients as a

population (Schwartzbaum et al., 2006a). Age is also known to be a prognostic factor in GBM, with poor overall survival seen in older patients (Latif et al., 1998, MRC, 1990). This is a finding that is reproduced in both datasets in this study. Similarly, patient performance status is a known prognostic factor in GBM and again in this study poor performance status is associated with poor overall survival, in keeping with the results of earlier studies (Latif et al., 1998, MRC, 1990). Taken together, these observations provide assurance that both the data used and the results obtained are valid and reliable.

3.3.3 Key drivers of the APE1, MGMT, NBN, PMS2 and PTEN networks are identified

Artificial neural network analysis has identified key genes that drive the *APE1*, *NBN*, *MGMT*, *PTEN* and *PMS2* networks. Of the fifteen gene hubs identified by ANN, five were found to be significantly associated with survival on Kaplan Meier analysis. Seven (*FOXG1*, *TOGLN2*, *DCLK2*, *THRA*, *RFX4*, *STXBP6* and *HPRT*) of the fifteen genes were significant on multivariate analysis which also included *APE1*, *NBN*, *MGMT*, *PTEN* and *PMS2*. Each of these five networks will be discussed separately.

3.3.3.1 CDH1, ACP5 and HI1AN are identified as key genes that drive the PTEN interaction network

The three main hubs within the *PTEN* interactome were *CDH1*, *ACP5* and *HIF1AN*. Of these three genes, only *HIF1AN* expression was found to be associated with overall survival, with low expression associated with poor survival. Hypoxia inducible factor 1, alpha subunit inhibitor (HIF1AN), is also

known as Factor inhibiting HIF-1 (FIH-1), and regulates HIF-1 alpha by hydroxylation of the C-terminal transactivation domain. HIF1AN senses oxygen levels and in normoxic conditions, hydroxylation of HIF-1 alpha occurs. This prevents the interaction between HIF-1 and its transcriptional coactivators (UniProt, 2015h). The association between HIF-1 and cancer, particularly response to treatment, has been widely reported (Semenza, 2003). More recent studies have investigated the relationship between HIF1AN and malignancy. Deng *et al* found that patients with locally advanced renal cell carcinoma, and HIF1AN located exclusively in the nuclear compartment, had a better progression free and overall survival than patients with negative expression of HIF1AN. The worst survival was found in individuals with exclusive cytoplasmic expression of HIF1AN (Deng et al., 2014). Similar results were also seen in breast cancer patients, with cytoplasmic location of HIF1AN associated with poor survival (Tan et al., 2007). HIF1AN has also been studied in GBM. Glioblastoma are extensively vascular tumours and the transcription of growth factors and cytokines that are responsible for angiogenesis are regulated by HIF-1. The *HIF1AN* gene is located at chromsome region 10q24 and is often deleted in GBM, implying that it may play a role in tumorgenesis and progression. A recent study by Wang *et al* found that HIF1AN inhibits HIF-1 mediated transcription of *GLUT1* and *VEGFA*, including in hypoxic human GBM cells. HIF1AN was found to be more potent than PTEN at inhibiting HIF-1 function. It has therefore been proposed that loss of HIF1AN, and the subsequent consitutive activation of HIF-1, results in increased angiogenesis and the survival of tumour cells in hypoxic conditions (Wang et al., 2014). This would be in keeping with the

findings of the current study, whereby low HIF1AN expression is associated with poor overall survival.

While neither *CDH1* nor *ACP5* expression were found to be associated with survival in this current study, both have either been linked to *PTEN* expression or malignancy in earlier studies. CDH1 is a co-activator protein of the anaphase-promoting complex/cylosome (APC/C) complex, alongside CDC20. CDH1 is active in the latter stages of mitosis and in G1 phase, whereas CDC20 exerts its effects in early mitosis. APC/C is an E3 ubiquitin ligase controlling cell cycle progression. Substrates for the APC/C complex include: mitotic cyclins and kinases, proteins involved in segregation of chromosomes, DNA replication proteins, SKP2 (an F-box protein) and transcription factors including FoxM1. Song *et al* have shown that nuclear PTEN interacts with APC/C and promotes the association of this complex with CDH1; this results in augmentation of the tumour suppressor activity of the APC-CDH1 complex (Song et al., 2011). The finding in the current study of *CDH1* as a key component of the *PTEN* interactome is therefore in agreement with the earlier study.

ACP5 plays an essential role in osteoclast differentiation and bone resorption. ACP5 has been shown to be over-expressed in hepatocellular carcinoma (HCC), with higher expression levels associated with worse survival. ACP5 has also been shown to be transcriptionally regulated by Forkhead box M1 (FoxM1) transcription factor, with co-expression of ACP5 and FoxM1 associated with poor prognosis in HCC (Xia et al., 2014). Taken together, this suggests that all three hubs may warrant further investigation in GBM and that

HIF1AN in particular should be investigated further as a potential prognostic and/or predictive marker in GBM.

3.3.3.2 ACACB, TGOLN2 and FOXG1 are identified as drivers of the APE1 interaction network

Within the *APE1* network, the three key hubs are *ACACB*, *TGOLN2* and *FOXG1*, with high expression of both *TGONL2* and *FOXG1* found to be associated with poor overall survival. On review of the literature, both *ACACB* and *FOXG1* have previously been associated with glioma. Acetyl-coenzyme A carboxylase beta (*ACACB*) is involved in fatty acid biosynthesis (UniProt, 2015e) and has been found to be significantly down-regulated in glioma-derived glial progenitor cells (GPCs) as compared to GPCs originating from normal white matter (Auvergne et al., 2013). *FOXG1* has been more extensively studied. Atypical Rett syndrome results from germ-line mutation in the *FOXG1* gene and is associated with microcephaly and psychomotor symptoms (Kortum et al., 2011). *FOXG1* has been shown to have roles in brain development and is a key transcriptional repressor (Danesin and Houart, 2012). Verginelli *et al* have recently shown that *in vitro* and *in vivo* knockdown of *FOXG1* results in impaired growth of glioblastoma. In the same study, and in keeping with the results of the current study, high *FOXG1* mRNA expression was shown to be associated with poor survival in GBM patients (Verginelli et al., 2013). Low *FOXG1* has also been shown to be associated with the midline location of gliomas (Sturm et al., 2012).

TOGLN2 is thought to play a role in the regulation of membrane traffic to and from the trans-Golgi network (UniProt, 2015a) and has not been previously

linked to glioma. *TOLGN2* has, however, been previously studied as a prognostic marker in head and neck cancer. Fountzilas *et al* describe a model that acts as a multigene predictor of recurrence in primary laryngeal cancer that includes *TGOLN2*. In this model, low expression of *TOGLN2* was associated with a good prognosis (Fountzilas et al., 2013), in agreement with the findings of this current study.

All three gene hubs identified from the *APE1* interactome show interesting links with glioma or malignancy and both *TOGLN2* and *FOXG1* would seem good candidates for future studies investigating prognostic and/or predictive markers in GBM.

3.3.3.3 THRA, RFX4 and CD55 are identified as drivers of the NBN interaction network

THRA, *RFX4* and *CD55* were found to be the most influential genes within the *NBN* interactome, with low levels of *THRA* associated with poor overall survival. In previous studies, both *THRA* and *RFX4* have been shown to be involved in glioma development. *THRA* encodes a nuclear hormone receptor with a high affinity for the thyroid hormones and can act as both a transcriptional activator and repressor (UniProt, 2015d). *RFX4* is thought to interact directly with the X-box and activate transcription (UniProt, 2015i). *THRA* has been shown to be essential for the growth of GBM (Davis et al., 2006) and *RFX4* has been found to be overexpressed in gliomas as compared to normal brain (Matsushita et al., 2005). Interestingly, in the study by Davis *et al* they find that the growth of glioblastoma cells are thyroid hormone dependent and infer that the induction of a mild hypothyroidism may improve

survival in glioblastoma patients (Davis et al., 2006). This would appear to be contrary to the results of the current study and would suggest that further studies are required to establish the role of thyroid hormones in glioblastoma.

Of the three hubs within the *NBN* interactome, *CD55* has been the most extensively studied in cancer and has been found to be over-expressed in the tumour microenvironment (Li et al., 2001). *CD55* encodes a gene called complement decay-accelerating factor and is involved in the regulation of complement activation and the innate immune response (UniProt, 2015c). Elevated levels of CD55 have been found to be associated with oesophageal adenocarcinoma risk in patients with Barrett's oesophagus (Murao et al., 2015) and to be associated with pancreatic cancer, including the risk of lymphatic spread, vascular invasion and metastasis (He et al., 2015). Similarly, in colorectal cancer elevated levels of CD55 have been found to be associated with poor survival (Durrant et al., 2003) whereas loss of CD55 was found to be associated with a worse prognosis in breast cancer (Madjd et al., 2004). Interestingly, in the current study CD55 expression was not found to be associated with survival.

Again all three hubs within the *NBN* interactome would appear suitable targets for future studies. Further work is clearly needed to explore the association between THRA and thyroid hormone status and prognosis in GBM.

3.3.3.4 PTER, ACSL4 and STXBP6 are identified as drivers of the PMS2 interaction network

Within the *PMS2* network *PTER*, *ACSL4* and *STXBP6* were identified as the key hubs, although none of these hubs were found to be associated with survival. Mutations in both *PTER* and *ACSL4* are known to be associated with syndromes that include a neurodevelopmental component. *PTER* is a gene that has been shown to be associated with Patau syndrome and Cri-du-chat syndrome (GeneCardsSuite, 2015). Patau syndrome is usually due to trisomy of chromosome 13 and is associated with congenital abnormalities including neurodevelopmental abnormalities. Cri-du-chat syndrome is due to deletion of part of chromosome 5 and again is associated with various congenital abnormalities, including microcephaly and also cognitive impairment (Rodriguez-Caballero et al., 2010). In terms of the actual function of PTER, little is known, except that it is involved in metal ion binding and potentially epithelial cell differentiation (GeneCardsSuite, 2015). Deletion of *ACSL4*, in combination with disruption of the *DMD* (Duchenne muscular dystrophy) gene, has been shown to be associated with developmental delay, hypotonia and congenital anomalies (Bhat et al., 2006). ACSL4 is known to be involved in the activation of long-chain fatty acids and is responsible for the synthesis of both cellular lipids and their degradation by beta-oxidation (UniProt, 2015b). While both *PTER* and *ACSL4* have been shown to be associated with neurodevelopmental disorders, no association has been described with malignancy. *STXBP6* is involved in the regulation of exocytosis (UniProt, 2015f) and has previously been shown to be differentially expressed in high versus low grade gliomas (Vital et al., 2010).

The *PMS2* network clearly identifies gene hubs that are involved in neurodevelopment and potentially gliomagenesis. While none of these genes were associated with survival on univariate analysis, *STXBP6* was found to be associated with survival on multivariate analysis. This, together with the earlier observation that *STXBP6* is expressed more in high grade glioma as compared to low grade, would suggest that further investigation of *STXBP6* expression is warranted across all glioma grades.

3.3.3.5 TANC2, DCLK2 and HPRT are identified as drivers of the MGMT interaction network

Of the three key hubs within the *MGMT* network only *HPRT* was found to be associated with survival, with high expression predicting poor overall survival. Hypoxanthine phosphoribosyltransferase (HPRT) plays a critical role in the generation of purine nucleotides and mutations in this gene result in Lesch-Nyhan syndrome or gout. HPRT is a transferase that drives the conversion of inosine monophosphate and guanine to guanine monophosphate (NCBI, 2015b). Mutations at the *HPRT* locus in human cancer cell lines has been linked to defective mismatch repair (Glaab and Tindall, 1997). Similarly, a specific polymorphism in the *MGMT* gene has been to be associated with increased *HPRT* mutations in the lymphocytes of tobacco smokers (Hill et al., 2007).

Although not found to be associated with survival in the current study, both *DCLK2* and *TANC2* have previously been found to be involved in brain development or malignancy. Doublecortin-like kinase 2 (*DCLK2*) encodes a member of the protein kinase superfamily and the doublecortin family. The

DCLK2 protein product is involved in the regulation of microtubule polymerization and in addition possesses a C-terminal domain that shows considerable homology to Ca^{2+}/calmodulin-dependent protein kinase and a serine/proline-rich domain that mediates protein-protein interactions. *DCX* belongs to the same gene family as *DCLK2* and in combination these two genes have been shown to be involved in hippocampal organization. Loss of these genes results in severe epilepsy and death (NCBI, 2015a). Tetratricopeptide repeat, ankyrin repeat and coiled-coil domain-containing protein 2 (*TANC2*) has a potential role *in utero* in embryonic development but otherwise the functions of this gene are unknown (UniProt, 2015g). The *TANC2* gene is located at 17q12 in a region commonly amplified in breast cancer. A study by Mahmood *et al* identified *TANC2* as a driver gene that is over-expressed in amplified breast tumours. *TANC2* was found to play a key role in regulating proliferation and survival of clonogenic breast cancer cells, suggesting that it may be involved in carcinogenesis. *TANC2* expression was not found to be associated with relapse-free survival in breast cancer patients (Mahmood et al., 2014).

Again, the gene hubs identified would suggest key roles for these genes in gliomagenesis. The finding of high *HPRT* expression as associated with poor overall survival is interesting and novel. Mutations within the *HPRT* gene have clearly been linked with defective DNA repair, but how this translates to the survival association seen is less clear and would be an interesting area for future studies.

3.3.4 Additional DNA repair genes are identified that may play important roles in gliomagenesis

Aside from the three key hubs identified within each interactome, other genes associated with DNA repair, or replication, were present in four out of the five networks. These additional DNA repair genes identified were *NIPBL* and *SMC3* (*PTEN* network), *MBD4* (*NBN* network), *MNAT1* (*APE1* network) and *BLM* (*PMS2* network). In common with the key hubs discussed earlier, several of these genes have been previously associated with malignancy or neurodevelopmental disorders. Expression of all of these additional DNA repair genes was associated with survival with the exception of *NIPBL*.

Both NIPBL and SMC3 are involved in the cohesion complex. NIPBL is responsible for keeping sister chromatids together during cell division. Recently an additional role has been identified for this complex outside of mitosis and meiosis. NIPBL is not a core component of the cohesion complex but is involved in loading the complex onto chromatin in the S-phase of the cell cycle. SMC3 is a core component of the cohesion complex. Mutations in *NIPBL* and *SMC3*, along with mutations in other components of the cohesion complex, are associated with Cornelia de Lange syndrome. This is a syndrome that is associated with characteristic facial features, growth retardation, abnormal upper limb development and low intelligence quotient. A role for NIPBL or SMC3 in the pathogenesis of glioma has yet to be defined. Other components of the cohesion complex have, however, been linked to malignancy, such as SMC1A. Mutations of the *SMC1A* gene have been identified in colorectal cancer and expression levels of SMC1A have been shown to be associated with survival. Recently, SMC1A has been shown to be

highly expressed in human glioma as compared to normal brain and a role proposed in the pathogenesis of glioma (Ma et al., 2013).

MBD4 (also known as MED1) contains an N-terminal 5-methylcytosine binding domain that enables binding to methylated DNA and also a C-terminal catalytic domain with homology to DNA damage specific glycosylases and lyases that are found in bacteria. MBD4 has been shown to interact with MLH1 and to act as a mismatch specific DNA *N*-glycosylase. Frameshift mutations in *MBD4* have also been shown to be associated with an increased risk of malignancy. The most widely studied tumour associated with *MBD4* mutations is colorectal cancer, but mutations have also been found to be associated with gastric, pancreatic and endometrial cancer (Bellacosa, 2001).

MNAT CDK-activating kinase assembly factor1 (MNAT1) is responsible for the formation of the CDK-activating kinase (CAK) enzymatic complex, along with cyclin H and CDK7. The CAK complex is involved in the activation of cyclin associated kinases and also associates with TFIIH resulting in the activation of transcription by RNA polymerase II. Mutations in CAK complex genes have been found to be associated with cancer susceptibility. For example, several polymorphisms within the *MNAT1* gene were found to influence lung cancer risk in a dominant genetic model, while the *MNAT* 42172A/G polymorphism modified lung cancer risk in a recessive model (Li et al., 2007).

Mutations within the *BLM* gene are associated with an autosomal recessive disorder called Bloom syndrome. Individuals with Bloom syndrome have an increased risk of a range of cancers including leukaemia, lymphoma and

epithelial derived tumours such as breast cancer. In addition, heterozygous carriers of *BLM* mutations also appear to be at an increased risk of malignancy, including an increased risk of colorectal cancer (Cleary et al., 2003). Also, polymorphisms in the *BLM* gene have been associated with an increased risk of breast cancer and tumour grade (Sassi et al., 2013). In a large study of breast cancer patients, high *BLM* mRNA expression was associated with poor prognostic features and poor breast cancer specific survival. Similarly, at the protein level, high cytoplasmic and low nuclear BLM expression was associated with aggressive tumour features and poor breast cancer specific survival (Arora et al., 2015). This is in keeping with the results of the current study, where high *BLM* gene expression levels are associated with poor overall survival in GBM.

Clearly, this component of the study has identified additional DNA repair genes that may have potential prognostic and/or predictive significance in GBM. Some of these DNA repair genes have been more widely studied in malignancy, such as *BLM*, while others are more novel targets for study such as *SMC3*. Validation of these findings in further datasets is required, followed by investigation of expression at the protein level. If the prognostic and/or predictive significance of these genes is confirmed, then cell line studies may be warranted. If, for example, high expression of one of these targets appears to be associated with resistance to chemotherapy then gene knock-down studies could be performed to confirm this *in vitro*. If this resulted in increased chemo-sensitivity then it would suggest that these genes may represent potential drug targets.

3.3.5 ANN inference can be used to identify recurrent themes and generate new hypotheses

Artificial neural network analysis techniques can generate vast amounts of information and a key to interpreting this data is to have a focused strategy. In this study, ANN methodology has been used to follow up the five most important DNA repair genes identified from the survival analyses. The use of ANN analysis has enabled key driver genes to be identified for each of the original five DNA repair genes of interest, but has also been used to identify more novel DNA repair associated genes. This methodology has previously been published in the literature (Lancashire et al., 2010) and is further validated by the identification of genes known to be associated with either malignancy or neural development.

Artificial neural network analysis enables the identification of common themes, for example pathways that may play key roles in GBM, and also allows the identification of individual genes that may warrant further investigation. For example, *FOXG1* was identified as a key hub within the *APE1* interactome and high expression was found to be associated with poor survival. In addition, on multivariate analysis *FOXG1* was found to be the most significant independent predictor of survival. This led on to a BMedSci project, undertaken by Miss Emily Robertson, investigating the prognostic significance of the FOX transcription factor family in GBM. In this follow-on study, in addition to *FOXG1*, expression of *FOXA2*, *FOXN2* and *FOXN3* were also found to be significantly associated with survival (Robertson et al., 2015). Future studies could take this forward to look at protein expression of these FOX genes and then lead in to cell line studies looking at the potential therapeutic benefit of

manipulating expression of these genes. A similar approach could be undertaken with the genes representing the other key hubs, particularly those found to be associated with survival, and also for the additional DNA repair genes identified from the network analysis. Another potential future study could be to undertake a similar body of work but looking at low grade glioma or paediatric glioma. The TCGA group is currently collecting samples for low grade glioma patients; although at the time of writing the number of samples available was too low to undertake this type of project.

3.3.6 Summary

In this chapter a panel of twelve key DNA repair genes are identified that are associated with survival in GBM. The ability to develop a DNA repair gene prognostic index is demonstrated and the five most important DNA repair genes associated with survival in GBM have been identified. Artificial neural network analysis techniques have been used to explore the genes that interact with these five key DNA repair genes. This has led to the identification of further genes, genome-wide, that may be involved in the development of gliomas or be associated with outcome, including the identification of additional genes involved in DNA repair.

This chapter has purely focused on gene expression data and therefore only represents a small proportion of the genetic and epigenetic changes seen in cancer. Additional levels of regulation are often seen other than transcriptional regulation, such as microRNAs, protein phosphorylation and ubiquitination. In order to attempt to overcome some of these factors APE1, NBN, PMS2 and

PTEN protein expression have been investigated in a cohort of high grade glioma patients in the following chapter.

CHAPTER 4

Exploring APE1, NBN, PMS2 and PTEN protein expression in high grade glioma

4 Results

4.1 Introduction

In the previous chapter five key DNA repair genes were identified whose expression, at the mRNA level, was found to be associated with survival. Low mRNA expression levels of *APE1*, *NBN* and *PTEN* were found to be associated with poor overall survival in glioblastoma. Conversely, high mRNA expression levels of *MGMT* and *PMS2* were found to be associated with poor overall survival. As discussed previously, MGMT expression has been extensively studied in glioblastoma (Costa et al., 2010, Hegi et al., 2005, Rivera et al., 2010, Zhang et al., 2013). This chapter will therefore focus on the assessment of APE1, NBN, PMS2 and PTEN protein expression in a cohort of high grade glioma patients to determine the pattern of expression and whether expression of these proteins is associated with survival.

4.1.1 An overview of the DNA repair proteins studied

The four DNA repair proteins selected for study are involved in different DNA repair pathways, which include BER, MMR and DSB repair.

4.1.1.1 APE1

APE1 is a key component of the BER pathway and is involved in the processing of AP sites, which are generated after the removal of a damaged

base by a DNA glycosylase. The phosphodiester bond immediately 5' to the AP site is hydrolysed by APE1 and results in a 3'-hydroxyl group and an abasic 5'-deoxyribose-phosphate (dRP) residue flanking the nucleotide gap (Abbotts and Madhusudan, 2010). Repair now proceeds by either the short patch (SP-BER) or long patch (LP-BER) pathway.

In addition, APE1 has also been found to have a redox function and to be involved in the regulation of various proteins, including some involved in DNA repair. The DNA repair activity of APE1 is controlled by the C-terminal domain of the protein, whereas the redox function is under control of the N-terminal region. While the BER repair function of APE1 is highly conserved from E. coli exonuclease III to human APE1, only mammalian APE1 has both endonuclease and redox activity (Georgiadis et al., 2008). Additional roles of APE1 have also been described including involvement in RNA quality control (Vascotto et al., 2009) and parathyroid gene expression (Okazaki et al., 1994).

The redox domain and redox function of APE1 is less well understood than the repair domain. APE1 is involved in the redox regulation of a variety of transcription factors and thereby controls gene expression. Transcription factors regulated by APE1 include ubiquitous proteins such as HIF-1α (Ema et al., 1999, Lando et al., 2000), p53 (Ueno et al., 1999), NFκβ (Hirota et al., 1999), CREB, Egr-1 (Huang and Adamson, 1993) and AP-1 (Xanthoudakis et al., 1992). APE1 also regulates the redox status of tissue specific transcription factors, including Pax-5 (Tell et al., 2000), Pax-8 (Cao et al., 2002), PEBP-2 (Akamatsu et al., 1997) and TTF-1. These transcription factors are involved in

pathways responsible for tumour promotion, progression and angiogenesis (Kelley et al., 2012).

4.1.1.2 NBN

NBN is a core component of the MRN complex which consists of MRE11-NBN-RAD50. This is involved in the repair of DNA double-strand breaks and promotes the resection of DNA ends in combination with CtIP and BRCA1. In HR, the MRN complex is the major sensor of DNA damage (Moreno-Herrero et al., 2005, Sartori et al., 2007). NBN serves to activate downstream kinases involved in the damage response including ATM (Lamarche et al., 2010). As discussed in Section 3.3.1, mutations within the *NBN* gene result in a cancer-predisposition syndrome known as Nijmengen breakage syndrome (Digweed and Sperling, 2004).

4.1.1.3 PMS2

PMS2 is a key component of the MMR pathway. PMS2 forms a complex with MLH1 which is known as the MutLα complex. MutLα co-ordinates the relationship between the MutS complex (damage recognition complex) and downstream mediators of the MMR pathway (Peltomäki, 2001). Deficiency of MMR repair is associated with an autosomal hereditary cancer syndrome known as Lynch Syndrome or Hereditary Non Polyposis Colorectal Cancer (HNPCC). This syndrome is associated with mutations in MLH1, MLH2, MSH6 and PMS2 and was first described in 1895, with the genetic basis first documented by Peltomaki *et al* in 1993 (Peltomaki et al., 1993). Patients with Lynch Syndrome are predisposed to early onset colorectal cancer and also

endometrial, ovarian, gastric and small bowel tumours. An increased risk of transitional cell carcinoma of the renal pelvis and ureter has also been demonstrated in this patient group (reviewed in (Lynch and Smyrk, 1996)).

4.1.1.4 PTEN

PTEN is a tumour suppressor gene and a negative regulator of the PI3K signalling pathway (Liu et al., 2009). In addition, a potential role for PTEN in DNA DSB repair has been proposed via an association between PTEN and RAD51 expression (Shen et al., 2007). PTEN has also been shown to influence NER by enhancing transcription of XPC (Ming et al., 2011). Furthermore, PTEN can indirectly influence DNA repair via cell cycle effects. PTEN induced suppression of cyclin D1 has been shown to result in stalling of the cell cycle at G1/S, enabling time for repair of DNA damage (Radu et al., 2003). The functions of PTEN are discussed in more detail in Section 1.8.

4.1.2 DNA repair and DDR protein expression in glioma

As illustrated above, all four DNA repair proteins play key roles in different DNA repair pathways and while all have been studied to an extent in malignancy, less is known about their role in glioma. Few studies have specifically investigated the expression of the DNA damage response proteins in glioblastoma. Seol *et al* have studied the expression of a panel of DDR response proteins in a retrospective cohort of 69 glioblastoma samples by immunohistochemistry. The 10 DDR-related proteins studied were ATM, Chk1, Chk2, TopBP1, Rad17, p53, NBN, MDC1, γH2AX and RPA1. Expression of ATM, Chk1, Chk2, TopBP1, p53 and NBN were significantly

higher in glioblastoma tissue as compared with normal brain. On Kaplan Meier survival analysis, high ATM expression was found to be associated with both better progression free and overall survival. Higher p53 expression was also associated with better progression free survival. In addition, higher ATM expression was an independent predictor of both overall and progression free survival. p53 was the only other DDR-related protein to be an independent predictor of progression free survival. Age >70 years, multiple lesions, extent of resection and TopBP1, ATM and p53 expression were all independent predictors of overall survival (Seol et al., 2011).

4.1.3 Immunohistochemistry is a widely used technique to evaluate protein expression

In order to study the expression of these four DNA repair proteins of interest an appropriate, simple and reproducible methodology was required. Immunohistochemistry is a technique that enables the visualisation of specific cellular components within cells by the interaction of a specific antibody with a target antigen. This enables the distribution of the antigen of interest to be studied within the tissue, including the location within individual cells (ThermoScientific, 2015). Immunohistochemistry is a commonly reported technique and has been used extensively within the Nottingham Translational DNA Repair Group to stain for a wide range of DNA repair proteins in large cohorts of various cancer types (Abdel-Fatah et al., 2014a, Abdel-Fatah et al., 2015, Arora et al., 2015).

While the principle of IHC has been known about since the 1930's, the first IHC study was reported by Coons *et al* in 1942 (Coons et al., 1942). It is a

process that involves a series of ordered steps that enable the specific identification of the target antigen. Firstly, the tissue should be suitably preserved; a commonly used technique is for the sample to be fixed in formalin and then embedded in paraffin, as was the case with the samples used in this study. In this scenario the paraffin must be completely removed prior to staining (ThermoScientific, 2015). In the current study this was achieved by a series of sequential washes in xylene, ethanol and water using a Leica autostainer that serves to both remove the wax and rehydrate the tissue.

During tissue fixation, cross-links are created that can prevent the antibody binding to the antigen. Antigen retrieval, also known as antigen unmasking or epitope retrieval, is a technique used to reverse these cross-links. Antigen retrieval can be achieved by a method known as heat-induced epitope retrieval (HIER) or by enzymatic antigen retrieval techniques. HIER is widely used and requires an appropriate buffer and heating device. The two most widely used buffers are 10mM citrate buffer (pH 6) and 1mM EDTA buffer (pH 8). Heating of the slides in the buffer solution can either be performed in a microwave, pressure cooker or, less commonly, a water bath (ThermoScientific, 2015). In this study HIER was employed using citrate buffer and heating in a Whirpool JT359 Jet Chef 1000W microwave.

The Novolink™ Polymer Detection system (as used in this study) uses polymerisation technology to produce polymeric HRP-linker antibody conjugates and involves a series of sequential steps interspersed with washing steps. After epitope retrieval, samples are prepared prior to the administration of primary antibody by exposure to blocking solutions. Firstly, Peroxidase

block is applied to neutralise endogenous peroxidase activity and secondly Protein Block is used to reduce non-specific binding of the primary antibody and polymer. Both of these steps serve to reduce background staining (LeicaBiosystems, 2016).

Tissue sections can now be exposed to the primary antibody diluted to a pre-determined concentration in antibody diluent. This both stabilises the antibody and assists even distribution across the sample, again reducing non-specific binding (LeicaBiosystems, 2016). Selection of the primary antibody is important and various quality checks can be applied, including sourcing an antibody that has been previously widely used. In this current study antibody specificity has been determined by Western blot analysis, whereby the presence of a single band at the correct molecular weight signifies a lack of non-specific binding.

After exposure to the primary antibody Post Primary solution is applied, this is a Rabbit anti Mouse IgG and is used to detect mouse antibodies. Novolink™ Polymer is used to detect rabbit immunoglobulins including the post primary and primary rabbit antibody bound to tissues. Sections are then incubated with DAB; with the peroxidase reaction produced resulting in the formation of a brown precipitate at the reaction site. Counterstaining is achieved with Haematoxylin, prior to the application of a coverslip and viewing under a light microscope (LeicaBiosystems, 2016).

Ensuring that the results achieved by IHC staining are reproducible is important. This can be achieved by the inclusion of both positive and negative control sections. In the current study a negative control section was included

that consisted of a section stained with the omission of primary antibody. This control serves to identify any background staining produced by reagents, other than the primary antibody, used in the process. Two positive control sections were included; one consisting of a tissue known to stain positive for the antigen of choice (antibody specific control) and one for beta 2 microglobulin, a commonly used housekeeper protein, that is known to stain positive in most tissues. The antibody specific positive control provided evidence that the antibody was staining as expected, and by using the same control section enabled the staining to be compared between batches to ensure consistency. Staining for beta 2 microglobulin provided evidence that the run of IHC had been successful.

4.1.4 Study rationale

Following on from the large scale DNA repair gene expression analysis performed in Chapter 3, in this chapter, expression of four of the most significant DNA repair genes that predict survival in glioblastoma will be explored at the protein level. Sixty-one samples from high grade glioma patients will be stained for APE1, NBN, PMS2 and PTEN protein expression by IHC. The aim of this chapter is to demonstrate whether these four DNA repair markers retain prognostic significance when expression is assessed at the protein level.

4.2 Results

4.2.1 APE1, NBN, PMS2 and PTEN antibody specificity is confirmed by western blot analysis

Western blot analysis was used to determine the specificity of the antibodies to APE1, NBN, PMS2 and PTEN. As shown in Figure 4.1 a single band is observed for all four glioblastoma cell line lines and all four antibodies. For APE1 a single band is seen at 37 kDa and for NBN at 85 kDa. Single bands are demonstrated for PTEN at 55 kDa and PMS2 at 110 kDa.

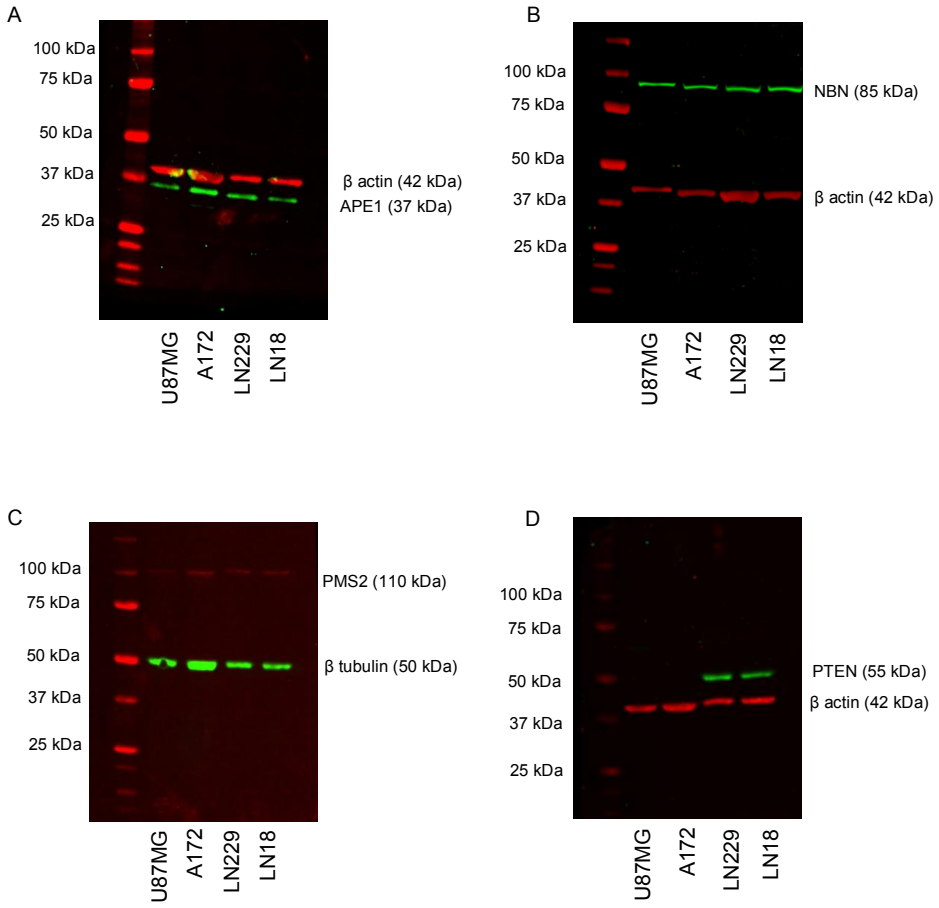


Figure 4.1. Western blots for APE1, NBN, PMS2 and PTEN.

Western blots confirm a single band for all four antibodies at the correct molecular weight. A. APE1 (single band at 37 kDa). B. NBN (single band at 85 kDa). C. PMS2 (single band at 110 kDa). D. PTEN (single band at 55 kDa).

4.2.2 Control sections ensure the quality of staining in maintained

As stated earlier, a negative control was included for each antibody, whereby the section was exposed to antibody diluent only and not primary antibody. An example is shown in (Figure 4.2) demonstrating an absence of staining. Positive control sections were also included in each run of IHC. A beta 2 microglobulin stained section was also included to confirm that the staining technique was successful (Figure 4.2).

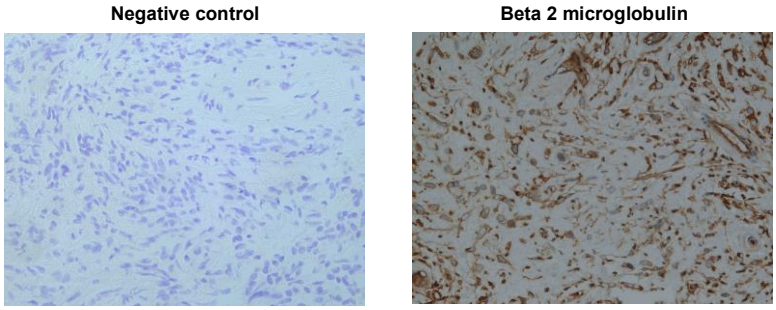


Figure 4.2. Negative and positive control sections.

Example of tissue stained with the omission of primary antibody, in this case a glioma full-face section (left). In addition, a breast cancer TMA was stained with beta 2 microglobulin (right) to confirm that the IHC technique for the batch of slides was successful.

In addition, a section of tissue known to stain positive for the antibody under investigation was included in each IHC run. This was used as an antibody-specific positive control. A list of positive controls used for each antibody is included in Section 2.3.4.

4.2.3 Determination of the optimum concentration of antibody for IHC analysis

For each antibody used, the optimal concentration for IHC analysis was determined on a tissue known to express the protein of interest. The tissues used for each antibody are listed in the Section 2.3.3.

Figure 4.3 shows the range of dilutions used for APE1 antibody concentration optimisation, staining glioma full-face sections taken from a single sample block. 1:300 was deemed the optimal concentration of APE1 antibody for IHC staining.

The optimum concentration of NBN antibody for IHC staining was determined using glioma full-face sections. A concentration of 1:150 was deemed to be the optimal concentration. The range of concentrations tested for the NBN antibody is shown in Figure 4.4.

PMS2 antibody optimisation was initially performed on colorectal cancer full-face sections, as colorectal cancer has previously been identified as a tissue that expresses high levels of PMS2. A range of concentrations (1:100 to 1:500) of PMS2 antibody were tested, with the best staining demonstrated between the concentrations of 1:100 to 1:200 (Figure 4.5). Full-face glioma sections were then stained at 1:100, 1:150 and 1:200 and 1:200 was deemed the optimal concentration for subsequent IHC analyses (Figure 4.6).

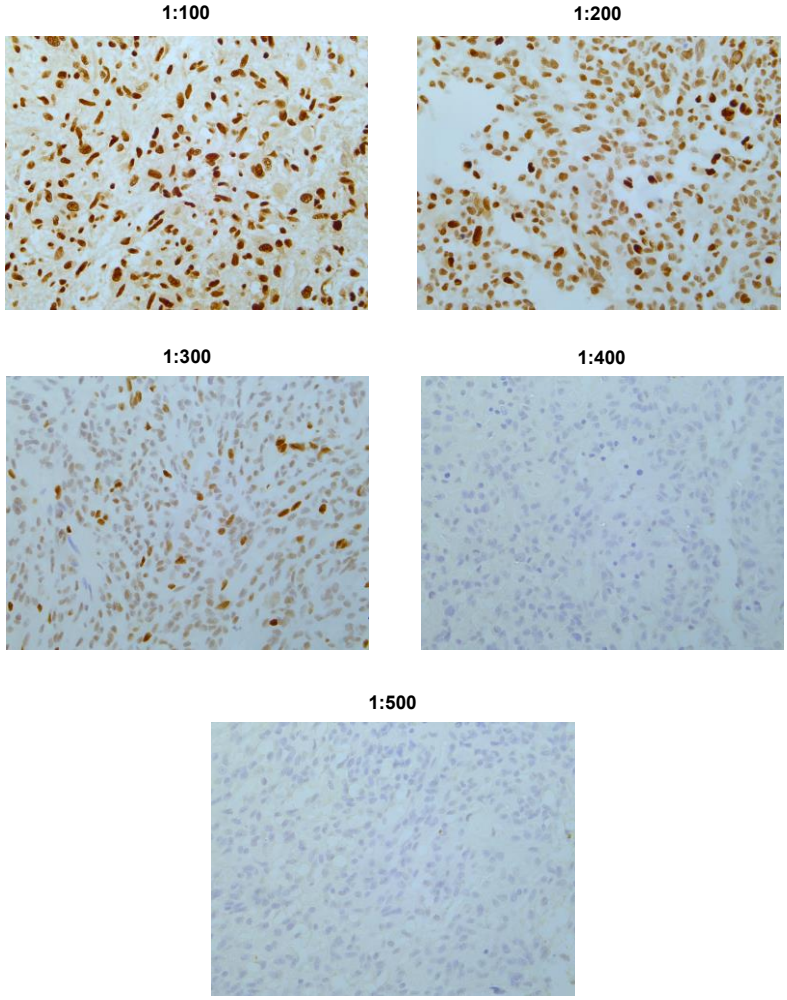


Figure 4.3. Range of APE1 antibody concentrations used for optimisation.

Optimisation of APE1 antibody concentration for IHC analysis. Full-face glioma sections were stained at APE1 antibody concentrations of 1:100 to 1:500. The optimal concentration for staining glioma tissue was found to be 1:300 dilution; at this concentration a range of staining intensities was demonstrated, including both positive and negative cells.

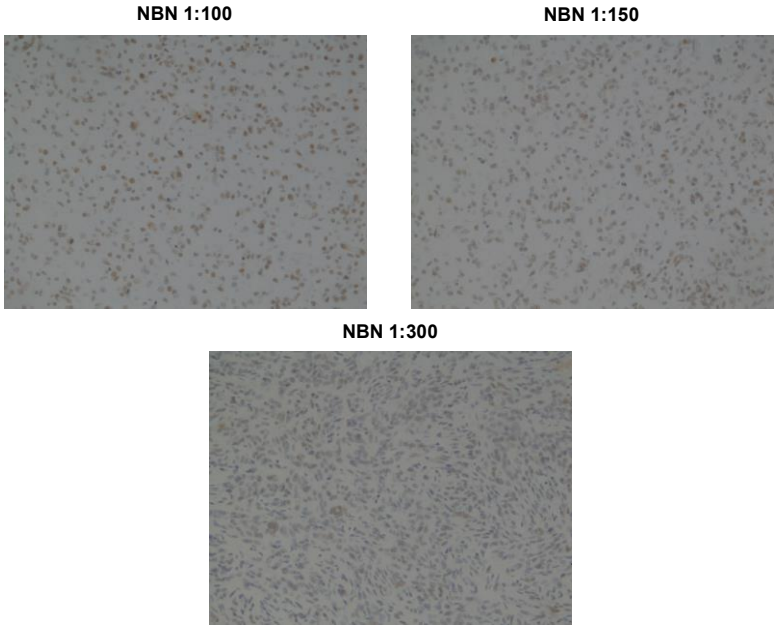


Figure 4.4. Range of NBN antibody concentrations used for optimisation.

Optimisation of NBN antibody concentration for IHC analysis. Full-face glioma sections were stained at NBN antibody concentrations of 1:100 to 1:300. The optimal concentration for staining glioma tissue was found to be 1:150 dilution; at this concentration a range of staining intensities was demonstrated, including both positive and negative cells.

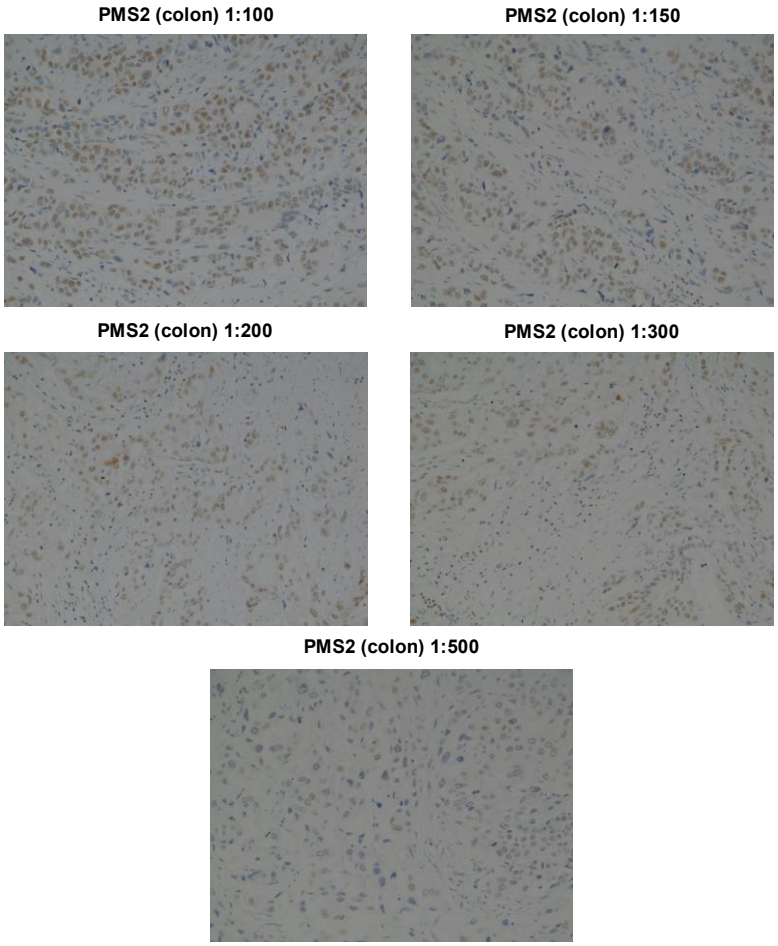


Figure 4.5. Range of PMS2 antibody concentrations used for optimisation on colon tissue.

Full-face colon cancer sections were stained at PMS2 antibody concentrations of 1:100 to 1:500. The optimal concentration for staining colon cancer tissue was found to be 1:200 dilution; at this concentration a range of staining intensities was demonstrated, including both positive and negative cells.

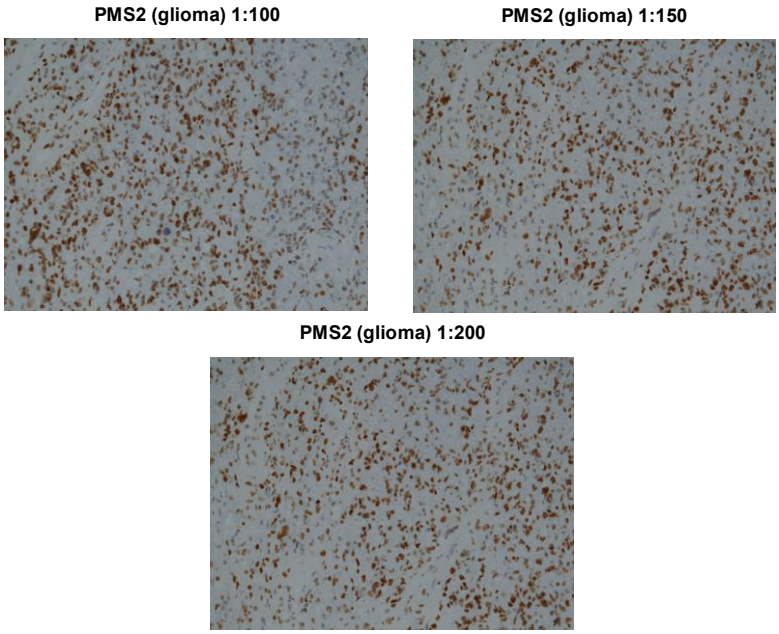


Figure 4.6. Range of PMS2 antibody concentrations used for optimisation on glioma tissue.

Full-face glioma sections were stained at PMS2 antibody concentrations of 1:100 to 1:200. The optimal concentration for staining glioma tissue was confirmed to be 1:200 dilution; at this concentration a range of staining intensities was demonstrated, including both positive and negative cells.

PTEN expression is commonly lost in human tumours, including in glioma (Li et al., 1997, Guldberg et al., 1997, Stemke-Hale et al., 2008, Cairns et al., 1997, Keniry and Parsons, 2008), and therefore TMAs were used for antibody optimisation. The optimal concentration of PTEN antibody for IHC analysis (1:150) was therefore determined using both gastric and ovarian TMAs (Figure 4.7) as these were readily available. For further validation that this concentration was suitable for glioma tissue, a glioma full-face section was also stained at a concentration of 1:150.

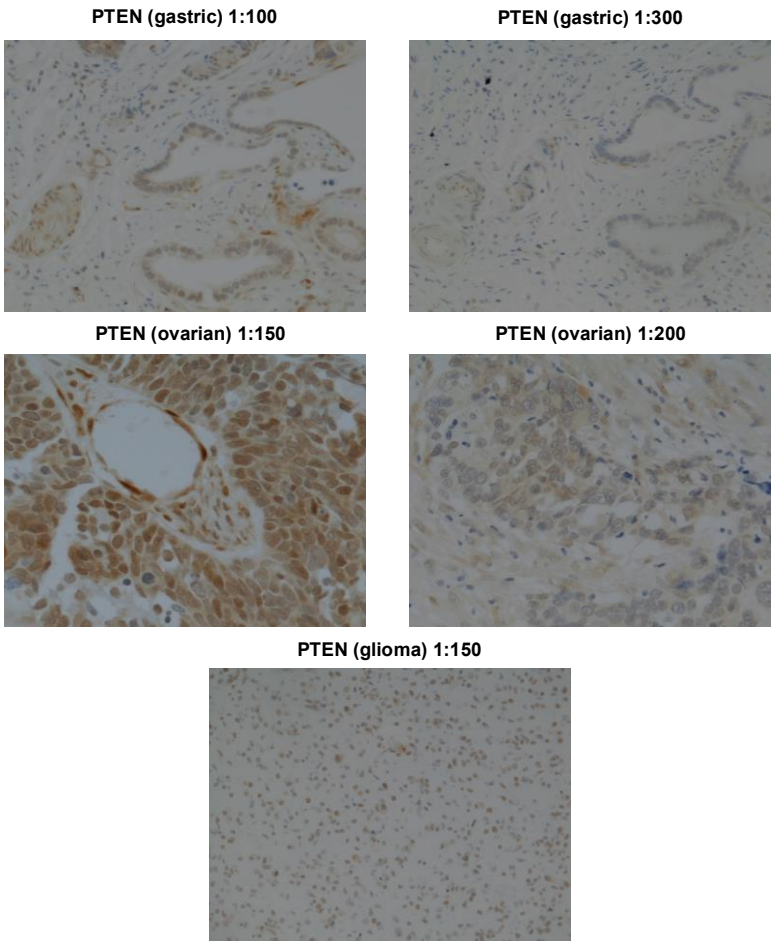


Figure 4.7. Range of PTEN antibody concentrations used for optimisation

Gastric and ovarian cancer TMAs were initially used for optimisation of PMS2 antibody concentration and stained at a range of concentrations from 1:100 to 1:300. Further staining of a glioma full-face section confirmed 1:150 to be the optimal concentration; at this concentration a range of staining intensities was demonstrated, including both positive and negative cells.

4.2.4 Older age and higher grade are significantly associated with poor overall survival in high grade glioma

As discussed previously, older age (MRC, 1990) and higher grade are known prognostic factors in glioma (Behin et al., 2003). To determine whether the cohort was representative of high grade glioma patients in general the association between both age and grade (3 versus 4) was explored in the cohort. On Kaplan Meier survival analysis both greater age (>58 years) and grade 4 tumour status were associated with poor overall survival (Figure 4.8).

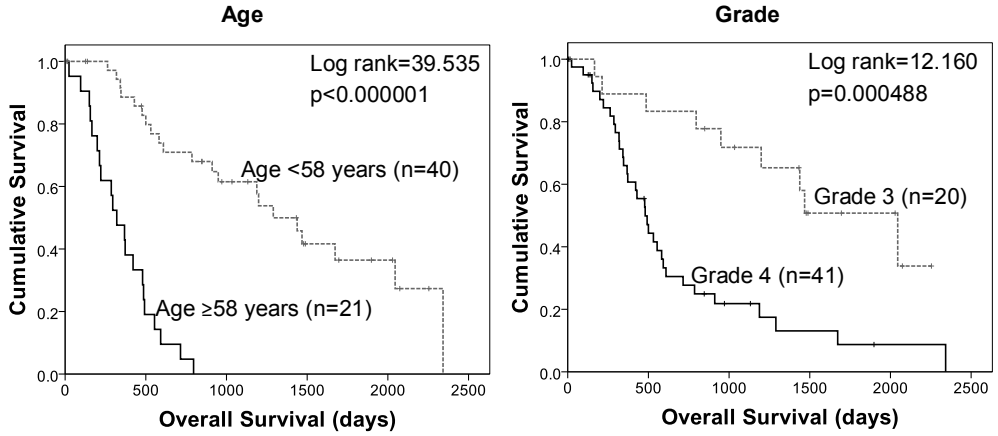


Figure 4.8. Kaplan Meier survival curves for age and grade in high grade glioma patients.

Kaplan Meier survival curves for age (left) and grade (right). Both age ≥58 years ($p<0.000001$) and grade 4 status ($p=0.000488$) were significantly associated with poor overall survival in high grade glioma, as determined by the Log Rank test.

4.2.5 APE1 protein is widely expressed in high grade glioma

Median APE1 score in high grade glioma patients in the cohort was 170 (range 50-250). Nuclear staining was demonstrated in all patients. Examples of tumours with low (H-score <170) and high APE1 expression are shown in Figure 4.9.

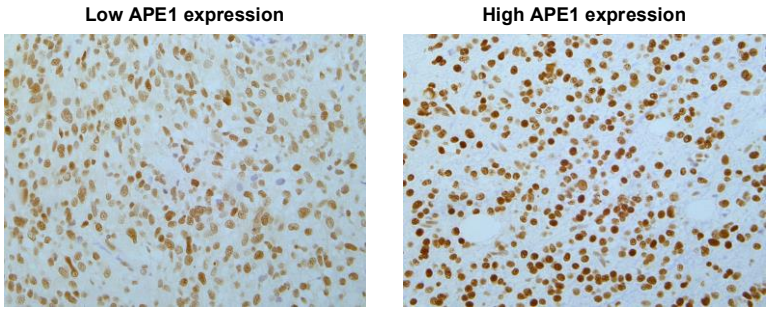


Figure 4.9. Examples of APE1 expression in high grade glioma sections.

An example of a high grade glioma with low APE1 expression (H score <170) is shown on the left, while an example of high APE1 expression (H score ≥170) is shown on the right.

4.2.6 APE1 protein expression is associated with nuclear PTEN and PMS2 expression

On Chi-squared analysis, APE1 protein expression is not associated with age (p=0.217), gender (p=0.891), a previous diagnosis of glioma (p=0.961) or grade (3 or 4) (p=0.101) (Table 4.1). Low APE1 expression is however, significantly associated with low nuclear PTEN expression (p=0.037) and low PMS2 expression (p=0.01).

Table 4.1. Clinicopathological associations of APE1 protein expression in high grade glioma.

	APE1 expression		
Variable	Low (N=25) Number (%)	High (N=35) Number (%)	X^2 **Adjusted p value**
	Demographic details		
Age			
<59 years	14 (56.0)	25 (71.4)	0.217
≥59 years	11 (44.0)	10 (28.6)	
Gender			
Male	16 (64.0)	23 (65.7)	0.891
Female	9 (36.0)	12 (34.3)	
Previous glioma diagnosis			
No	18 (72.0)	25 (71.4)	0.961
Yes	7 (28.0)	10 (28.6)	
Grade			
3	5 (20.0)	14 (40.0)	0.101
4	20 (80.0)	21 (60.0)	
	Treatment given		
Radiotherapy			
No	0 (0.0)	2 (6.2)	0.666*
Yes	21 (100.0)	30 (93.8)	
Chemotherapy			
No	7 (33.3)	9 (27.3)	0.634
Yes	14 (66.7)	24 (72.7)	
	DNA repair gene expression		
NBN expression			
Low	17 (68.0)	20 (57.1)	0.394
High	8 (32.0)	15 (42.9)	
Nuclear PTEN expression			
Low	24 (96.0)	25 (71.4)	**0.037***
High	1 (4.0)	10 (28.6)	
Cytoplasmic PTEN expression			
Low	13 (52.0)	14 (40.0)	0.357
High	12 (48.0)	21 (60.0)	
PMS2 expression			
Low	10 (40.0)	4 (11.4)	**0.010**
High	15 (60.0)	31 (88.6)	

* Yates' continuity correction. Significant p values (≤0.05) are shown in bold.

4.2.7 Variable NBN protein expression is present in high grade glioma

Nuclear NBN expression was demonstrated in all high grade glioma samples. Two samples also demonstrated cytoplasmic NBN expression. Median H-score for nuclear NBN expression was 70 (range 0-250). Examples of low (H-score ≤85) and high NBN (H-score >85) staining observed in the cohort are shown in Figure 4.10.

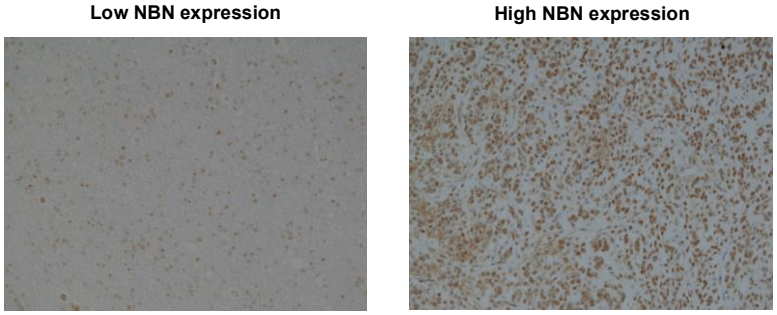


Figure 4.10. Examples of NBN expression in high grade glioma sections.

An example of a high grade glioma with low NBN expression (H score ≤85) is shown on the left, while an example of high NBN expression (H score >85) is shown on the right.

4.2.8 NBN protein expression is not associated with clinicopathological features

NBN protein expression is not associated with age (p=0.885), gender (p=0.720), a previous diagnosis of glioma (p=0.270) or grade (3 or 4) (p=0.080). Similarly NBN expression is not associated with APE1 (p=0.394), nuclear (p=0.139) or cytoplasmic (p=0.948) PTEN or PMS2 (p=0.996) expression (Table 4.2).

Table 4.2. Clinicopathological associations of NBN protein expression in high grade glioma.

	NBN expression		
Variable	Low (N=37) Number (%)	High (N=24) Number (%)	X^2 **Adjusted p value**
	Demographic details		
Age			
<59 years	24 (64.9)	16 (66.7)	0.885
≥59 years	13 (35.1)	8 (33.3)	
Gender			
Male	23 (62.2)	16 (66.7)	0.720
Female	14 (37.8)	8 (33.3)	
Previous glioma diagnosis			
No	28 (75.7)	15 (62.5)	0.270
Yes	9 (24.3)	9 (37.5)	
Grade			
3	9 (24.3)	11 (45.8)	0.080
4	28 (75.7)	13 (54.2)	
	Treatment given		
Radiotherapy			
No	1 (3.1)	1 (4.5)	1.000*
Yes	31 (96.9)	21 (95.5)	
Chemotherapy			
No	9 (27.3)	7 (31.8)	0.716
Yes	24 (72.7)	15 (68.2)	
	DNA repair gene expression		
APE1 expression			
Low	17 (45.9)	8 (34.8)	0.394
High	20 (54.1)	15 (65.2)	
Nuclear PTEN expression			
Low	33 (89.2)	17 (70.8)	0.139*
High	4 (10.8)	7 (29.2)	
Cytoplasmic PTEN expression			
Low	17 (45.9)	10 (41.7)	0.948
High	20 (54.1)	14 (58.3)	
PMS2 expression			
Low	9 (24.3)	5 (20.8)	0.996*
High	28 (75.7)	19 (79.2)	

* Yates' continuity correction. Significant p values (≤0.05) are shown in bold.

4.2.9 PMS2 is widely expressed in high grade glioma

Nuclear PMS2 expression was observed in all high grade gliomas studied. A broad range of H scores was demonstrated, ranging from 20-250, with a median H-score of 155. Examples of low (H-score ≤120) and high PMS2 (H-score >120) expression are shown in Figure 4.11.

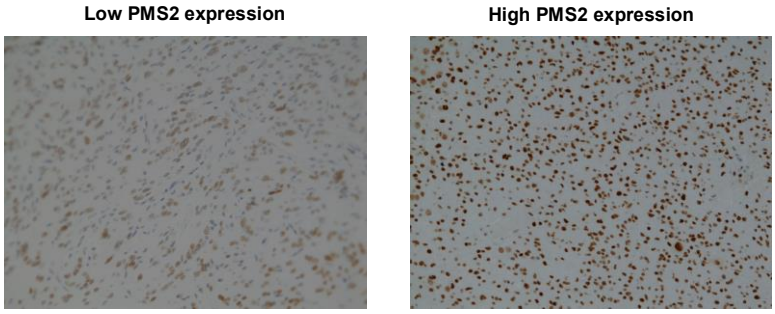


Figure 4.11. Examples of PMS2 expression in high grade glioma sections.

An example of a high grade glioma with low PMS2 expression (H score ≤120) is shown on the left, while an example of high PMS2 expression (H score >120) is shown on the right.

4.2.10 PMS2 expression was not associated with clinicopathological features, except APE1 expression

PMS2 expression was not associated with age (p=0.663), gender (p=0.728), a previous history of glioma (p=0.276) or grade (3 or 4) (p=0.480). Low PMS2 expression was associated with low APE1 expression (p=0.01), as stated previously. PMS2 expression was not associated with nuclear (p=0.417) or cytoplasmic (p=0.269) PTEN expression or NBN expression (p=0.996) (Table 4.3).

Table 4.3. Clinicopathological associations of PMS2 protein expression in high grade glioma.

	PMS2 expression		
Variable	Low (N=14) Number (%)	High (N=47) Number (%)	X^2 **Adjusted p value**
	Demographic details		
Age			
<59 years	8 (57.1)	32 (68.1)	0.663*
≥59 years	6 (42.9)	15 (31.9)	
Gender			
Male	10 (71.4)	29 (61.7)	0.728*
Female	4 (28.6)	18 (38.3)	
Previous glioma diagnosis			
No	12 (85.7)	31 (66.0)	0.276*
Yes	2 (14.3)	16 (34.0)	
Grade			
3	3 (21.4)	17 (36.2)	0.480*
4	11 (78.6)	30 (63.8)	
	Treatment given		
Radiotherapy			
No	0 (0.0)	2 (4.7)	1.00*
Yes	11 (100.0)	41 (95.3)	
Chemotherapy			
No	3 (27.3)	13 (29.5)	1.00*
Yes	8 (72.7)	31 (70.5)	
	DNA repair gene expression		
APE1 expression			
Low	10 (71.4)	15 (32.6)	**0.010**
High	4 (28.6)	31 (67.4)	
NBN expression			
Low	9 (64.3)	28 (59.6)	0.996*
High	5 (35.7)	19 (18.5)	
Nuclear PTEN expression			
Low	13 (92.9)	37 (38.5)	0.417*
High	1 (7.1)	10 (21.3)	
Cytoplasmic PTEN expression			
Low	8 (57.1)	19 (40.4)	0.269
High	6 (42.9)	28 (59.6)	

* Yates' continuity correction. Significant p values (≤0.05) are shown in bold.

4.2.11 Both nuclear and cytoplasmic PTEN expression patterns are present in high grade glioma

Both nuclear and cytoplasmic expression of PTEN were seen in the cohort. Median H-score for nuclear PTEN expression was 10 (range 0-105). H-scores were higher for cytoplasmic PTEN expression, with a median of 50 (range 0-200). Examples of both PTEN nuclear and cytoplasmic expression are shown in Figure 4.12.

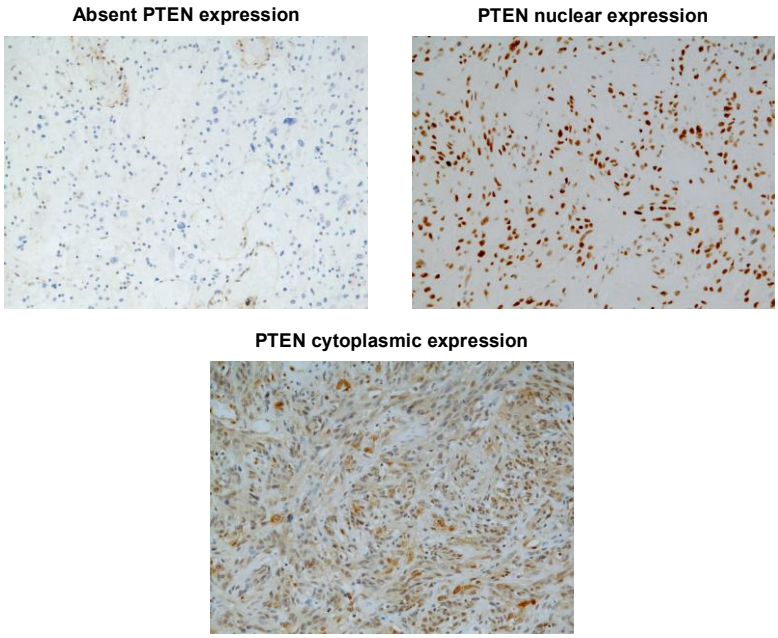


Figure 4.12. Examples of PTEN expression in high grade glioma.

An example of a high grade glioma with low PTEN expression is shown in the top left, while an example of nuclear PTEN expression is shown in the top right. Cytoplasmic PTEN expression is shown at the bottom.

Fifty (82.0%) high grade glioma patients demonstrated low nuclear PTEN expression (H-score ≤50), while low cytoplasmic PTEN expression (H-score <20) was observed in 27 (44.3%) patients. Thirteen patients had negative

PTEN expression (H-score=0 for both nuclear and cytoplasmic expression). No association was seen between nuclear and cytoplasmic PTEN expression (p=0.112).

4.2.12 PTEN expression is associated with grade

Nuclear PTEN expression was not associated with age (p=0.367), gender (p=0.746) or a previous diagnosis of glioma (p=1.000). Likewise, cytoplasmic PTEN expression was not associated with age (p=0.355), gender (p=0.498) or a past history of glioma (p=0.094). Low nuclear PTEN expression was, however, associated with grade 4 tumours (p=0.006) (Table 4.4). Similarly, low cytoplasmic PTEN expression was associated with grade 4 glioma (p=0.034) (Table 4.5).

As mentioned earlier, low nuclear PTEN expression was associated with low APE1 expression (p=0.037), but not with NBN (p=0.139) or PMS2 (p=0.417) expression. Cytoplasmic PTEN expression was not associated with APE1 (p=0.357), NBN (p=0.742) or PMS2 (p=0.269) expression.

Table 4.4. Clinicopathological associations of nuclear PTEN protein expression in high grade glioma.

	Nuclear PTEN expression		
Variable	Low (N=50) Number (%)	High (N=11) Number (%)	X^2 **Adjusted p value**
	Demographic details		
Age			
<59 years	31 (62.0)	9 (81.8)	0.367*
≥59 years	19 (38.0)	2 (18.2)	
Gender			
Male	31 (62.0)	8 (72.7)	0.746*
Female	19 (38.0)	3 (27.3)	
Previous glioma diagnosis			
No	35 (70.0)	8 (72.7)	1.000*
Yes	15 (30.0)	3 (27.3)	
Grade			
3	12 (24.0)	8 (72.7)	**0.006**
4	38 (76.0)	3 (27.3)	
	Treatment given		
Radiotherapy			
No	1 (2.3)	1 (10.0)	0.810*
Yes	43 (97.7)	9 (90.0)	
Chemotherapy			
No	14 (31.1)	2 (20.0)	0.753
Yes	31 (68.9)	8 (80.0)	
	DNA repair gene expression		
APE1 expression			
Low	24 (49.0)	1 (9.1)	**0.037***
High	25 (51.0)	10 (90.9)	
NBN expression			
Low	33 (66.0)	4 (36.4)	0.139*
High	17 (34.0)	7 (63.6)	
Cytoplasmic PTEN expression			
Low	25 (50.0)	2 (18.2)	0.112*
High	25 (50.0)	9 (81.8)	
PMS2 expression			
Low	13 (26.0)	1 (9.1)	0.417*
High	37 (74.0)	10 (90.9)	

* Yates' continuity correction. Significant p values (≤0.05) are shown in bold.

Table 4.5. Clinicopathological associations of cytoplasmic PTEN protein expression in high grade glioma.

	Cytoplasmic PTEN expression		
Variable	Low (N=27) Number (%)	High (N=34) Number (%)	X^2 **Adjusted p value**
	Demographic details		
Age			
<59 years	16 (59.3)	24 (70.6)	0.355
≥59 years	11 (40.7)	10 (29.4)	
Gender			
Male	16 (59.3)	23 (67.6)	0.498
Female	11 (40.7)	11 (32.4)	
Previous glioma diagnosis			
No	22 (81.5)	21 (61.8)	0.094
Yes	5 (18.5)	13 (38.2)	
Grade			
3	5 (18.5)	15 (44.1)	**0.034**
4	22 (81.5)	19 (55.9)	
	Treatment given		
Radiotherapy			
No	1 (4.2)	1 (3.3)	1.000*
Yes	23 (95.8)	29 (96.7)	
Chemotherapy			
No	7 (29.2)	9 (29.0)	0.991
Yes	17 (70.8)	22 (71.0)	
	DNA repair gene expression		
APE1 expression			
Low	13 (48.1)	12 (36.4)	0.357
High	14 (51.9)	21 (63.6)	
NBN expression			
Low	17 (63.0)	20 (58.8)	0.742
High	10 (37.0)	14 (41.2)	
Nuclear PTEN expression			
Low	25 (92.6)	25 (73.5)	0.112*
High	2 (7.4)	9 (26.5)	
PMS2 expression			
Low	8 (29.6)	6 (17.6)	0.269
High	19 (70.4)	28 (82.4)	

* Yates' continuity correction. Significant p values (≤0.05) are shown in bold.

4.2.13 Low APE1 and low PTEN nuclear expression are associated with poor overall survival in high grade glioma

On Kaplan Meier survival analysis, both low APE1 (p=0.031) and low nuclear PTEN expression were significantly associated with poor overall survival. NBN (p=0.388), PMS2 (p=0.464) and cytoplasmic PTEN (p=0.545) expression were not significantly associated with survival in high grade glioma. Kaplan Meier survival cures for all five markers are shown in Figure 4.13.

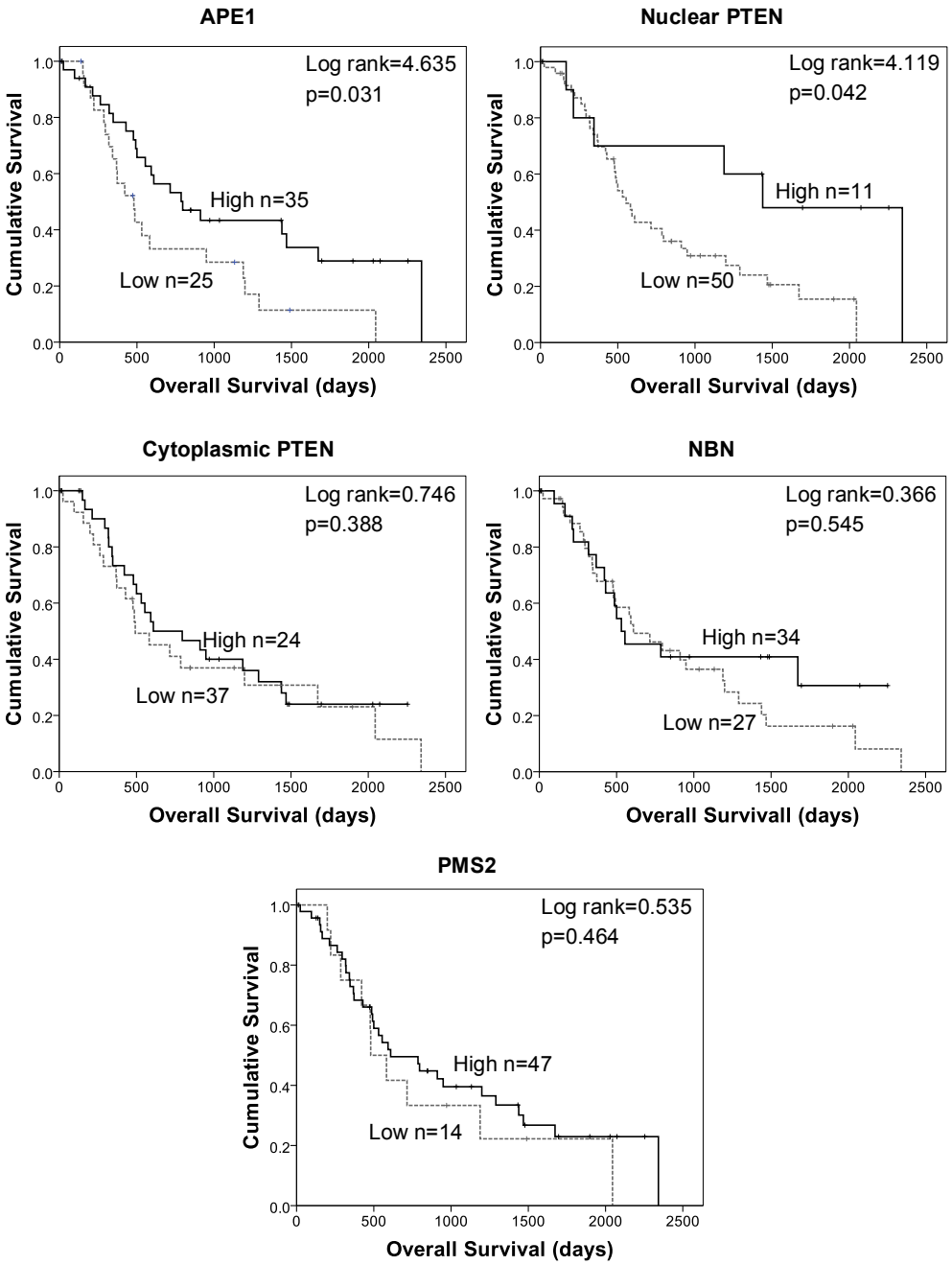


Figure 4.13. Kaplan Meier survival curves for the APE1, nuclear PTEN, cytoplasmic PTEN, NBN and PMS2 protein expression in high grade glioma.

Kaplan Meier survival curves for low/high expression of APE1, nuclear PTEN, cytoplasmic PTEN, NBN and PMS2 in 61 high grade glioma patients. Both low APE1 (p=0.031) and low nuclear PTEN (p=0.042) expression are associated with poor overall survival, as determined by the Log Rank test.

4.2.14 Combination of low nuclear PTEN and low APE1 expression is associated with poor overall survival

As both APE1 and nuclear PTEN are significantly associated with survival in high grade glioma the expression of the two markers may be connected. Kaplan Meier survival analysis was performed to investigate the combined effect of PTEN and APE1 expression on survival (Figure 4.14). Patients with both high APE1 and high nuclear PTEN expression had the best overall survival whereas patients with both low APE1 and low nuclear PTEN expression had the worst survival (p=0.027). Of the 26 patients with high expression of one marker and low expression of the other only one individual had low APE1 and high PTEN expression, hence patients with a mixed pattern of APE1 and nuclear PTEN expression were included together in one group (Group 2 in Figure 4.14).

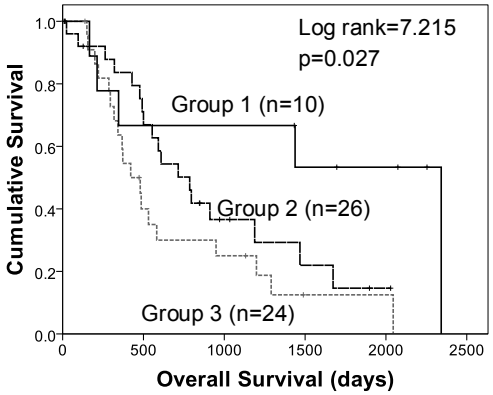


Figure 4.14. Kaplan Meier survival curve for the combined expression of APE1 and nuclear PTEN.

Group 1 = high APE1/ high nuclear PTEN, Group 2 = high APE1/low nuclear PTEN or low APE1/high nuclear PTEN and Group 3 = low APE1/low nuclear PTEN. In patients with both low APE1 and low nuclear PTEN expression survival was significantly worse than patients with both high APE1 and high nuclear PTEN expression (p=0.027), as determined by the Log Rank test.

4.2.15 Low nuclear PTEN expression is associated with survival in patients treated with chemotherapy

On Kaplan Meier survival analysis of patients who received radiotherapy neither APE1 (p=0.069), NBN (p=0.661), nuclear PTEN (p=0.106), cytoplasmic PTEN (p=0.895) or PMS2 (p=0.575) expression was significantly associated with survival (Figure 4.15).

In patients treated with chemotherapy, low nuclear PTEN expression (p=0.023) was associated with poor overall survival. APE1 (p=0.156), NBN (p=0.453), cytoplasmic PTEN (p=0.973) and PMS2 (p=0.210) expression were not associated with survival in patients treated with chemotherapy (Figure 4.16).

Analysis was not performed in patients who did not receive chemotherapy or in those who did not receive radiotherapy due to low patient numbers.

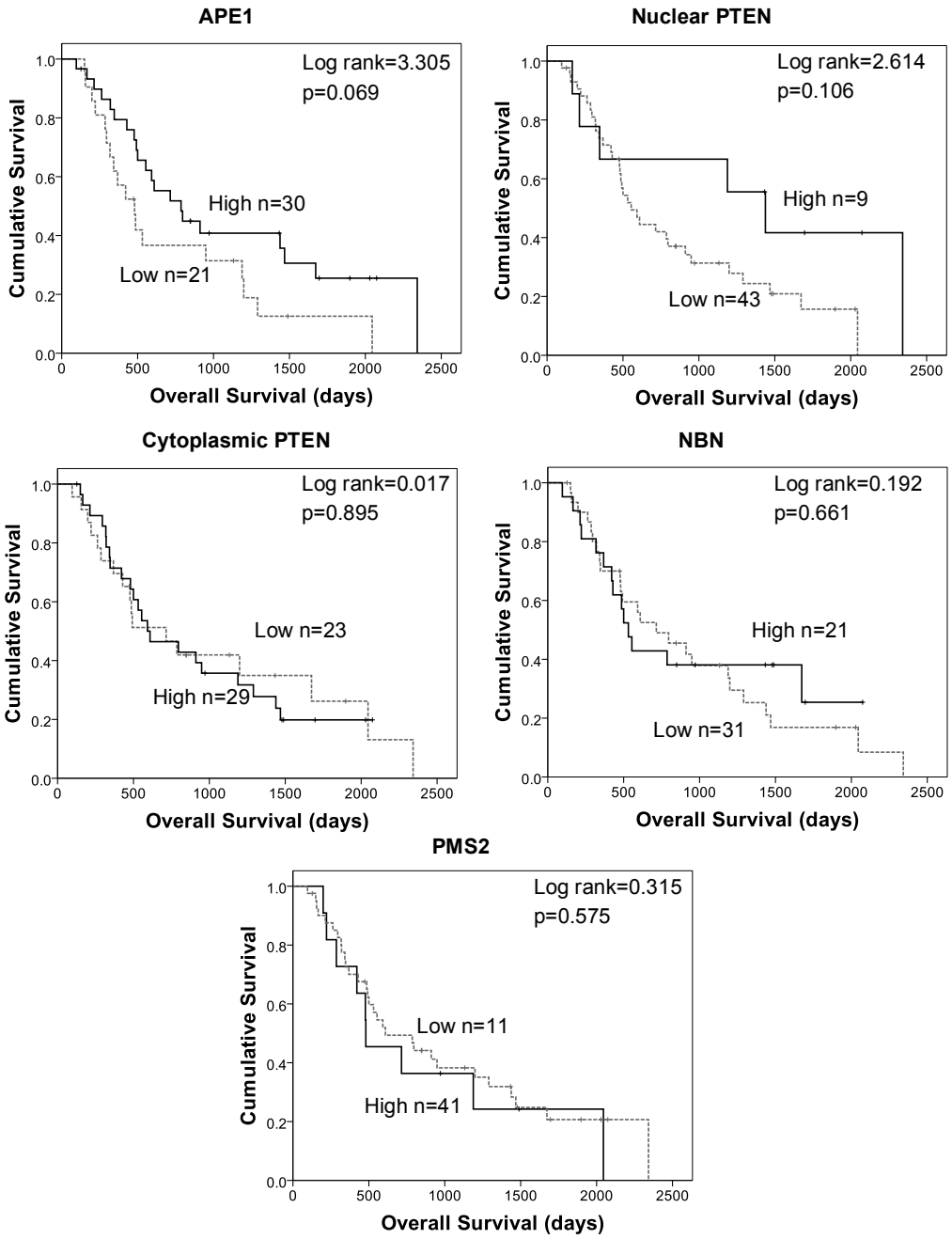


Figure 4.15. Kaplan Meier survival analysis for APE1, PTEN, NBN and PMS2 expression in patients who received radiotherapy.

On Kaplan Meier survival analysis neither APE1, nuclear PTEN, cytoplasmic PTEN, NBN or PMS2 expression were associated with overall survival in high grade glioma patients (as determined by the Log Rank test).

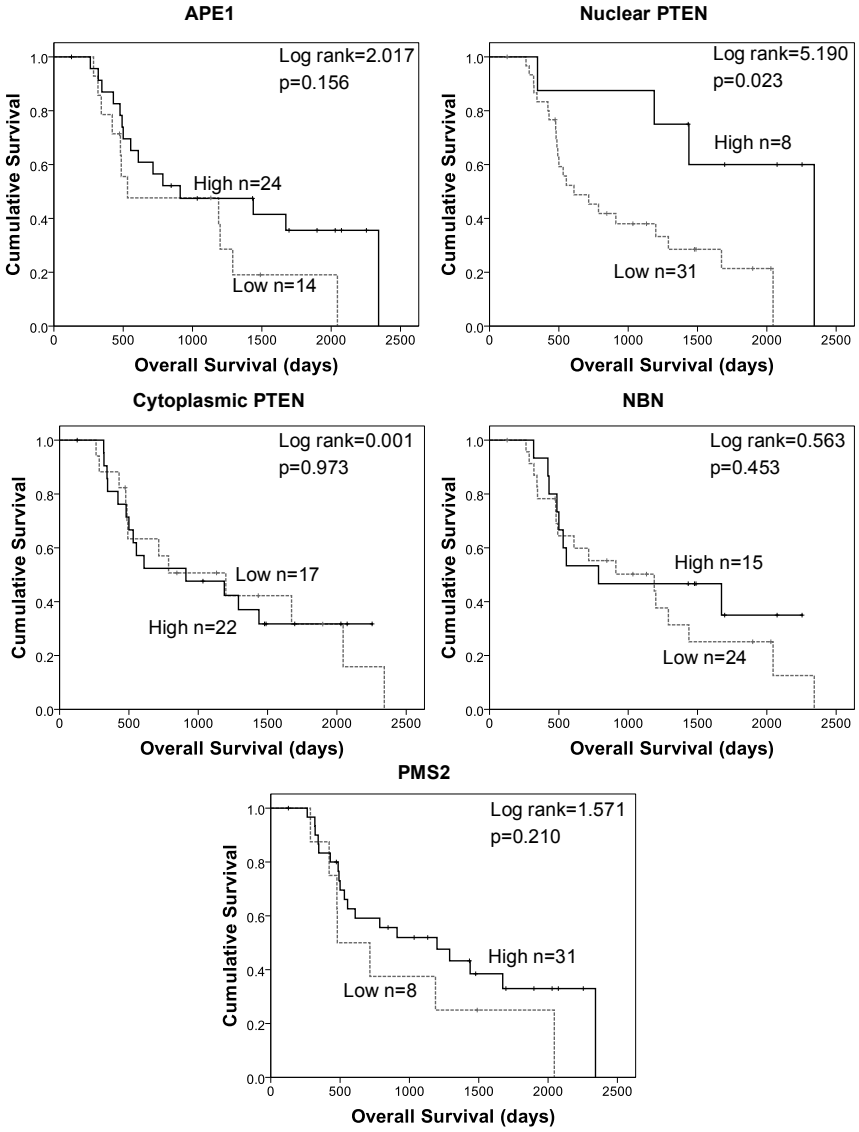


Figure 4.16. Kaplan Meier survival analysis for APE1, PTEN, NBN and PMS2 expression in patients who received chemotherapy.

On Kaplan Meier survival analysis only low nuclear PTEN expression (p=0.023) was associated with poor overall survival in high grade glioma. Neither APE1, cytoplasmic PTEN, NBN or PMS2 expression were associated with overall survival in high grade glioma patients (as determined by the Log Rank test).

4.3 Discussion

In this Chapter the expression pattern, and prognostic significance, of APE1, NBN, PMS2 and PTEN has been evaluated in 61 high grade glioma patients by immunohistochemistry. Both APE1 and nuclear PTEN expression have been found to be associated with prognosis in high grade glioma.

4.3.1 Low APE1 protein expression is associated with poor overall survival in high grade glioma

In this study, low APE1 protein expression was significantly associated with poor overall survival in high grade glioma, in keeping with the findings in Chapter 3 for *APE1* mRNA expression. Previous studies, in a variety of tumour types, have found the opposite results, with high APE1 expression associated with poor overall survival (Wang et al., 2004, Al-Attar et al., 2010). For example, in osteosarcoma, high APE1 protein expression was found to be associated with poor overall survival and siRNA knockdown of APE1 resulted in increased sensitivity to DNA damaging agents (Wang et al., 2004). In ovarian cancer, expression of APE1 was found to be associated with poor overall survival as compared to patients with negative APE1 expression. In the same study nuclear APE1 expression was associated with poor overall survival in patients treated with neoadjuvant platinum chemotherapy (Al-Attar et al., 2010). However, a recent large study of over one thousand breast cancer cases found low APE1 protein expression, as assessed by IHC, to be associated with poor survival, aggressive phenotypes and resistance to endocrine treatment (Abdel-Fatah et al., 2014c). While Bobola *et al* found that high AP

endonuclease activity was present in high grade glioma, as compared to both low grade glioma and normal brain (Bobola et al., 2001), no studies have investigated the prognostic significance of APE1 protein expression in glioma.

Interestingly, in the current study APE1 expression was exclusively nuclear and positive in all tumour sections. In an earlier study by Al-Attar *et al*, 71.9% of ovarian cancers were found to have nuclear APE1 expression, as compared to only 34.8% of gastro-oesophageal tumours (Al-Attar et al., 2010). This would suggest that differences exist between the expression patterns of APE1 between tumour types, in a similar way to how APE1 levels appear to have different prognostic implications in different tumour types.

APE1 protein expression does not appear to have predictive significance in this cohort, with APE1 protein expression not associated with survival in either patients treated with chemotherapy or with radiotherapy. In the TCGA gene expression cohort, low *APE1* mRNA expression was associated with poor overall survival, but only in patients that received chemotherapy, or that received radiotherapy, suggesting that this marker may have predictive significance. Previous cell line studies have shown that knock-down of *APE1* is associated with radiosensitivity (Naidu et al., 2010) and that APE1 inhibition results in chemosensitivity (Wang et al., 2004). These earlier studies are *in vitro* and not *in vivo* but, based on these findings, it would seem likely that patients with low APE1 expression treated with chemo- or radiotherapy would have better overall survival than those with high APE1 expression. Clearly this is in disagreement to the findings of this current study. In the current study, patients were either classified as having received chemotherapy or radiotherapy

at some point during the course of their disease. Many patients received both chemotherapy and radiotherapy during the course of their illness and were therefore included in both the analysis of chemotherapy and also radiotherapy treated patients. This classification was due to difficulties processing the large volume of treatment data available in the TCGA cohort and ascertaining the time order in which treatments were delivered. In the Nottingham cohort, the small cohort size hindered multiple subgroup analyses. Ideally, prospective treatment data would be collected for a large cohort of patients to enable subgroup analysis, for example in patients that received chemo-radiotherapy, to assess whether high or low APE1 expression is associated with a response to treatment.

Taken together these results suggest that APE1 plays a complex role in human tumours, and that this role varies with tumour type. The association between low APE1 expression and poor survival, as observed at both the mRNA and protein level, may be as a consequence of low APE1 levels promoting a mutator phenotype. In this scenario, accelerated mutagenesis may occur and promote aggressive features in tumours (Loeb, 2001). This is supported, as discussed in Chapter 3, by the finding that at the mRNA level, low *APE1* expression is associated with poor prognostic features. The lack of association between low APE1 expression and poor prognostic features in the IHC study is most likely due to the small sample size and this would need to be tested in a large cohort of glioblastoma patients.

4.3.2 Low PTEN expression is associated with poor overall survival in high grade glioma

The finding that low nuclear PTEN protein expression is associated with poor overall survival in high grade glioma is again in keeping with the results seen for *PTEN* mRNA expression in Chapter 3. Low PTEN protein expression has been previously reported to be associated with poor survival in a variety of tumour types, including mesothelioma (Opitz et al., 2008), endometrial (Kanamori et al., 2002) and colorectal cancer (Sawai et al., 2008). In previous studies, *PTEN* mutation has been shown to be associated with poor survival in high grade glioma (Smith et al., 2001). Similarly, low levels of PTEN protein expression have been found to be associated with poor survival in glioblastoma in a previous study (Ermoian et al., 2002a).

In this current study, PTEN expression has been classified as nuclear and/or cytoplasmic. No significant association was observed between nuclear and cytoplasmic PTEN expression and only nuclear PTEN expression was associated with overall survival. In previous studies of PTEN protein expression, a distinction between nuclear and cytoplasmic expression is often not made. However, in a study of patients with oesophageal squamous cell carcinoma, nuclear PTEN expression was found not to be associated with cytoplasmic PTEN expression. In the same study, only negative nuclear PTEN expression was associated with poor overall survival (Tachibana et al., 2002).

In the current study both low nuclear and low cytoplasmic PTEN expression were found to be associated with grade, with grade four glioma found to have lower levels of expression than grade three tumours. Again, this is in keeping

with earlier studies where PTEN levels have been found to be highest in normal brain, with intermediate levels seen in low grade tumours, and low levels in high grade glioma (Ermoian et al., 2002a).

The results of this study have shown low nuclear PTEN protein expression to be associated with poor overall survival in patients treated with chemotherapy. However, in the TCGA cohort low *PTEN* mRNA expression was associated with poor overall survival in patients both treated with, and not treated with, chemotherapy. This suggests that *PTEN* status does not predict response to chemotherapy.

Earlier studies have focused on PTEN expression as a predictor of radiotherapy response. One such study, in head and neck cancer patients, found that patients with low PTEN expression treated with post-operative radiotherapy had higher loco-regional recurrence rates than patients with high PTEN expression (Snietura et al., 2012). Similarly, in patients with cervical cancer treated with radiotherapy, *PTEN* mutation was associated with poor survival (Harima et al., 2001). In this current study, PTEN expression was not associated with survival in GBM patients treated with radiotherapy.

Taken together, these results suggest that PTEN expression, in particular nuclear PTEN expression, may act as a prognostic marker in high grade glioma.

4.3.3 Low PTEN expression is associated with low APE1 expression

In this current study low nuclear PTEN expression was found to be associated with low APE1 expression. Fantini *et al* have previously shown that APE1 may regulate PTEN expression via the transcription factor erg-1. In addition, the same study also demonstrated that *APE1* knockdown in HeLa cells resulted in a significant decrease in *PTEN* expression (Fantini et al., 2008). The work by Fantini *et al* may therefore offer an explanation as to the association between APE1 and nuclear PTEN expression in the current study. Furthermore, in the current study, tumours with low APE1 and low PTEN expression, either at the mRNA or at the protein level, had significantly worse overall survival as compared to patients with both high APE1 and high PTEN expression. Individuals with either low APE1/high PTEN or high APE1/low PTEN expression had an intermediate outcome. Taken together, this data would suggest that individuals with either low APE1 or low PTEN expression represent a poor prognostic group and could be targeted with more intensive, individualised treatment regimens; this is particularly the case for patients with both low APE1 and low nuclear PTEN expression. Exactly what these treatment regimens would entail would need to be the subject of future studies, but may involve using a synthetic lethality strategy.

4.3.4 NBN and PMS2 protein expression are not associated with overall survival in high grade glioma patients

In this study, NBN protein expression was not associated with survival, although, as shown in Chapter 3, low *NBN* mRNA expression was associated

with poor overall survival. The results at the protein level for NBN are in keeping with the results of an earlier study by Seol *et al*, who found that NBN protein expression was not associated with either progression free or overall survival in glioblastoma. The study by Seol *et al* was also small and, unlike this study, only included patients with a diagnosis of glioblastoma. Other patient characteristics are similar between the two studies, with 66% of patients being male and a median age of 54.4 years in the Seol study, as compared to 70.5% and 54 years respectively in the current study. In other tumours, such as head and neck cancers, and in myelodysplastic syndrome, high NBN expression has been found to be associated with poor survival (Yang et al., 2006, Kefala et al., 2013).

PMS2 protein expression was also found not to be associated with survival in the current study, despite high *PMS2* mRNA expression being found to be associated with poor overall survival. PMS2 expression has generally not been explored alone. The concept of microsatellite instability (MSI), however, has been extensively studied, particularly in colorectal cancer. MSI has been found to be present in approximately 15% of colorectal cancer and represents a hypermutable phenotype that is caused by a loss of DNA MMR activity. While a small proportion of these individuals have an inherited cancer predisposition syndrome known as Hereditary non-polyposis colorectal cancer (HNPCC) or Lynch syndrome, the vast majority have sporadic MSI, commonly associated with the loss of expression of MLH1 and PMS2 (Boland and Goel, 2010). Multiple studies have investigated the prognostic, and indeed predictive, significance of MSI in colorectal tumours with the results of a large

meta-analysis suggesting that the presence of MSI is associated with a favourable prognosis (Guastadisegni et al., 2010). In the same study, microsatellite stability (MSS) was associated with a significant benefit from 5-flurouracil (5-FU) based chemotherapy. In MSI colorectal cancers, it was not possible to determine if 5-FU based chemotherapy conferred a benefit due to study heterogeneity (Guastadisegni et al., 2010). As loss of PMS2 is commonly seen in MSI, then it could be concluded that the findings of the current mRNA expression study are in keeping with the results of the earlier studies. This was not, however, reproduced at the protein level.

The differences seen between the results for NBN and PMS2 expression at the mRNA and protein levels could be due simply to the small size of the Nottingham cohort. An alternative explanation could be the inclusion of grade 3 glioma patients in the protein expression cohort. It may be that the differences observed are a reflection the mRNA expression cohorts representing a different patient group to those included in the protein expression cohort. Alternatively, a true difference may be present between the expression of NBN and PMS2 at the mRNA and protein level, which could be due to factors that are post-transcriptional, such as the presence of microRNAs, protein phosphorylation or ubiquitination.

4.3.5 Limitations and future work

A major limitation of the protein expression study has been the number of tumour samples available, although a cohort size of over 60 patients is not dissimilar to other studies investigating this tumour type. Reassuringly, despite

the cohort size, analysis of the patient demographics would imply that this is a representative sample of high grade glioma patients, with old age and higher tumour grade associated with poor overall survival as would be expected. Unfortunately, performance status data was not available and therefore the prognostic significance of this could not be confirmed.

Another potential limitation is that four patients had already received radiotherapy treatment, and three chemotherapy treatment, prior to collection of the study sample which may impact on the study results. Also the natural history of glioma has made the collection of survival data challenging. Survival time was calculated from the date that the tumour sample was collected, as this is the date for which DNA repair protein expression could be assessed. Forty-three patients had not previously been diagnosed with a glioma and therefore the study sample represented their initial diagnostic sample. Seven patients had previously been diagnosed with a low grade glioma on imaging and therefore the trial sample represented the initial diagnosis of a high grade tumour. Similarly, eight patients had a previous histological diagnosis of a grade 2 tumour but the samples included in the study are the first diagnostic samples representing high grade transformation. Three patients, however, had previously been diagnosed with a high grade tumour and therefore the survival time calculated for these individuals represents survival from time of tissue collection and DNA repair protein expression assessment, rather than survival from the time of diagnosis with a high grade tumour. This may potentially influence the study results.

Also, again due to the number of samples available in the current study, both grade 3 and grade 4 glioma have been included together. In the mRNA study (Chapter 3), only glioblastoma patients were included and therefore it is possible that the results observed in the Nottingham cohort may reflect the inclusion of grade 3 tumours. Study size has also hindered the investigation of the predictive significance of the four DNA repair proteins. As with the mRNA study, survival was assessed for each marker in patients that received chemotherapy and in those that received radiotherapy. Due to the low numbers of patients that did not receive these treatments it was not possible to perform survival analysis in the untreated subgroups. Additionally, some patients had received both chemotherapy and radiotherapy, and many had received several courses of treatment over the duration of their illness. While this component of the study has served as a useful screen for the predictive significance of these markers, this would be better studied with a greater number of patients, whereby specific treatment subgroups of patients could be investigated.

Ideally, a large multicentre prospective study is required, collecting tissue prior to any treatment for both gene and protein expression analysis. This would also enable the prospective collection of treatment and demographic data, such as performance status. Availability of normal brain tissue, plus samples from lower grade tumours, would also enable any differences in expression of DNA repair markers with malignant transformation to be assessed.

4.3.6 Summary

This study, in combination with the mRNA expression study, has identified low PTEN and low APE1 expression as poor prognostic markers in high grade glioma. It may therefore be possible to target these poor prognosis APE1 and/or PTEN deficient tumours with a personalised treatment approach. One such strategy would be to employ a synthetic lethality approach, and hence target these tumours with an inhibitor of a second DNA repair pathway. As discussed in Chapter 1, targeting PTEN deficient tumours with a synthetic lethality strategy has been the topic of several studies in recent years, with variable results. In the following chapter the potential of targeting glioma cell lines, both proficient and deficient in PTEN expression, with an ATM inhibitor will be explored.

CHAPTER 5

Targeting ATM in PTEN deficient and proficient glioblastoma cell lines

5 Results

5.1 Introduction

PTEN mutations are common in a range of human tumours including melanoma (Guldberg et al., 1997), prostate (Li et al., 1997, Cairns et al., 1997) and breast cancer (Li et al., 1997, Stemke-Hale et al., 2008). PTEN mutations have been reported in up to 80% of cases of glioblastoma, and are more commonly seen in high grade tumours than in low grade glioma (Rasheed et al., 1997, Sano et al., 1999). Over recent years, evidence has emerged that tumours which are deficient in PTEN may have down-regulation of DNA DSB repair and, as a consequence, may be amenable to targeting with a synthetic lethality strategy (McCabe et al., 2006, Mendes-Pereira et al., 2009, Minami et al., 2012). Earlier studies have shown that PARP inhibition is synthetically lethal in PTEN deficient endometrial and colorectal cancer cell lines (146). Similarly, PTEN deficient melanoma cell lines have been shown to be amenable to selective targeting with APE1 inhibitors (Abbotts et al., 2014). Another potential synthetic lethality strategy involves the use of ATM inhibitors in PTEN deficient tumours (McCabe et al., 2015). This potential treatment approach will be explored further in this chapter in a panel of glioma cell lines, as well as in isogenic *PTEN* knock-down glioma cell lines.

5.1.1 Rationale for using an ATM inhibitor in PTEN deficient glioblastoma

PTEN is involved in the regulation of multiple cellular processes, including DNA repair and cell cycle checkpoint control, in addition to its well established role as a negative regulator of the PI3K signalling pathway (Keniry and Parsons, 2008) (See Section 1.8). PTEN deficiency has been shown to be associated with both chromosomal instability and defective DNA repair (Shen et al., 2007, Puc et al., 2005). Defective DNA DSB repair has been demonstrated in PTEN deficient tumours (Abbotts, 2013, Mendes-Pereira et al., 2009, Abbotts et al., 2014)). While the mechanism behind this remains controversial, it has been proposed that PTEN may transcriptionally regulate expression of the critical HR protein RAD51, via the transcription factor E2F-1 (Gupta et al., 2009, Puc and Parsons, 2005, Shen et al., 2007). Other studies have proposed that PTEN status may influence the recruitment of HR factors such as RAD51 to DNA DSBs (Bassi et al., 2013). In order to perform its role in DNA repair, PTEN is required to be located in the nucleus. Nuclear PTEN also plays a role in cell cycle checkpoint control by suppressing cyclin D1 and stalling progression through the G1/S checkpoint. Cells deficient in nuclear PTEN demonstrate loss of the G1/S checkpoint and therefore accumulate DNA damage that has not had time to repair (Diao and Chen, 2007, Weng et al., 2001, Radu et al., 2003). Previous studies have shown that PTEN deficient colorectal and endometrial cancer cell lines can be targeted with PARP inhibition (Mendes-Pereira et al., 2009), although the same does not hold true for prostate cancer cell lines (Fraser et al., 2012). PTEN deficient melanoma

cell lines have also been shown to be amenable to a synthetic lethality approach with the use of APE1 inhibitors (Abbotts et al., 2014).

ATM is a key sensor of DNA damage and triggers the DDR (Abraham, 2001, Weber and Ryan, 2015). While traditionally ATM has been thought to be activated and recruited to the site of DNA DSBs via the MRN complex, more recently ATM has also been shown to be activated by DNA SSBs (Khoronenkova and Dianov, 2015). As a consequence of its activation, ATM further activates downstream targets that are involved in the regulation of cell cycle checkpoints, DNA repair and apoptosis (Khalil, 2012). ATM has been shown to play a role in the control of all cell cycle checkpoints, including the G2/M checkpoint. Manipulation of ATM levels, either via knock-down technology or pharmaceutical inhibition, has been shown to increase radio-sensitivity in glioblastoma cell lines (Hickson et al., 2004).

Taken together, the evidence would suggest that targeting PTEN deficient tumours with ATM inhibitors could represent a potential synthetic lethality strategy in glioblastoma. It is proposed that inhibition of ATM would result in loss of activation of the DDR in response to DNA DSBs and potentially SSBs. In PTEN deficient tumours, defective HR would result in the massive accumulation of this damage and hence cell death. As PTEN deficient cells have an impaired G1/S cell cycle checkpoint, they are reliant on the G2/M checkpoint. ATM inhibition would result in loss of the G2/M checkpoint and hence increase the amount of DNA damage to such an extent that cells undergo apoptosis.

5.1.2 PTEN mutation in glioma cell lines

In order to investigate ATM synthetic lethality in PTEN deficient glioblastoma, a suitable cell line models are required. Previous studies have demonstrated that disruption of PTEN expression is common in glioma cell lines, and report a frequency of between 60-80% (Ishii et al., 1999). In a study by Ishii *et al*, 34 glioma cell lines were investigated for PTEN status; 73.5% of these cell lines were found to have mutations or deletion in the PTEN gene, including the U87MG and A172 cell lines. Additional studies have confirmed the PTEN status of the U87MG and A172 cell lines (Li and Sun, 1998, Bassi et al., 2013, Ishii et al., 1999). Review of the COSMIC database (Forbes et al., 2015) (www.cancer.sanger.ac.uk) records the presence of a homozygous intronic substitution at 209+1G>T in the U87MG cell line and a loss of copy number for the PTEN gene in the A172 cell line. U87MG and A172 cell lines were therefore included in this study as PTEN deficient cell lines. The LN18 and LN229 cell lines are widely reported in the literature as *PTEN* wild-type (Forbes et al., 2015, Wick et al., 1999, Ishii et al., 1999).

5.1.3 Overcoming genetic diversity in glioma cell lines

Glioblastoma have commonly undergone a massive number of genomic aberrations (Fischer and Meese, 2007). DNA repair has been shown to play a key role in the development of these genetic modifications; for example, reduced DSB repair ability has been shown to be involved in the process of gene amplification in glioblastoma (Fischer and Meese, 2007). In view of this high burden of genetic aberration, it would seem logical that glioblastoma

tumours, and indeed glioma cell lines, would have defects in DNA repair and other key cellular pathways.

While the four glioma cell lines studied have been selected based on their PTEN status, due to the aberrant genetic background of glioblastoma, it is clear that these cell lines will differ in more than just PTEN expression. For example, in the study by Ishii *et al*, mutations or deletions were present in *p53* in 76.5% of the glioma cell lines studied. 64.7% of the glioma cell lines studied had mutations or deletions in $p14^{ARF}$ and similarly 64.7% in *p16*. Of the 34 cell lines studied, nine had mutations or deletions in all four genes (*PTEN*, *TP53*, $p14^{ARF}$ and *p16*), while 22 had aberrations in at least three of these genes (Ishii et al., 1999). Exploring ATM synthetic lethality in standard glioblastoma cell lines with varied PTEN status provides a proof of concept; however, it is difficult to disentangle the influence of PTEN status versus the effect of the wider genetic background. This problem can potentially be overcome by gene knock-down techniques.

5.1.4 Gene silencing techniques

Over the years various techniques have been employed to investigate gene function, including antisense oligonucleotide and ribozyme techniques. In recent years RNA interference (RNAi) techniques have been widely used to 'knock-down' expression of a gene in vertebrate cells. RNAi by siRNA is a commonly used technique but effectiveness can be influenced by a variety of factors, including the cell line, effectiveness of the siRNA and the efficiency of transduction. siRNA mediated RNAi is only a temporary method of gene

'knock-down', lasting for approximately three to five cell doublings (McManus and Sharp, 2002).

The use of short hairpin RNA (shRNA) offers a more stable method of long-term gene 'knock-down'. shRNA is a modified form of siRNA that consists of two strands of RNA that are joined by a short loop, and as a consequence are more stable than siRNAs. shRNAs are delivered into cells using a vector. Lentiviruses are commonly used vectors and are able to infect non-dividing cells. Once the vector DNA is integrated into the host cell chromosomes, the use of antibiotic selection markers can be utilised to select for transduced cells (Lambeth and Smith, 2013).

5.1.5 Development of isogenic cell lines by lentiviral transduction

The Thermo Scientific Open Biosystems Expression Arrest TRIPZ Lentiviral shRNAmir system has been used in this study to develop *PTEN* knock-down LN18 and LN229 cell lines by lentiviral transduction. The steps involved in this process include: plasmid preparation and amplification, lentiviral production and transduction of glioma cell lines, confirmation of shRNA induction and confirmation of *PTEN* knock-down.

shRNA constructs can be purchased as a stock solution of ready-transformed *Escherichia coli* that express the plasmid of interest. Plasmids are amplified by overnight culture prior to transfection into Lenti-X 293T human embryonic kidney cells for virus production. After 48 hours incubation, supernatants are harvested and viral concentration determined. This Lentivirus-containing supernatant can then be used to transduce the target cell line(s).

After transduction a cDNA copy of the viral genome is made by an RNA-dependent DNA polymerase (reverse transcriptase). Viral cDNA is subsequently incorporated into the host genome with the assistance of a viral integrase. The shRNA can now be transcribed in the nucleus by RNA polymerase II or III and then further processed into pre-shRNA by Drosha, a Class 2 ribonculease III enzyme. These pre-shRNAs are now transported to the cytoplasm by Exportin-5 where they are processed to form siRNAs by Dicer, a double-stranded RNA-specific RNase III enzyme. The siRNA molecules can now bind to the RNA-induced silencing complex (RISC), where the duplex is unwound and cleavage of the sense strand is performed by endonucleases. The anti-sense strand then binds to a complimentary mRNA molecule, which is subsequently degraded by the RNase H-like activity of Argonaute 2. This results in silencing of the target gene (Moore et al., 2010, ThermoScientific, 2009).

Three Thermo Scientific TRIPZ PTEN shRNA constructs (158, 317 and 477) were used to induce PTEN knock-down, each targeting a different region of the gene. These were selected as they had previously been shown to result in the highest levels of *PTEN* knock-down in the MeWo melanoma cell line by Dr Rachel Abbotts (Abbotts, 2013). The Thermo Scientific Open Biosystems Expression Arrest short hairpin RNA constructs contain a hairpin stem of 22-nucleotide dsRNA with a sequence directed against the mRNA of interest. A 19-nucleotide loop from human microRNA-30 (miR-30) forms the hairpin loop. The presence of the miR-30 loop, together with an additional 125 nucleotide flanking sequence increases the Drosha and Dicer processing 10-fold, as compared to conventional shRNA without the miRNA (ThermoScientific, 2009).

The pTRIPZ vector has been designed so that exposure to doxycycline results in induction of shRNA production and is therefore described as 'Tet-On'. In the presence of tetracycline, the TRIPZ transactivator (reverse tetracycline transactivator (rtTA3)) binds to, and activates, the tetracycline response element (TRE) promotor. This process drives expression of the shRNA and also of a TurboRFP reporter, allowing visualization of transcript expression. An additional feature of the TRIPZ vector is the presence of a puromycin resistance gene, enabling selection of stable cell lines in the presence of puromycin at a dose lethal to non-transduced cells (ThermoScientific, 2009).

An overview of the pTRIPZ lentiviral vector is shown in Figure 5.1 and a summary of the components of the vector is shown in Table 5.1 (ThermoScientific, 2009).

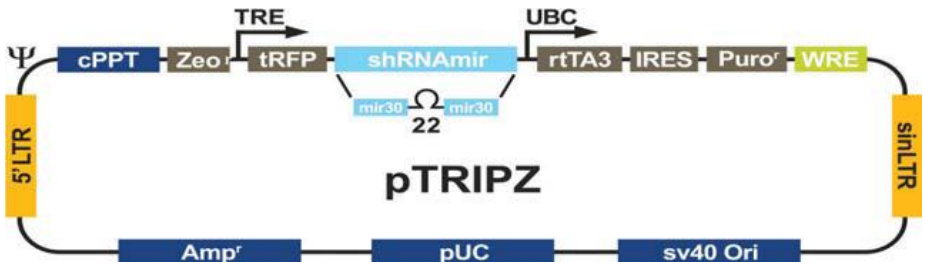


Figure 5.1. A simplified vector map of the pTRIPZ lentiviral vector.

The various components of the vector system are summarised in Table 5.1. Modified from (ThermoScientific, 2009).

Table 5.1. An overview of the function of the pTRIPZ lentiviral vector elements.

Vector Element	Function
5'LTR	5' long terminal repeat
Amp'	Ampicillin (carbenicillin) bacterial selectable marker
cPPT	Central Polypurine tract helps translocation into the nucleus of non-dividing cells
IRES	Internal ribosome entry site. Allows rtTA3 and Puro' expression in a single transcript
pUC Ori	High copy replication and maintenance plasmid in *E. coli*
Puro'	Puromycin resistance for selection of stable integrants
shRNAmir	miR-30-adpated shRNA
sinLTR	3' self-inactivating terminal repeat
TRE	Tetracycline-inducible promoter
tRFP	TurboRFP reporter
UBC	Drives expression of rtTA3 and IRES-puro
WRE	Woodchuck hepatitis post-transcriptional regulatory element enhances transgene expression in target cells
Zeo	Zeocin selectable marker

Modified from (ThermoScientific, 2009).

5.1.6 Study rationale

The data presented in Chapters 3 and 4 indicate that PTEN deficiency is associated with a poor prognosis in high grade glioma. Earlier studies have shown conflicting results regarding the effectiveness of targeting PTEN deficient tumours using a synthetic lethality strategy (Mendes-Pereira et al., 2009, Fraser et al., 2012, Abbotts et al., 2014, McEllin et al., 2010). Previous work from our group has demonstrated that PTEN deficient melanoma cell lines can be targeted successfully with APE1 inhibitors (Abbotts R et al., 2013). Further work has revealed the potential to target BRCA1 deficient tumours with inhibitors of ATM and DNA-PKcs (Albarakati et al., 2015). More recently, a potential synthetic lethality approach for targeting PTEN deficient cell lines with ATM inhibitors has been reported (McCabe et al., 2015). In this chapter the therapeutic potential of targeting PTEN deficient glioblastomas with ATM inhibitors using a synthetic lethality strategy will be investigated.

5.2 Results

5.2.1 An exploration of ATM expression in high grade glioma

The TCGA cohort has been used to explore the association between *ATM* mRNA expression and survival. Kaplan Meier survival analysis was performed for each probe of *ATM* in the whole cohort and separately in patients with high and low *PTEN* expression. Associations were also explored between *ATM* and *PTEN* expression.

In addition, 60 high grade glioma samples (Nottingham cohort) were stained for ATM protein expression using immunohistochemistry. Kaplan Meier survival analysis was performed for ATM expression in the whole cohort and separately in patients with high and low PTEN expression.

5.2.1.1 ATM mRNA expression is associated with survival in GBM with high levels of PTEN mRNA expression

Three probes for *ATM* were identified in the TCGA dataset: 208442_s_at, 210858_x_at and 212672_at. *ATM* expression for all three probes was not associated with survival in the TCGA cohort (208442_s_at p=0.052, 210858_x_at p=0.241, 212672_at p=0.291) (Figure 5.2).

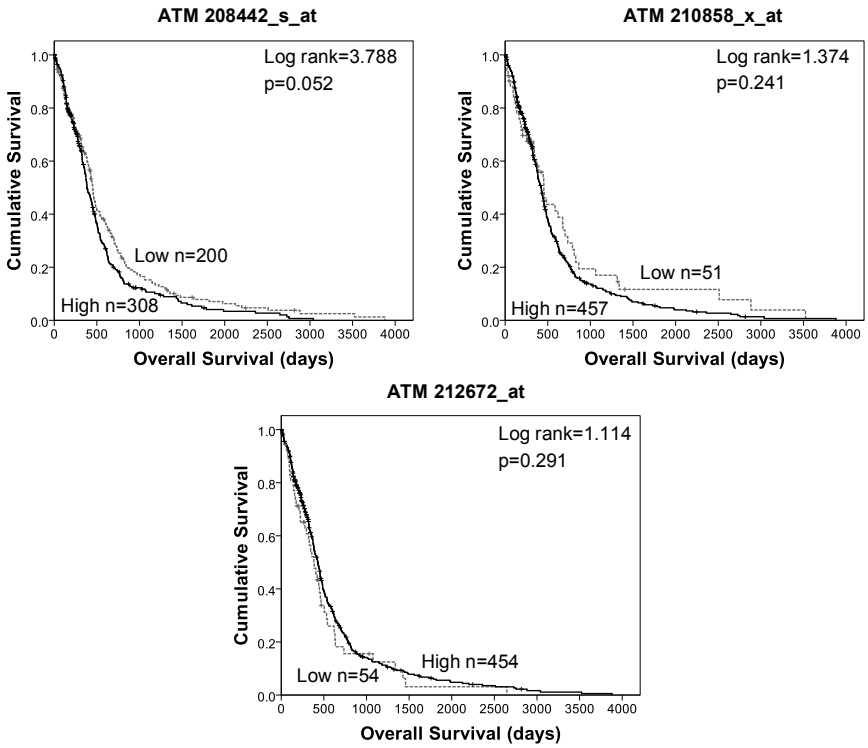


Figure 5.2. Kaplan Meier survival curves for the three probes representing *ATM* expression in the TCGA dataset.

mRNA expression levels for the three *ATM* probes were dichotomised into low/high expression using Xtile. Kaplan Meier survival analysis was performed and an association between expression level and overall survival was sought using the Log Rank test. Expression levels of none of the 3 probes tested were found to be associated with survival.

ATM mRNA expression was also investigated in patients with both high and low *PTEN* mRNA expression (probe 204054_at). On Kaplan Meier survival analysis, in patients with low *PTEN* mRNA expression, none of the probes for *ATM* (208442_s_at p=0.124, 210858_x_at p=0.656, 212672_at p=0.547) were significantly associated with survival (Figure 5.3).

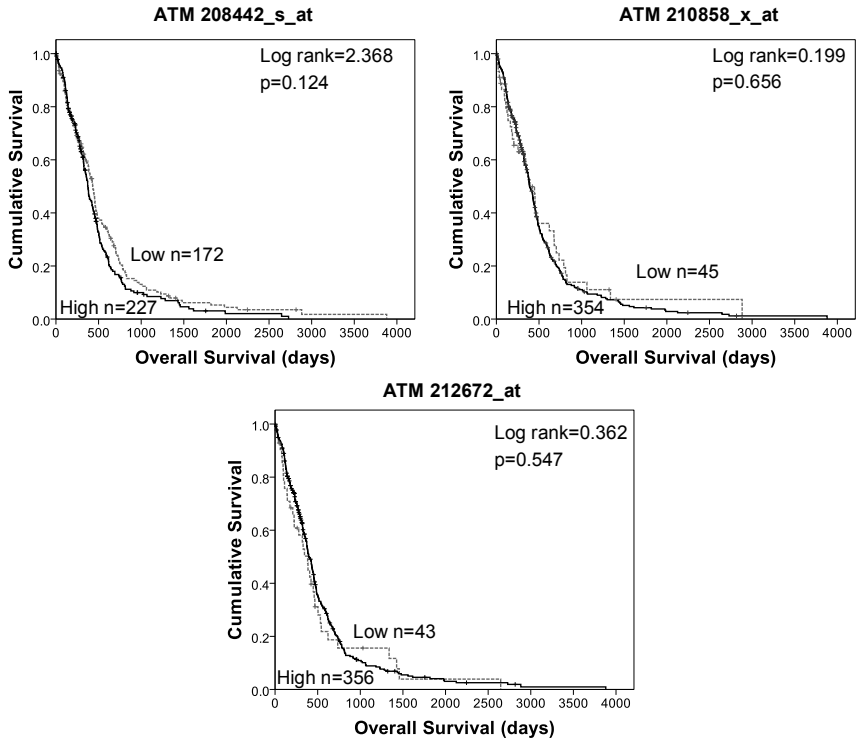


Figure 5.3. Kaplan Meier survival curves for the three probes representing *ATM* expression in the TCGA cohort for patients with low *PTEN* mRNA expression.

mRNA expression levels for the three *ATM* probes were dichotomised into low/high expression using Xtile. Kaplan Meier survival analysis was performed and an association between expression level and overall survival was sought using the Log Rank test. Expression levels of none of the 3 probes tested were found to be associated with survival in patients with low *PTEN* mRNA expression.

However, in patients with high *PTEN* mRNA expression, high expression of *ATM* probes 208442_s_at (p=0.030) and 210858_x_at (p=0.025) was associated with poor overall survival, but not probe 212672_at (p=0.295) (Figure 5.4).

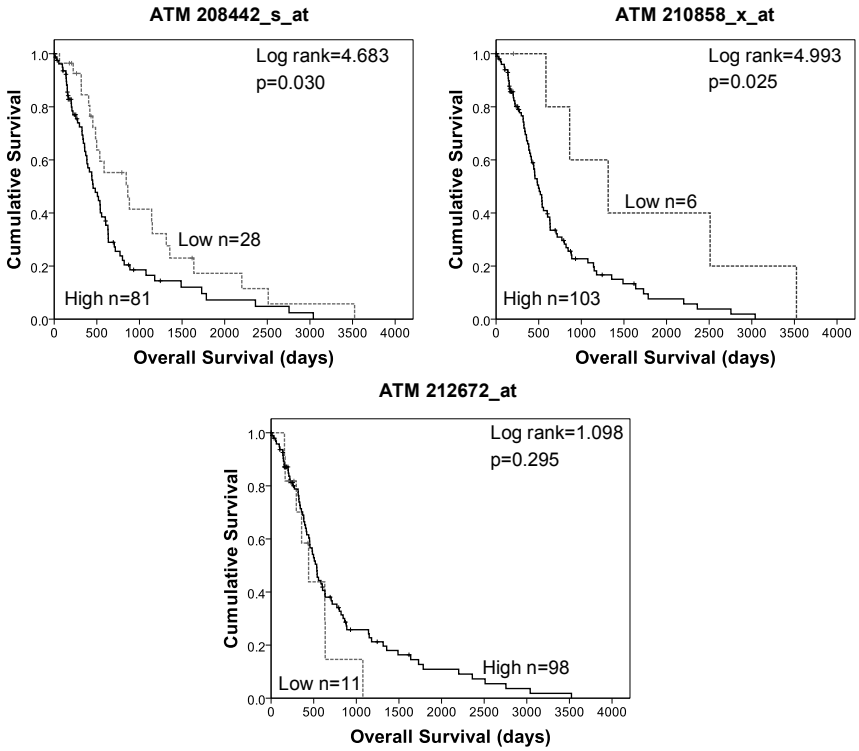


Figure 5.4. Kaplan Meier survival curves for the three probes representing *ATM* expression in the TCGA cohort for patients with high *PTEN* mRNA expression.

mRNA expression levels for the three *ATM* probes were dichotomised into low/high expression using Xtile. Kaplan Meier survival analysis was performed and an association between expression level and overall survival was sought using the Log Rank test. Expression levels of both the 208442_s_at (p=0.030) and 210858_x_at (p=0.025) *ATM* probes were found to be significantly associated with survival in patients with high *PTEN* mRNA expression.

5.2.1.2 PTEN mRNA expression is associated with ATM mRNA expression in the TCGA cohort

On chi squared analysis in the TCGA cohort, low *PTEN* (probe 204054_at) expression was associated with low *ATM* (probe 208442_s_at) expression (p=0.001). No association was seen between *PTEN* expression and expression of *ATM* probes 210858_x_at (p=0.075) and 212672_at (p=0.837).

Further associations were sought between *ATM* expression and clinical variables. No significant associations were demonstrated between age, gender, performance status and treatment for any of the *ATM* probes studied. The associations between *ATM* probe 208442_s_at and clinical variables are shown in Table 5.2.

Table 5.2. Clinicopathological associations of *ATM* (probe 208442_s_at) mRNA expression in the TCGA Cohort.

	***ATM* (208442_s_at) expression**		
Variable	**Low (N=200)** Number (%)	**High (N=308)** Number (%)	X^2 **p value**
Age			
<59 years	97 (48.5)	145 (47.1)	0.754
≥59 years	103 (51.5)	163 (52.9)	
Gender			
Male	78 (39.0)	122 (39.6)	0.891
Female	122 (61.0)	186 (60.4)	
Performance status			
0-1	110 (72.8)	175 (76.1)	0.476
2-4	41 (27.2)	55 (23.9)	
Radiotherapy treatment			
No	52 (26.0)	83 (26.9)	0.813
Yes	148 (74.0)	225 (73.1)	
Chemotherapy treatment			
No	62 (31.0)	95 (30.8)	0.970
Yes	138 (69.0)	213 (69.2)	

5.2.1.3 Evaluation of ATM protein expression in high grade glioma

5.2.1.3.1 Optimisation of ATM staining in glioma

Immunohistochemistry was performed as described in Section 2.3. The optimal concentration, and conditions, for ATM antibody staining of paraffin embedded tissue sections by IHC had previously been determined by Mr Paul Moseley in breast tissue (Abdel-Fatah et al., 2014a). In the previous study, ATM antibody was used at a concentration of 1:100 and incubated overnight at 4°C. In order to ascertain if this was an appropriate antibody concentration for staining glioma sections, full-face glioma optimisation sections were stained at the following concentrations: 1:75, 1:100, 1:125, 1:150 and 1:200 (Figure 5.5). After review by an expert histopathologist (TA), 1:100 was found to be the optimal concentration for staining brain tissue.

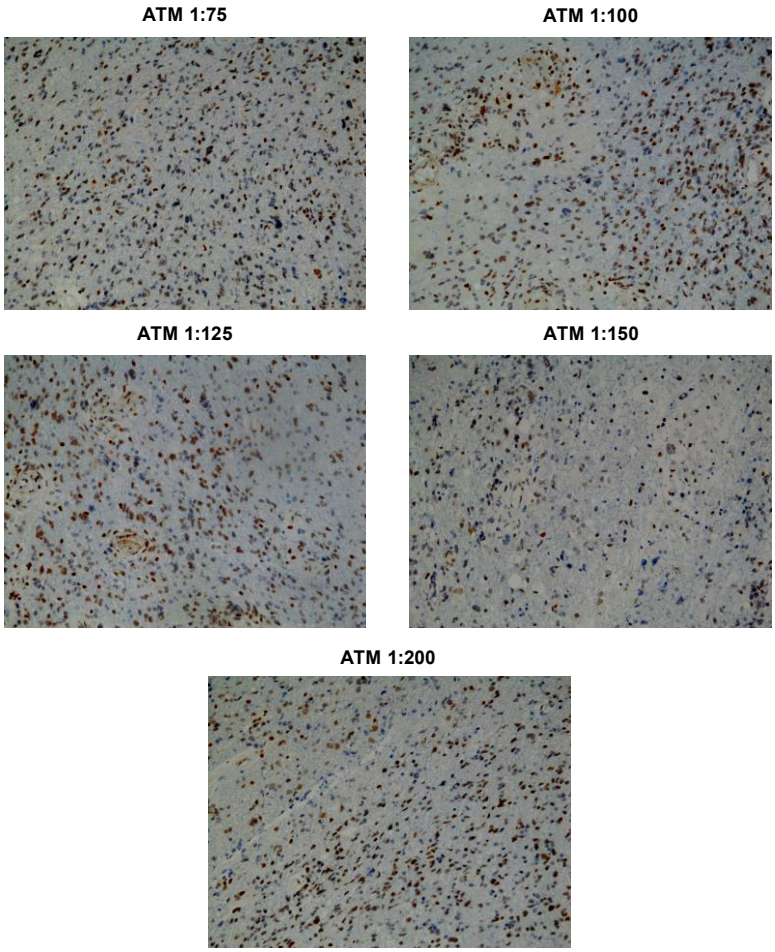


Figure 5.5. Range of ATM antibody concentrations used for optimisation.

Full-face glioma sections were stained at ATM antibody concentrations of 1:75 to 1:200. The optimal concentration for staining glioma tissue was found to be 1:100 dilution.

5.2.1.3.2 High grade gliomas demonstrate variable amounts of ATM protein expression

ATM expression was recorded as a percentage of cells positive for ATM and ranged from 1-85% (median=35%). All samples demonstrated nuclear ATM expression and no cytoplasmic staining was observed. Examples of high and low ATM expression are shown in Figure 5.6.

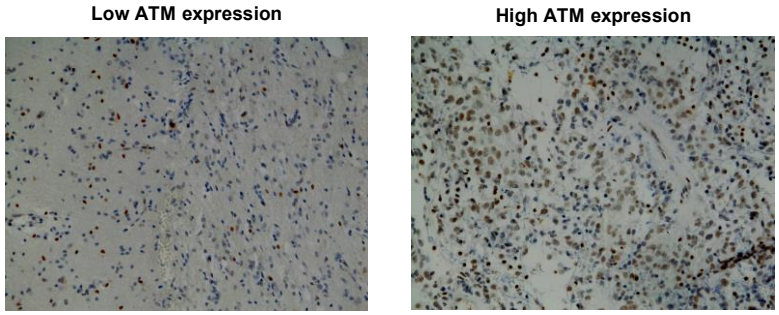


Figure 5.6. Examples of ATM expression in high grade glioma sections.

An example of a high grade glioma with low ATM protein expression (≤75% cells positive) is shown on the left, while an example of high ATM protein expression (>75% cells positive) is shown on the right.

On Chi-squared analysis, ATM protein expression was not associated with clinicopathological features in high grade glioma (Table 5.3). ATM protein expression was also not associated with either nuclear (p=0.427, Yates continuity correction) or cytoplasmic (p=0.770, Yates continuity correction) PTEN protein expression.

Table 5.3. Clinicopathological associations of ATM expression in the Nottingham Cohort.

	ATM expression		
Variable	**Low (N=51)** Number (%)	**High (N=9)** Number (%)	X^2 **p value**
Age			
<59 years	34 (66.7%)	5 (55.6%)	0.791*
≥59 years	17 (33.3%)	4 (44.4%)	
Gender			
Male	34 (66.7%)	4 (44.4%)	0.368*
Female	17 (33.3%)	5 (55.6%)	
Previous glioma diagnosis			
No	36 (70.6%)	7 (6.5%)	0.968*
Yes	15 (29.4%)	2 (22.2%)	
Grade			
3	16 (31.4%)	3 (33.3%)	1.000*
4	35 (68.6%)	6 (66.7%)	
Radiotherapy treatment			
No	2 (4.5%)	0 (0.0%)	1.000*
Yes	42 (95.5%)	9 (100.0%)	
Chemotherapy treatment			
No	13 (28.9%)	2 (22.2%)	1.000*
Yes	32 (71.1%)	7 (77.8%)	

* Yates' continuity correction. Significant p values (≤0.05) as shown in bold.

5.2.1.3.3 ATM protein expression is associated with survival in high grade glioma patients with high levels of PTEN protein expression

In the Nottingham cohort of high grade glioma patients, ATM protein expression is not significantly associated with survival (p=0.074) (Figure 5.7).

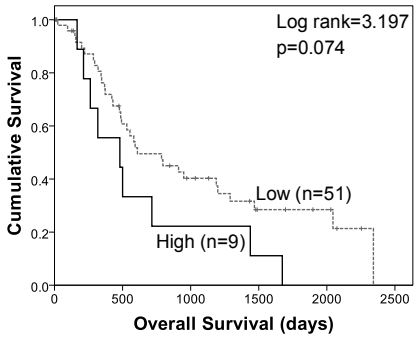


Figure 5.7. Kaplan Meier survival curves for ATM protein expression in the Nottingham Cohort.

ATM expression level has been dichotomised into low/high expression using Xtile. Kaplan Meier survival analysis was performed and an association between expression level and overall survival was sought using the Log Rank test. Expression level of ATM protein was not found to be significantly associated with survival.

In high grade glioma patients with low PTEN protein expression, ATM protein expression is not associated with survival (p=0.371) (Figure 5.8). However, in high grade glioma patients with high PTEN protein expression, high levels of ATM protein expression are significantly associated with poor overall survival (p=0.022) (Figure 5.8).

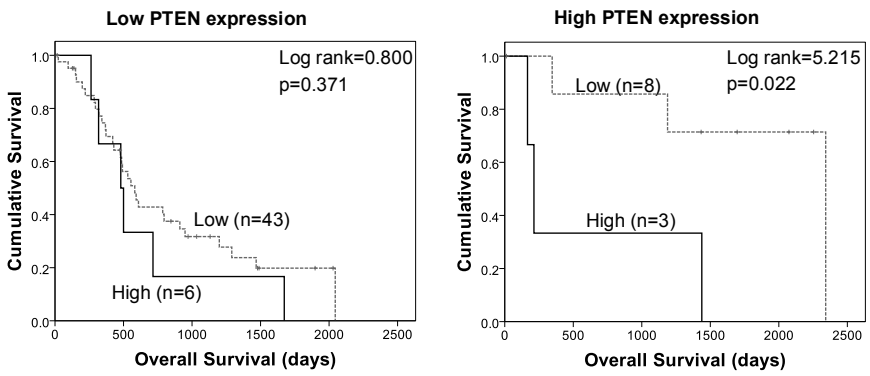


Figure 5.8. Kaplan Meier survival curves for ATM protein expression in the Nottingham Cohort, in patients with low PTEN (left) and high PTEN (right) protein expression.

ATM protein expression level has been dichotomised into low/high expression using Xtile. Kaplan Meier survival analysis was performed and an association between expression level and overall survival was sought using the Log Rank test. Expression level of ATM protein was not found to be significantly associated with survival in patients with low PTEN protein expression (p=0.371) but was significantly associated with survival in patients with high PTEN protein expression (p=0.022).

5.2.2 Exploring the potential of targeting PTEN proficient and deficient glioblastoma cell lines with an ATM inhibitor

The aim of this chapter was to explore the potential of targeting PTEN deficient glioblastoma cell lines with an ATM inhibitor to induce synthetic lethality. However, the results discussed previously show an association between high ATM expression and poor survival in patients with high PTEN expression, but not in patients with low PTEN expression. This unexpected finding is consistent across both the large TCGA gene expression dataset and the smaller Nottingham Cohort for protein expression. Potentially, therefore, targeting PTEN proficient glioblastoma with an ATM inhibitor may induce synthetic lethality. To explore this further, an in depth investigation of the therapeutic potential of targeting ATM in GBM cell lines, and the influence of PTEN status, was performed.

Firstly, a panel of four glioblastoma cell lines was screened for PTEN status. Expression of ATM was confirmed prior to survival assays using an ATM inhibitor (KU55933). Functional studies were then performed to ascertain the functional consequences of ATM inhibitor treatment on the cell lines.

5.2.2.1 Western blot analysis demonstrates reduced PTEN protein expression in A172 and U87MG cell lines

Western blot analysis was used to determine PTEN protein expression using an antibody directed against full length PTEN. PTEN expression was calculated relative to the LN18 cell line. Results for all four cell lines are shown in Figure 5.9. No significant difference in PTEN expression is observed between the LN18 and LN229 cell lines; these cell lines will be referred to as 'PTEN proficient'. Reduced levels of PTEN expression are demonstrated in both the A172 and U87MG cell lines as compared with the LN18 ($p<0.05$) cell line. The A172 and U87MG cell lines will therefore be referred to as 'PTEN deficient'.

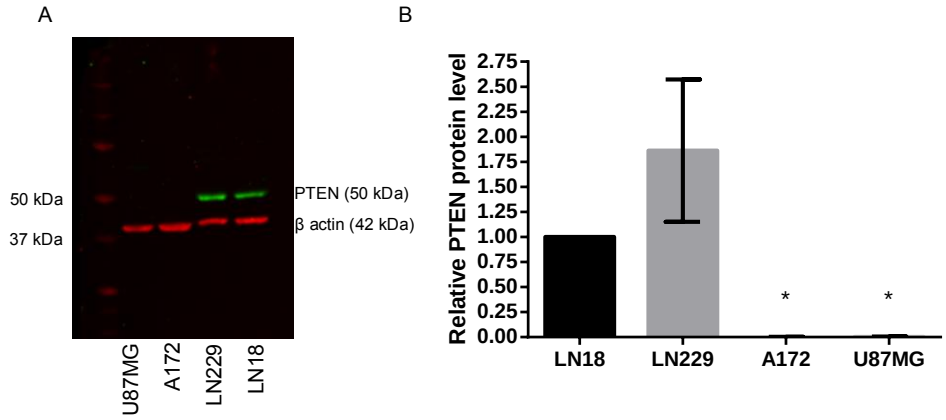


Figure 5.9. Western blot for PTEN protein expression in glioblastoma cell lines.

A. A representative Western blot for PTEN protein expression in four glioblastoma cell lines. Beta actin has been used as a housekeeper (molecular weight=42 kDa). A single PTEN band of approximately 50 kDa is demonstrated for the LN229 and LN18 cell lines. B. Graphical representation of relative PTEN protein expression in four glioblastoma cell lines. PTEN expression was first normalised to beta actin expression for each cell line and then normalised expression stated relative to expression in the LN18 cell line. Relative PTEN expression was significantly higher in the LN18 cell line as compared to the A172 and U87MG cell lines. *$p<0.05$, calculated by one-way ANOVA. Data is presented as mean ± standard error of the mean (SEM) (n=3).

5.2.2.2 Endpoint PCR analysis also confirms reduced PTEN mRNA expression in the A172 and U87MG cell lines

PTEN mRNA expression in the four glioma cell lines was assessed by endpoint PCR analysis. All samples were deemed clear of genomic DNA contamination by the inclusion of a 'no reverse transcriptase control' (see Section 2.4.3.5.2).

5.2.2.2.1 Endpoint PCR confirms absent PTEN mRNA expression in the A172 and U87MG cell lines

In keeping with the results of Western blot analysis, PTEN mRNA expression was also absent in the A172 and U87MG glioblastoma cell lines by endpoint PCR analysis (Figure 5.10B). Beta actin was included as a housekeeper gene and is expressed to a similar level in all four glioblastoma cell lines (Figure 5.10A).

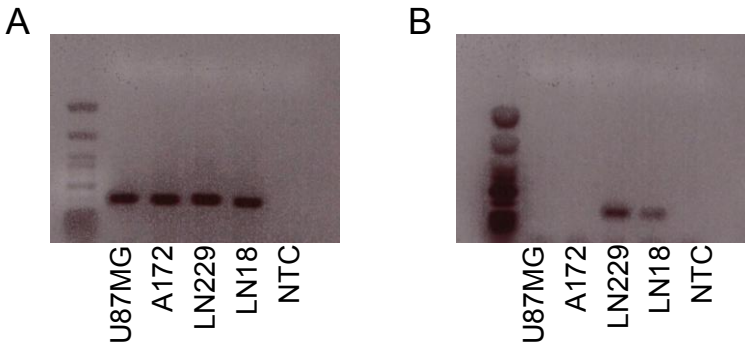


Figure 5.10. Endpoint PCR for beta actin (A) and PTEN (B) mRNA expression in glioblastoma cell lines

A. A representative endpoint PCR gel for beta actin mRNA expression in four glioblastoma cell line samples. Beta actin is expressed to a similar level in all four cell lines. B. A representative endpoint PCR gel for PTEN mRNA expression in four glioblastoma cell line samples. PTEN expression is absent in the U87MG and A172 cell lines. NTC=no template control.

5.2.2.3 Western blot analysis demonstrates that all four glioma cell lines express ATM protein

On Western blot analysis all four glioma cell lines express ATM (Figure 5.11). The highest level of ATM protein expression is seen in the LN18 cell line, which exhibits a significantly higher level of ATM expression than LN229 ($p<0.05$).

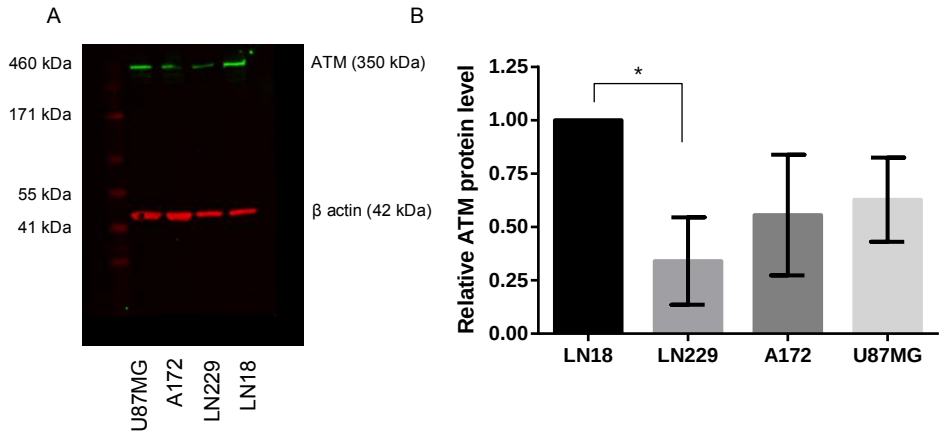


Figure 5.11. Western blot for ATM protein expression in glioblastoma cell lines.

A. Representative Western blot for ATM protein expression in four glioblastoma cell lines. Beta actin has been used as a housekeeper (molecular weight=42 kDa). A single ATM band of approximately 350 kDa is demonstrated for all four cell lines. B. Graphical representation of relative ATM protein expression in four glioblastoma cell lines. ATM expression was first normalised to beta actin expression for each cell line and then normalised expression stated relative to expression in the LN18 cell line. Relative ATM expression was significantly higher in LN18 as compared to the LN229 cell line, *$p<0.05$, calculated by one-way ANOVA. Data is presented as mean ± SEM (n=3).

5.2.2.4 ATM mRNA expression was confirmed by endpoint PCR analysis

Endpoint PCR analysis confirmed that ATM mRNA was expressed in all four glioma cell lines (Figure 5.12B). Beta actin was included as a housekeeper gene (Figure 5.12A).

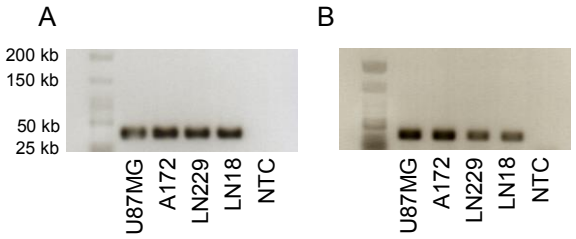


Figure 5.12. Endpoint PCR for beta actin (A) and ATM (B) mRNA expression in glioblastoma cell lines.

A. A representative endpoint PCR gel for beta actin mRNA expression in four glioblastoma cell line samples. Beta actin is expressed to a similar level in all four cell lines. B. A representative endpoint PCR gel for ATM mRNA expression in four glioblastoma cell line samples. NTC=no template control.

5.2.2.5 Survival of glioblastoma cell lines in response to an ATM inhibitor

A variety of techniques can be used to determine cell survival in response to differing concentrations of drug. A widely used method is the clonogenic survival assay. In this experiment, cells are plated out at low concentrations, allowed to adhere for at least 4 hours, or overnight, and drug added at varying concentrations. After approximately 10-14 days cells are fixed, stained with crystal violet and the colonies counted. This experiment follows the assumption that each colony arises from a single cell, but if too many cells are seeded the colonies will coalesce and be impossible to count (Munshi et al., 2005).

Prior to performing a clonogenic assay, the optimum number of cells to be seeded needs to be determined, often referred to as a plating efficiency (PE) experiment. The optimum number of cells must result in countable individual colonies after 10-14 days and the highest plating efficiency [PE (%) = (number of colonies formed/number of cells seeded) x 100] (Munshi et al., 2005).

Alternatively, colorimetric methods can be used to determine the number of viable cells and this can be used for proliferation, cytotoxicity or chemo-sensitivity assays. In this study, the CellTiter 96® AQ_{ueous} One Solution Cell Proliferation (MTS) Assay (Promega) was used to determine sensitivity to ATM inhibitor-treatment and also, as described in Section 2.4.7, cell proliferation. Cells are plated into 96 well plates, allowed to adhere and then treated with drug at a range of concentrations. On day 7 the CellTiter 96® AQ_{ueous} One Solution Reagent is added; this contains a tetrazolium compound

[3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2H-tetrazolium, inner salt; MTS] and an electron coupling reagent (phenazine ethosulfate; PES). As PES demonstrates chemical stability, this means that it can be combined with MTS to form a stable solution. This is unlike other assays that use the less stable PMS as the coupling reagent and hence PMS and MTS have to be supplied separately. In this study cells were incubated with the CellTiter 96® AQ_{ueous} One Solution Reagent for 2 hours and then absorbance recorded at 490nm on a 96-well plate reader. The absorbance recorded is a measure of formazan product and is directly proportional to the number of living cells (Promega, 2015b).

5.2.2.5.1 Glioma cell lines do not grow well at low density

Several attempts were made to perform plating efficiency experiments for the four glioma cell lines. Different numbers of cells were seeded, ranging from 100-5000 in both 6 well tissue culture plates and T25 tissue culture flasks. At densities of ≤600 cells per well, no colonies formed for any cell line. At higher densities (≥2000 cells) the U87MG cell line grew as a monolayer, while the plating efficiency for the other three cell lines ranged from 4.7 to 15.7%.

5.2.2.5.2 The LN18 cell line is most sensitive to the ATM inhibitor KU55933

Due to the inability of the GBM cell lines to grow at low density, the MTS assay was used to evaluate sensitivity to ATM inhibitor treatment rather than the clonogenic survival assay. As shown in Figure 5.13, the LN18 cell line (PTEN proficient) was most sensitive to treatment with KU55933 (an inhibitor

of ATM). Increasing concentrations of KU55933 resulted in significantly decreased survival of the LN18 cell line ($p<0.01$).

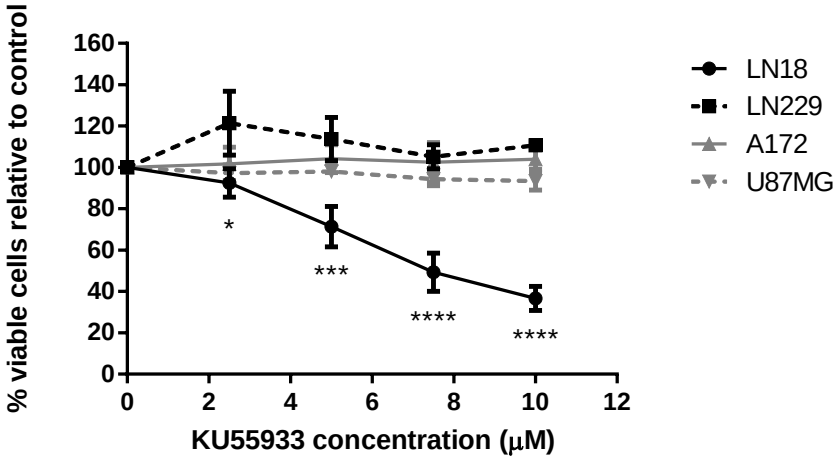


Figure 5.13. MTS assay for ATM inhibitor KU55933 in glioblastoma cell lines.

The LN18 cell line (PTEN proficient) exhibited significantly less viable cells after exposure to ATM inhibitor. * $p<0.05$, ***$p<0.001$, ****$p<0.0001$ compared to LN229 cell line, calculated by two-way ANOVA. A statistically significant difference was also demonstrated between LN229 and U87MG cell lines at 2.5μM KU55933 only ($p<0.05$). Data presented as mean ± SEM (n=3).

For additional validation of the results seen in the MTS assay with KU55933, a growth curve experiment was performed for LN18 and U87MG and a 10μM dose of KU55933. As shown in Figure 5.14 the LN18 cells treated with KU55933 grew slower than the LN18 cells not treated with KU55933. No difference was seen between the growth of the U87MG cells, treated and untreated, with KU55933.

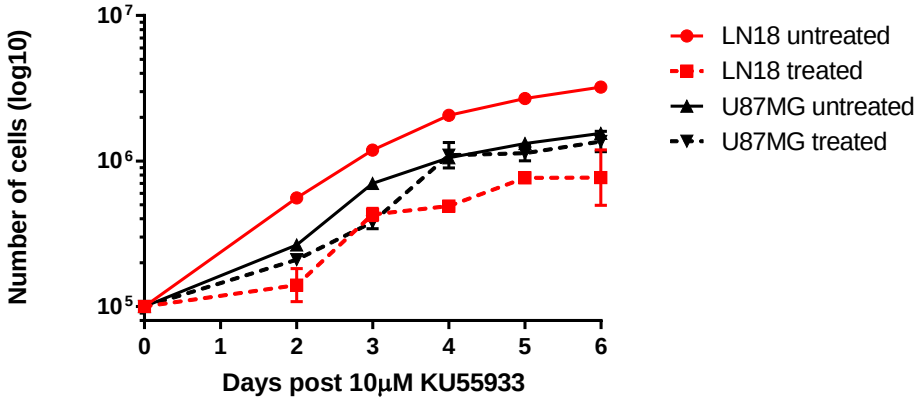


Figure 5.14. Growth assay with 10µM KU55933 in the LN18 and U87MG glioblastoma cell lines.

The LN18 cell line (PTEN proficient) exhibited slower growth after exposure to 10µM KU55933. No difference was seen between the growth of U87MG untreated and treated cells (n=1). Data presented as mean ± standard deviation (SD).

5.2.2.6 Functional consequences of ATM inhibitor treatment in glioblastoma cell lines

In order to establish the functional consequences of ATM inhibitor treatment in in glioblastoma cell lines, four assays were performed: the neutral comet assay; and flow cytometry assessment of gamma H2AX accumulation, cell cycle progression and apoptosis.

5.2.2.6.1 Background to the neutral comet assay

A comet assay is used to measure DNA damage and can either be performed under neutral or alkaline conditions. While it has previously been stated that the neutral comet assay detects only DNA double strand breaks, and the alkaline comet assay detects single strand breaks, the neutral comet assay is widely accepted to detect both single and double DNA strand breaks. The alkaline comet assay can detect other types of damage which under alkaline conditions are converted to DNA strand breaks (Olive and Banath, 2006). Neutral comet assays were used in this study as loss of both PTEN and ATM expression would be predicted to result in the accumulation of DSBs, plus potentially SSBs.

In this assay cells are harvested, mixed with low melting point molten agarose and immobilized on at CometSlide™. Samples are lysed in a salt-detergent buffer to remove cell membranes and histones from DNA. Application of an electrical current to negatively charged DNA that has undergone relaxation of supercoiling results in the migration of broken ends of DNA towards the anode. Cells with greater numbers of DNA breaks demonstrate greater levels

of migration in an electric current and can be visualised by the application of a fluorescent DNA stain. The visible effects of this procedure are referred to as comets, with cells with higher levels of DNA damage having a longer comet 'tail' (Olive and Banath, 2006).

Comets can either be scored by visual inspection by a skilled observer or using image analysis software. Commercially available software has the advantage of providing information on a variety of parameters such as the total intensity of the comet, percentage head and tail DNA, tail length and tail moment. Comet tail moment is widely reported in the scientific literature and is a way of expressing tail length and tail intensity in a single measurement. Comet tail length is inversely proportional to the size of DNA fragments, as smaller fragments migrate further in an electric current. Tail intensity is a measure of the fraction of DNA within the tail. Mean tail moment has been used throughout this thesis (Collins et al., 2008).

5.2.2.6.2 Determination of the optimal time point for the neutral comet assay

The neutral comet protocol was followed as per Section 2.4.6.1. Optimisation was performed at three time points: 6, 24 and 48 hours post drug administration. KU55933 at a concentration of 10μM was used, with MMS (100μM) as a positive control to induce DNA breaks. Each sample was plated out in duplicate. The highest levels of DNA damage were demonstrated at 6 hours post drug administration, with lower levels at 24 hours. The lowest levels of DNA damage were present at 48 hours post drug administration, suggesting that repair has occurred in the remaining live cells (Appendix 5).

For subsequent experiments, the 6 and 24 hour time points were included, with the LN18 cell line treated with 100µM MMS as a positive control.

5.2.2.6.3 Glioblastoma cell lines do not accumulate significant amounts of DNA damage in response to ATM inhibitor treatment

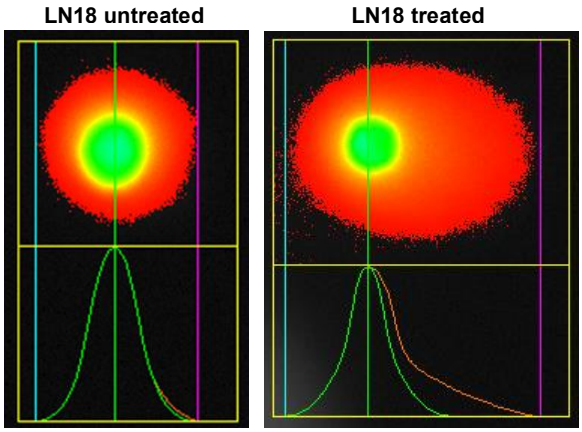


Figure 5.15. Example Comet images for LN18 cell line.

Representative Comet images are shown for the LN18 cell line untreated (left) and treated with 10µM KU55933 (right).

Examples of comets observed in the LN18 treated cell line as compared to an untreated LN18 cell are shown in Figure 5.15. Higher levels of DNA damage were seen in KU55933 treated LN18 cells as compared to untreated LN18 cells; however, no statistically significant difference was seen between mean tail moment for any cell lines (untreated vs treated with KU55933) at either 6 hours (Figure 5.16) or 24 hours (Figure 5.17) post drug administration.

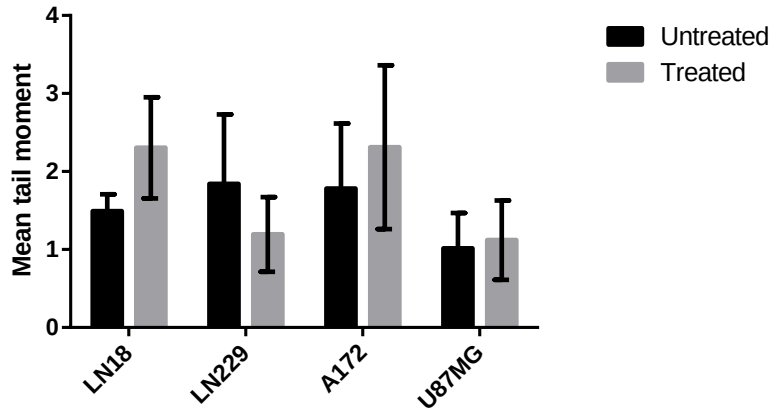


Figure 5.16. Neutral comet assay in glioblastoma cell lines 6 hours after exposure to 10μM KU55933.

No statistically significant difference was seen between the mean tail moment in untreated and treated samples for any cell line at 6 hours post drug exposure, as calculated by two-way ANOVA. Data presented as mean ± SEM (n=3).

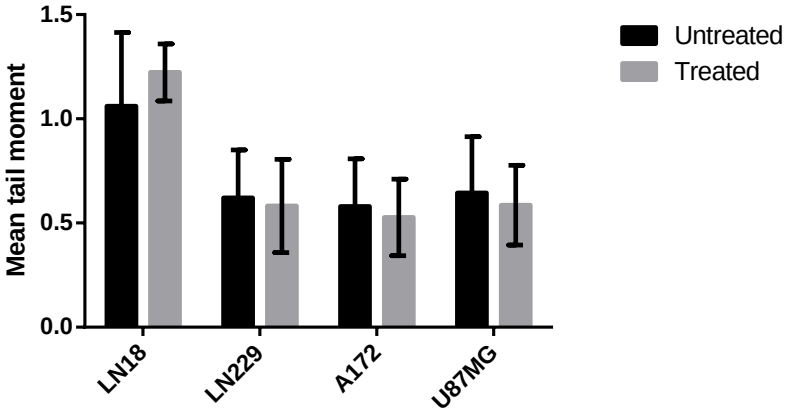


Figure 5.17. Neutral comet assay in glioblastoma cell lines 24 hours after exposure to 10μM KU55933.

No statistically significant difference was seen between the mean tail moment in untreated and treated samples for any cell line at 24 hours post drug exposure, as calculated by two-way ANOVA. Data presented as mean ± SEM (n=3).

Figure 5.18 demonstrates that LN18 cells accumulate significantly higher amounts of DNA damage in response to the alkylating agent MMS (positive control) at 6 hours post drug administration ($p<0.05$). However, after 24 hours of exposure to MMS, no statistically significant differences are seen between untreated and treated LN18 cells.

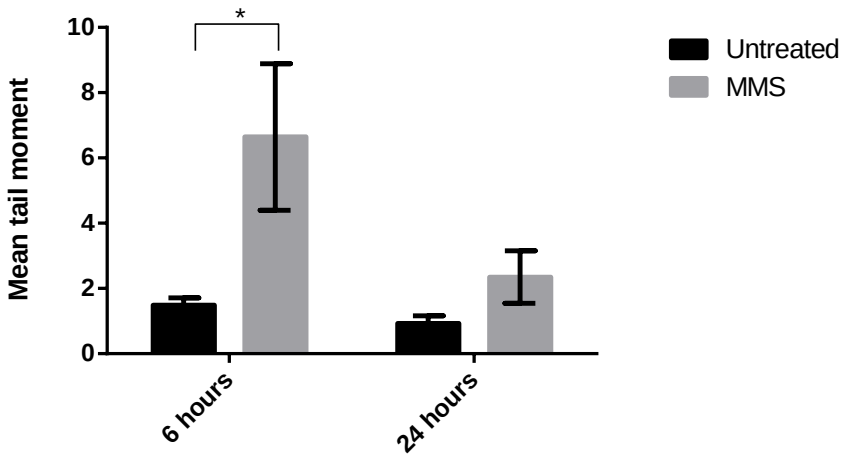


Figure 5.18. Neutral comet assay in LN18 glioblastoma cell lines 6 and 24 hours after exposure to 100µM MMS.

The LN18 cell line accumulates DNA damage after exposure to 100µM MMS as compared to untreated control at 6 hours post drug administration. *$p<0.05$, as calculated by two-way ANOVA. This damage appears to be resolving by 24 hours post drug administration. Data presented as mean ± SEM (n=3).

5.2.2.6.4 Flow cytometry

Flow cytometry has been used to assess cell cycle progression and to determine the accumulation of gamma H2AX in response to ATM inhibitor treatment using a dual staining protocol described in Section 2.4.6.2. In addition, combined Annexin V and propidium iodide staining was used to determine the apoptotic fraction of cells after treatment with an ATM inhibitor (Section 2.4.6.3).

5.2.2.6.5 Background of flow cytometry

Flow cytometry is a high throughput method that enables the sequential analysis of single cells. The technique involves the excitation of cells, commonly using a laser, and subsequent collection of the emitted light. Flow cytometry allows the collection of data on multiple parameters - including size, granularity and fluorescence – rapidly, from large populations of cells. The process involves three key steps. Firstly, the cells are labelled with a fluorophore, which are chemicals that can fluoresce and that absorb light at one wavelength and emit light at another, longer, wavelength. Secondly, the samples are processed by the flow cytometer. Cells are subjected to hydrodynamic focusing in the flow cell which results in the alignment of single cells that are then passed through a laser beam. For each cell, scattered light is collected of the same wavelength as the laser. Light collected in the forward scatter direction provides information on the cell size; the bigger the cell the greater the forward scatter. Side scatter data provides information on the structure and granularity of the cells; higher side scatter is associated with less

uniformity of cell shape and a larger amount of internal membranes. Emitted fluorescent light is also collected, but only if the fluorophore has bound to the cell. The intensity of fluorescent light is proportional to the amount of fluorophore bound to the cell and wavelength is dependent on the fluorophore used (Brown and Wittwer, 2000).

5.2.2.6.6 γH2AX analysis by flow cytometry

H2AX is a member of the H2A subfamily and accounts for 10-15% of the H2A of the core nucleosome structure. As an early event in the response to DNA DSBs, H2AX undergoes phosphorylation at serine 139 on the carboxyl-terminus and is subsequently known as γH2AX. This phosphorylation is performed by members of the PI3K family, including ATM, ATR and DNA-PKcs. H2AX phosphorylation occurs at sites flanking the DSB and results in foci that recruit DSB repair proteins, including RAD50, RAD51 and BRCA1 (Kinner et al., 2008).

Various methods have been used to detect the presence of γH2AX as a marker of DNA DSBs, including immunofluorescence and confocal microscopy. γH2AX can also be measured using a flow cytometric assay. One advantage of this assay is that it can be combined with cycle cell analysis, enabling the assessment of γH2AX levels in the various stages of cell cycle (Porcedda et al., 2008).

5.2.2.6.7 Anti-phospho-H2AX (Ser139) FITC conjugate is specific for γH2AX

In addition to binding to the desired antigen, fluorescently labelled antibodies can also bind to Fc receptors on cells, participate in non-specific protein-protein interactions, or the fluorophore can bind directly to the cell. An isotype control is a monoclonal antibody of the same immunoglobulin class or subclass as the fluorescently labelled specific antibody. The isotype control is itself labelled with the same fluorophore as the specific antibody but is raised against an irrelevant antigen. Samples stained with the isotype control and the unstained cells can be compared, and if there is a large difference in staining then this indicates non-specific staining has occurred.

An isotype control was included for each cell line in all γH2AX flow cytometry experiments using a normal mouse IgG FITC conjugate (Merck Millipore). No difference was seen between unstained cells and normal mouse IgG FITC conjugate isotype controls for all four cell lines (see example for LN18 cell line in Figure 5.19).

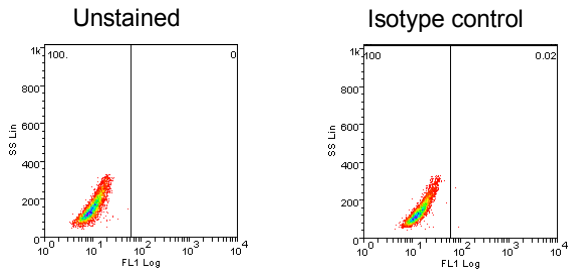


Figure 5.19. Example of isotype control staining in the LN18 glioblastoma cell line by flow cytometry.

No difference was seen between the staining pattern for unstained cells and normal mouse IgG FITC conjugate isotype control for all four cell lines (example for the LN18 cell line shown above).

5.2.2.6.8 Glioblastoma cell lines do not accumulate DNA DSBs in response to ATM inhibitor treatment

Flow cytometry analysis was performed to determine the levels of γH2AX for all four glioblastoma cell lines; both untreated and treated with 10μM KU55933. Representative dot plots for each condition are shown in Figure 5.21. No significant difference in the percentage of cells staining positive for γH2AX was demonstrated between untreated and KU55933 treated glioma cell lines after 24 hours of drug exposure (Figure 5.20). Analysis of γH2AX by stage of cell cycle was therefore not performed.

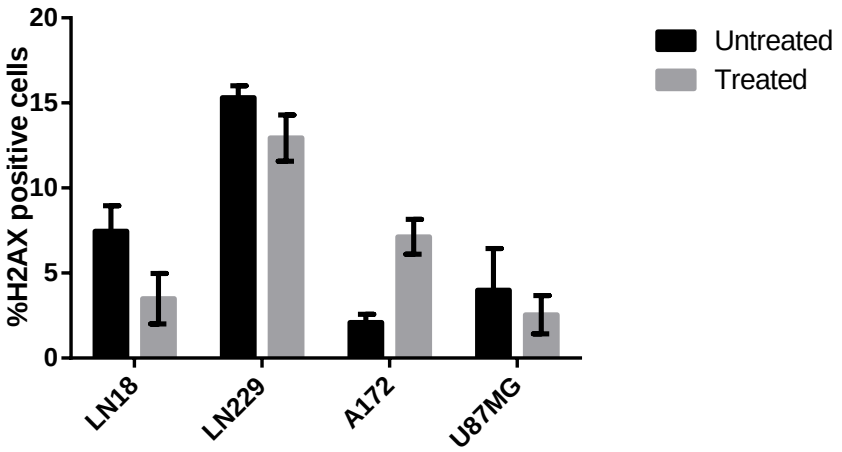


Figure 5.20. Percentage of cells staining positive for γH2AX in cells both untreated and treated with 10 μM KU55933.

No statistically significant difference was seen between the percentage of cells that stain positive for γH2AX in untreated and treated samples for any cell line at 24 hours post drug exposure, as calculated by two-way ANOVA. Data presented as mean ± SEM (n=2).

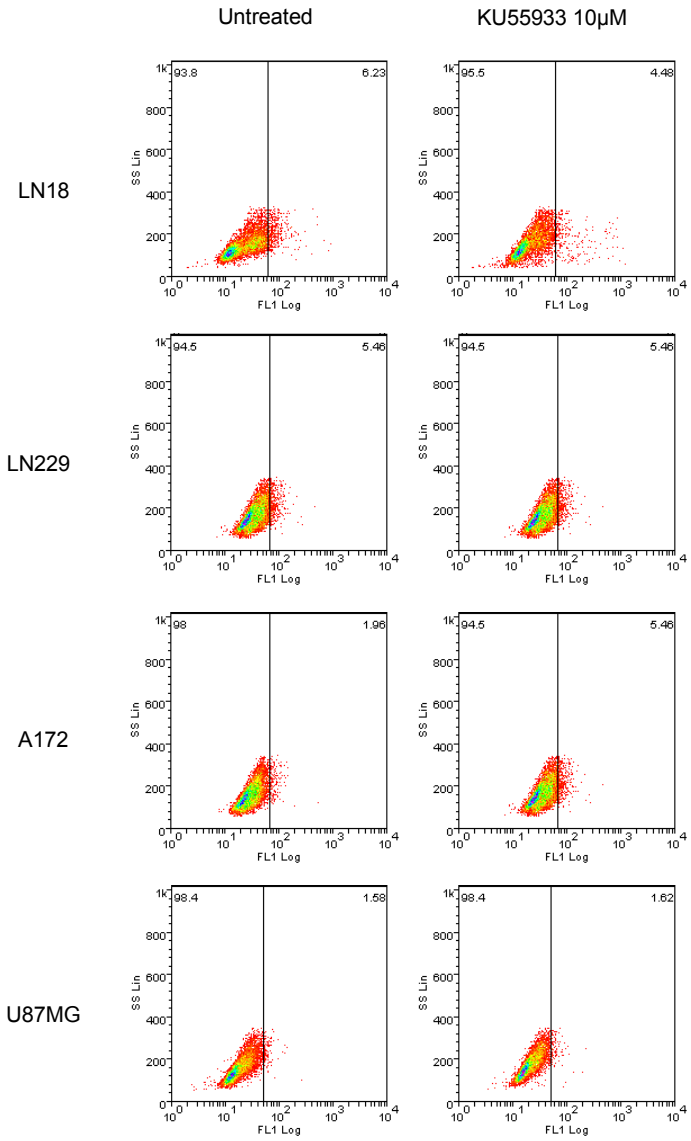


Figure 5.21. Representative dotplots for γH2AX staining in glioblastoma cell lines untreated and treated with 10 μM KU55933.

Intensity of γH2AX staining is shown on the horizontal axis as FL1 log. Cells to the right of the vertical line are classified as staining positive for γH2AX. No significant difference is seen in the numbers of cells that stain positive for γH2AX in the untreated and treated samples for all four cell lines. Percentages of cells that stain positive and negative for γH2AX are shown in the top right (positive) and left (negative) corners of each dotplot.

5.2.2.6.9 Cell cycle analysis by propidium iodide flow cytometry

Cell cycle analysis can be performed by quantification of the DNA content of cells. Propidium iodide is a DNA binding fluorophore that binds to DNA in a stoichiometric manner and therefore the DNA binding is proportional to the DNA content. As a consequence, cells in S phase bind more PI and fluoresce more intensely than cells in G1. Cells in G2 phase of cell cycle contain double the DNA content of cells in G1 and therefore should fluoresce twice as brightly. In order for PI to bind to DNA, the cell's nuclear membrane needs to be permeabilised. 70% ethanol can be used to both fix and permeabilise the cells.

5.2.2.6.10 LN18 glioblastoma cell line accumulates cells in G1 phase in response to ATM inhibitor treatment

Propidium iodide cell cycle analysis was performed to determine the proportion of cells present in each stage of the cell cycle for all four glioblastoma cell lines; both untreated and treated with 10μM KU55933. Representative cell cycle histograms are shown in Figure 5.22. The LN18 cell line has significantly more cells in the G1 phase of cell cycle after treatment with KU55933 and significantly less cells in S phase ($p<0.01$), implying that ATM inhibitor treatment is inducing a G1/S checkpoint arrest (Figure 5.23). None of the other three glioblastoma cell lines show a significant difference in the distribution of cells across the phases of the cell cycle in response to ATM inhibitor treatment.

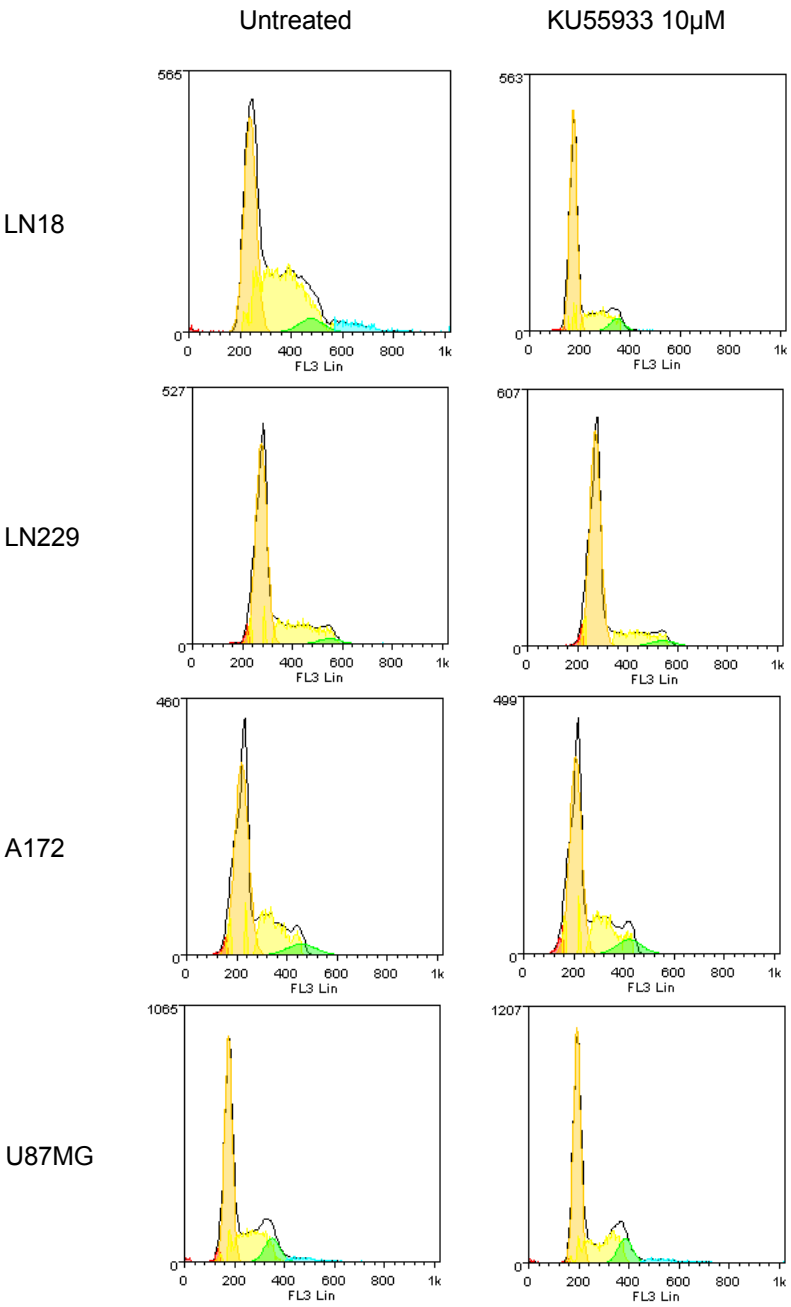


Figure 5.22. Representative cell cycle plots for glioblastoma cell lines untreated and treated with 10 µM KU55933.

Intensity of propidium iodide staining is represented by FL3 Lin on the x axis. The y axis displays the number of cells that stain at a particular intensity. Untreated, the LN18 cell line demonstrates an accumulation of cells in S phase, whereas after treatment with KU55933 cells accumulate in G1 phase. Similar cell cycle profiles are demonstrated for both untreated and treated LN229, U87MG and A172 cell lines. Orange=G1 phase; yellow=S phase, green=G2/M phase.

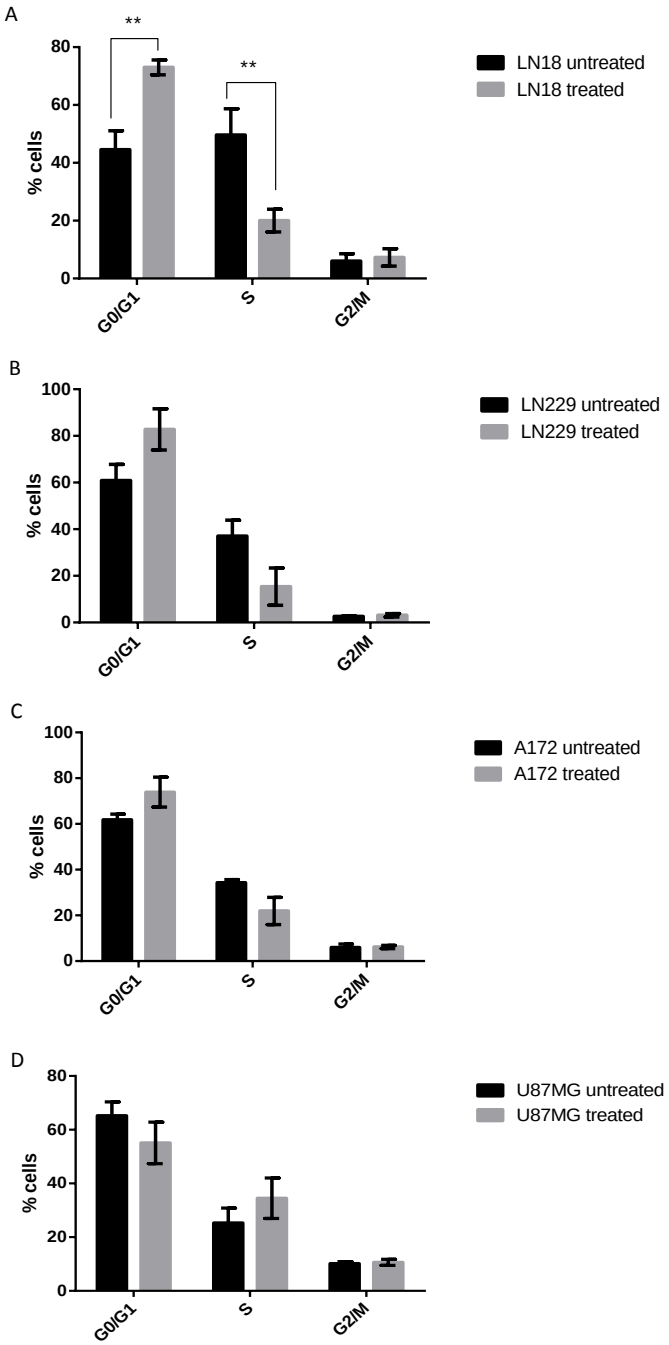


Figure 5.23. Graphical representation of the percentage of cells in each phase of cell cycle in glioblastoma cell lines both untreated and treated with 10 µM KU55933.

The LN18 (A) cell line has significantly more cells in G1 phase of cell cycle after treatment with KU55933 (**$p<0.01$) and significantly fewer cells in S phase (**$p<0.01$), as determined by two-way ANOVA. No statistically significant differences are demonstrated between the distribution of cells throughout cell cycle in untreated and treated LN229 (B), A172 (C) and U87MG (D) cell lines. Data presented as mean ± SEM (n=3).

5.2.2.6.11 Apoptosis detection by Annexin V-FITC flow cytometry

Cells generally die by two main mechanisms, apoptosis and necrosis. Apoptosis is often referred to as programmed cell death and involves cell dehydration and shrinkage, chromatin condensation and nuclear fragmentation. Apoptotic bodies are formed but the cell membrane remains intact. *In vivo*, apoptotic cells undergo rapid phagocytosis and the process is not associated with inflammation.

In a normal cell, phosphatidylserine residues are located on the internal leaflet of the plasma membrane. When a cell undergoes apoptosis these residues are 'flipped' to the cell surface and can bind Annexin V in the presence of calcium (Figure 5.24). Annexin V-FITC can be used to identify cells that are expressing Annexin V on the cell surface and are therefore undergoing apoptosis.

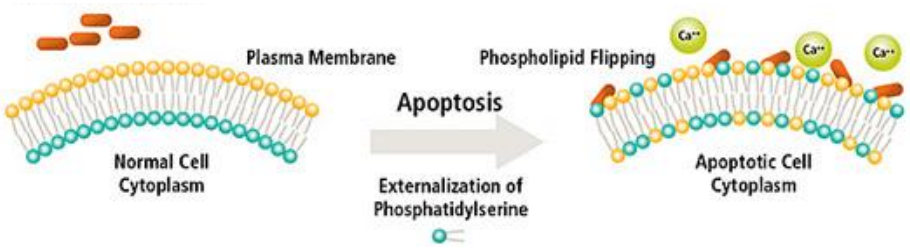


Figure 5.24. Diagrammatic representation of apoptosis.

In the normal cellular environment phosphatidylserine residues are located on the internal side of the plasma membrane. During the process of apoptosis these residues flip to the external surface of the cell membrane where, in the presence of calcium, they can bind Annexin V. In this image, Annexin V is represented by the orange lozenges. Image modified from (Promega, 2015a)

Necrosis is associated with both mitochondrial and cell swelling, followed by membrane rupture with release of the intracellular contents. Unlike apoptosis, necrosis is associated with inflammation. Propidium iodide (PI) is a fluorophore that cannot cross the plasma membrane and therefore does not stain live or apoptotic cells but does stain those that have undergone necrosis.

5.2.2.6.12 LN18 glioma cell line undergoes apoptosis in response to ATM inhibitor treatment

Flow cytometry assessment of the marker Annexin V, in combination with propidium iodide, was used to determine whether KU55933 induces apoptotic cell death in glioblastoma cell lines. Significantly increased rates of apoptosis were demonstrated in the LN18 (PTEN proficient) glioblastoma cell line in response to 24 hours exposure to 10μM KU55933 as compared to the untreated control ($p<0.001$). The LN18 cell line also exhibited the highest baseline rate of apoptosis of the four cell lines. No significant differences in apoptosis were demonstrated between the treated and untreated samples for the A172, U87MG or LN229 cell lines (Figure 5.24 and Figure 5.25).

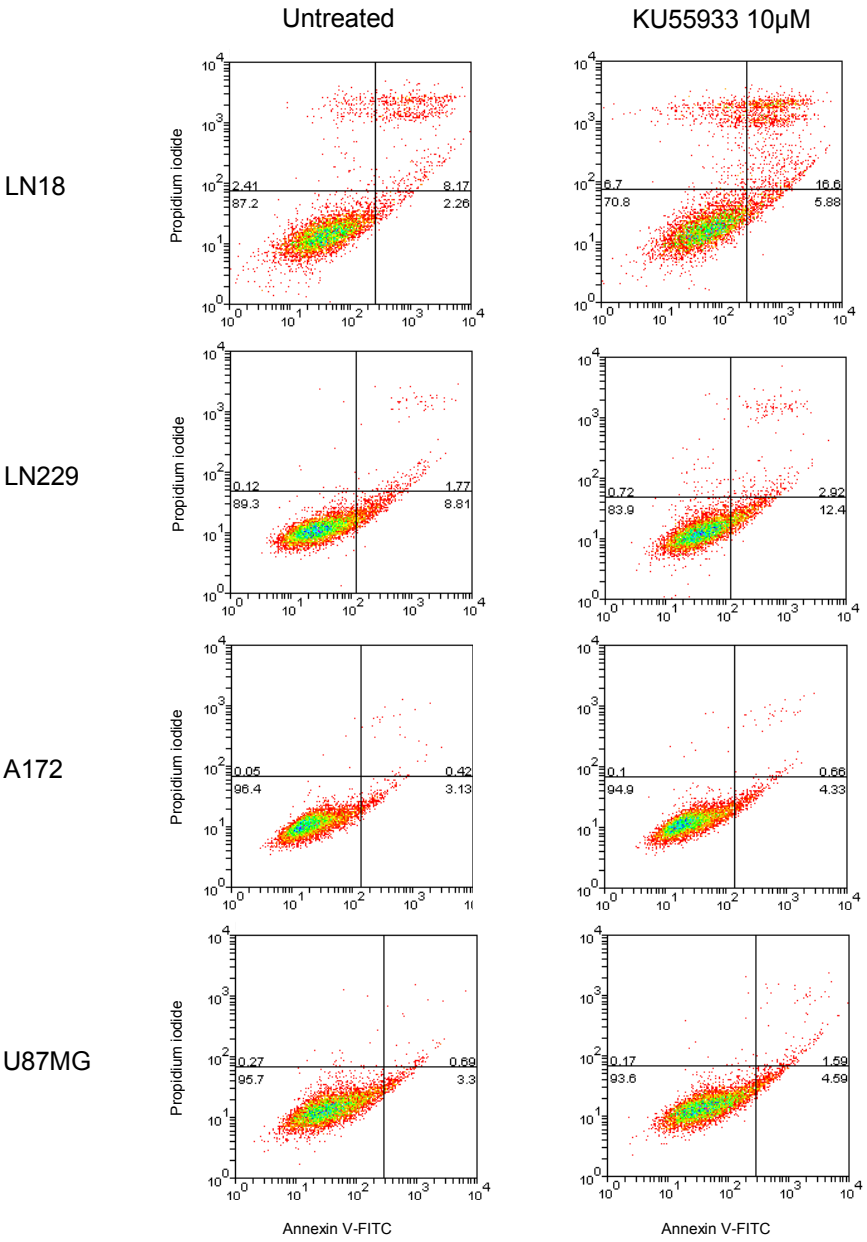


Figure 5.25. Dotplots for Annexin V-FITC and propidium iodide combination flow cytometry for glioblastoma cell lines untreated and treated with 10μM KU55933.

The greatest increase in percentage of apoptotic cells between untreated and treated samples is seen in the LN18 cell line. Apoptotic cells are either Annexin V positive/propidium iodide positive (considered late apoptotic, top right quadrant) or Annexin V positive/propidium iodide negative (bottom right quadrant).

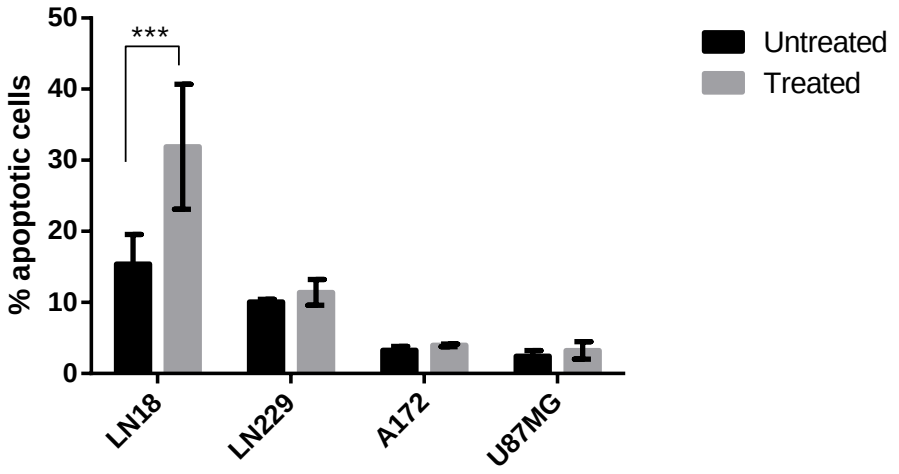


Figure 5.26. Percentage of apoptotic cells, as detected by Annexin V-FITC staining, for glioblastoma cell lines untreated and treated with 10µM KU55933.

A statistically significant increase in the percentage of apoptotic cells is demonstrated in LN18 treated cells as compared to LN18 untreated cells. ***$p<0.001$, calculated by two-way ANOVA. The highest baseline level of apoptosis is also seen in the LN18 cell line. No significant difference is seen between the percentage of apoptotic cells in untreated and treated samples for the LN229, A172 and U87MG cell lines. Data presented as mean ± SEM (n=3).

5.2.2.7 Potential explanations for the selective sensitivity of the LN18 cell line to ATM inhibitor treatment

Several potential explanations as to why the LN18 cell line was more sensitive to ATM inhibitor treatment have been explored in more detail. Firstly, differences in DNA repair gene expression between the four cell lines have been explored. Secondly, expression of p53 and PI3K family members has been evaluated as earlier studies have shown these proteins may influence response to ATM inhibitor treatment. Finally, knock-down of *PTEN* using doxycycline-inducible cell lines models has been performed to establish the role of PTEN expression on ATM sensitivity in both the LN18 and LN229 cell lines. This is described in more detail in Section 5.2.3.

5.2.2.7.1 RT^2 Profiler DNA repair PCR Array

The RT^2 Profiler DNA repair PCR Array was used to screen the four glioblastoma cell lines to identify DNA repair genes whose expression is potentially influenced by PTEN status. This 96-well array is a commercially available high-throughput screen that enables the expression of 84 DNA repair genes to be simultaneously assessed. In addition, the array measures the expression of five housekeeper genes (*ACTB*, *B2M*, *GAPDH*, *HPRT1* and *RPLPO*) and incorporates internal controls for reverse transcription, PCR efficiency and genomic DNA contamination. A list of the 84 DNA repair genes evaluated by the RT^2 Profiler DNA repair PCR Array is included in Appendix 6.

This technique involves the extraction of RNA from individual cell lines. RNA is then reverse transcribed to cDNA. The cDNA is combined with a PCR Mastermix that contains a mix of deoxyribonucleotides, a thermostable DNA polymerase and a SYBR Green probe. This cDNA containing Mastermix is then plated onto the RT^2 Profiler DNA repair PCR Array plate, where each well contains a gene-specific primer set. The plate is then subjected to thermal cycling which results in amplification of the double-stranded DNA (dsDNA) component. Fluorescent intensity increases with the amount of dsDNA, until this reaches a pre-defined threshold known as the C_T value (or threshold cycle). The PCR cycle at which this occurs is dependent on the starting concentration of RNA specific to the individual primer. The $\Delta\Delta C_T$ analysis method is commonly used to normalise the C_T of the gene of interest against the C_T of a housekeeper gene. This enables gene expression to be compared between different samples or cell lines.

An amplification plot demonstrating the accumulation of fluorescent signal over time is shown in Figure 5.27 and demonstrates the increase in signal with increasing input of cDNA. As cDNA input quantity increases, the C_T value decreases, indicating that the fluorescence intensity threshold is reached after fewer cycles of PCR. A background, or baseline, level of fluorescence is always present and is automatically determined by the PCR cycler. During the exponential phase, theoretically each cycle of PCR results in doubling of the dsDNA product, and therefore the fluorescent signal. This is followed by a linear phase whereby consumption of the PCR reagents occurs, gradually

reducing PCR efficiency, and ultimately resulting in the cessation of amplification (plateau phase).

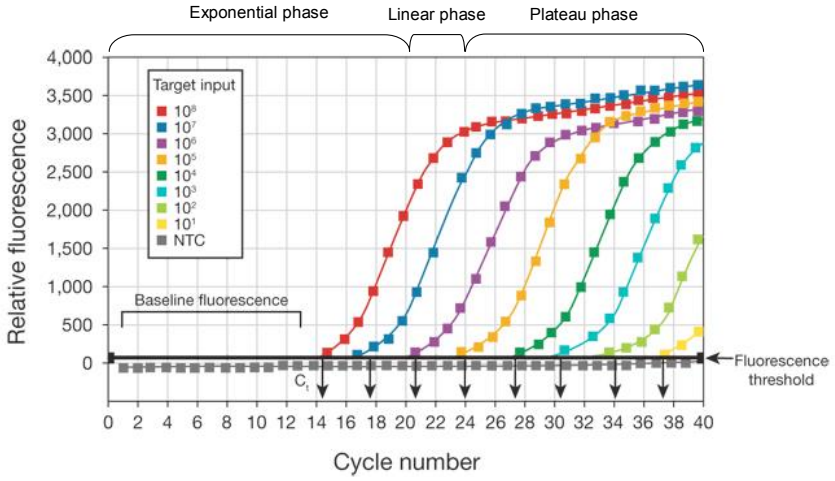


Figure 5.27. Example of a real-time PCR amplification plot.

The three phases of the PCR reaction are shown at the top of the image. During the exponential phase each PCR cycle theoretically results in a doubling of the dsDNA product. With higher input amounts of cDNA greater fluorescence intensity is observed at a lower PCR cycle number. During the linear phase consumption of PCR products occurs and PCR efficiency decreases. Amplification ceases in the plateau phase. Image modified from (LifeTechnologies, 2012).

5.2.2.7.2 Post-amplification melting curve analysis can be used to assess for primer specificity

Melting curve analysis was performed for each cell line for each PCR analysis. Example melting curves for all four glioma cell lines are shown in Figure 5.28. A double melting peak was seen for the LN18 cell line on the second PCR run (Figure 5.28 bottom) for the DMC1 primer; this primer was therefore excluded from subsequent analyses.

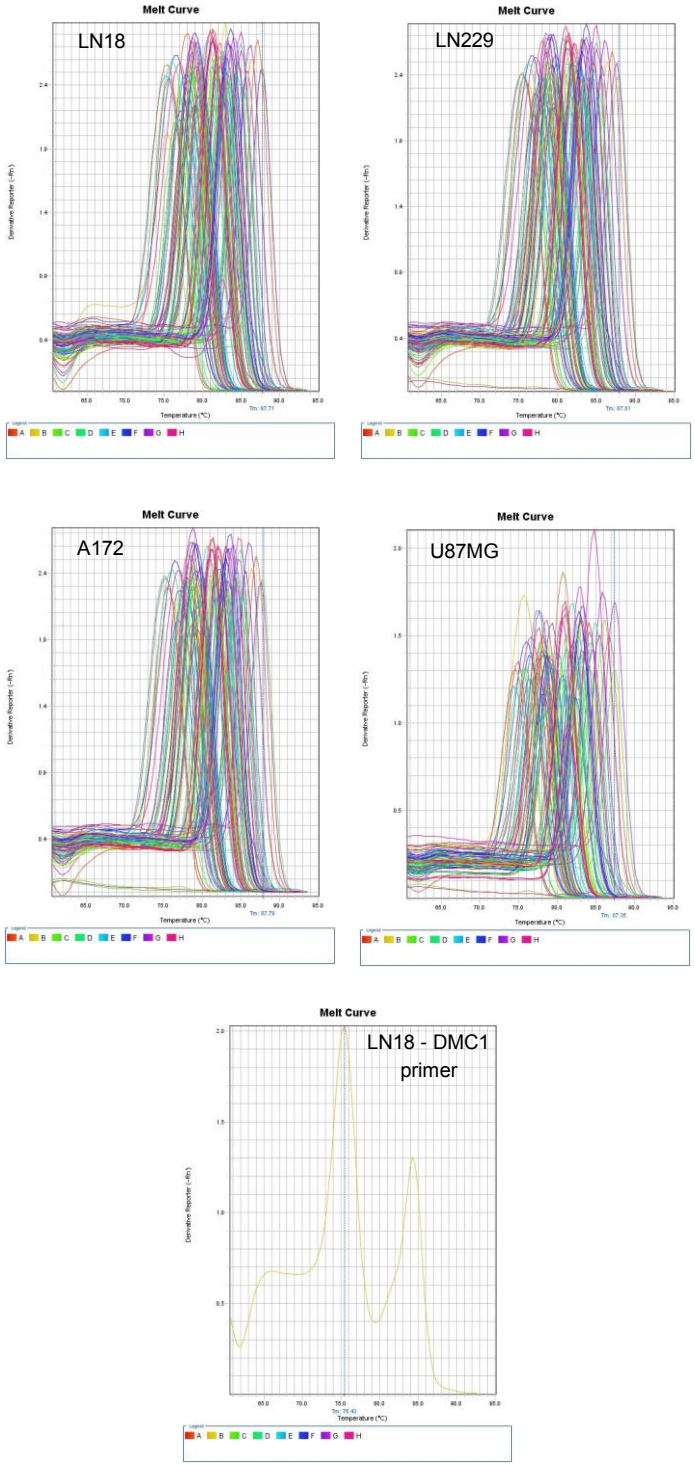


Figure 5.28. Melting curves for the four glioblastoma cell lines, plus melting curve for the non-specific DMC1 primer in the LN18 cell line.

A single melting curve is demonstrated for all primers in the LN18 (top left), LN229 (top right), A172 (central left) and U87MG (central right) cell lines, with the exception of the DMC1 primer in the LN18 cell line (bottom). A double peak is seen for this primer and therefore this has been excluded from all subsequent analyses in all four cell lines.

5.2.2.7.3 PTEN deficiency is associated with reduced DNA repair gene expression

C_T values for each RT-PCR profiler experiment for each cell line (n=2) were amalgamated into a single Excel spreadsheet and uploaded into the Qiagen data analysis software. Beta actin was selected as the house keeper gene with the least variation in C_T values across the samples and repeat experiments from a panel of five commonly used housekeeper genes. Target gene expression was then normalised against beta actin expression using the $\Delta\Delta C_T$ method, allowing gene expression to be compared between different samples. In this study DNA repair genes that demonstrate a greater than 2-fold expression difference between cell lines were identified, as described in Section 2.4.3.6.1.

DNA repair genes that are up- or down-regulated in LN18 as compared to U87MG are shown in Figure 5.29. Five DNA repair genes were identified that are under-expressed in the LN18 cell line as compared with U87MG: ERCC1, ERCC5, LIG4, NEIL2 and ATR. A total of 18 genes were up-regulated in LN18 as compared to U87MG; the greatest level of 'up-regulation' was seen for MGMT. All major DNA repair pathways are represented in the up-regulated genes. A list of all the up- and down-regulated genes (and the fold changes) in LN18 as compared to U87MG is shown in Appendix 7 and Appendix 8.

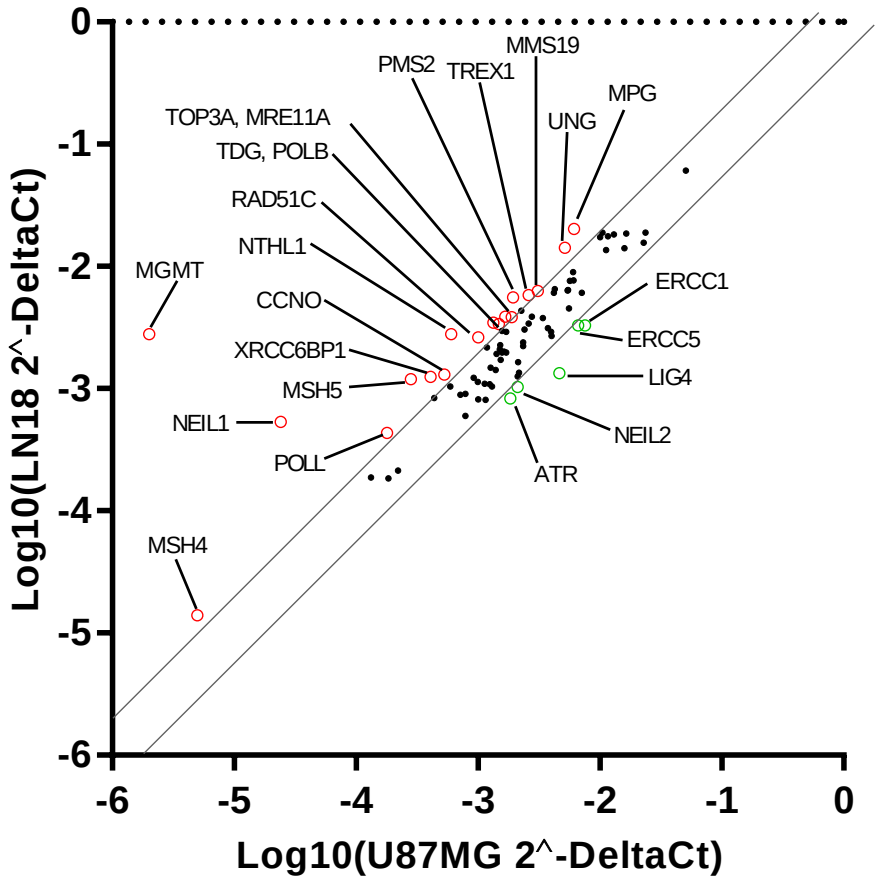


Figure 5.29. Scatterplot comparing DNA repair gene expression in U87MG (PTEN deficient) (x axis) against LN18 (PTEN proficient) (y axis)glioblastoma cell lines.

The area outside of the diagonal grey lines represents DNA repair genes with a greater than two-fold difference in expression between the two cell lines (n=2). Green circles represent genes that are under-expressed in LN18 as compared to U87MG. Red circles represent genes that are over-expressed in LN18 as compared to U87MG.

DNA repair genes that are up- or down-regulated in LN18 as compared to A172 are shown in Figure 5.30. Two DNA repair genes were identified that are under-expressed in the LN18 cell line as compared with A172, these are DDB2 and LIG4. A total of 31 genes were up-regulated in LN18 as compared to A172; the greatest level of up-regulation was seen for MGMT. All major DNA repair pathways are represented in the up-regulated genes. A list of all

the up- and down-regulated genes (and the fold changes) in LN18 as compared to A172 is shown in Appendix 9 and Appendix 10.

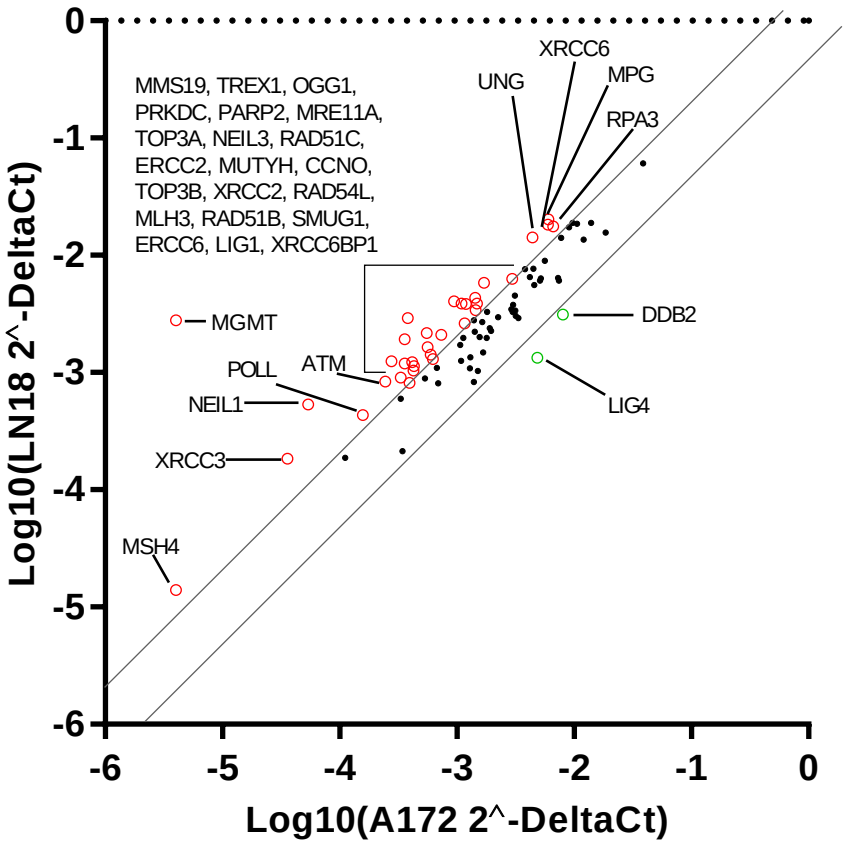


Figure 5.30. Scatterplot comparing DNA repair gene expression in A172 (PTEN deficient) (x axis) against LN18 (PTEN proficient) (y axis) glioblastoma cell lines.

The area outside of the diagonal grey lines represents DNA repair genes with a greater than two-fold difference in expression between the two cell lines (n=2). Green circles represent genes that are under-expressed in LN18 as compared to A172. Red circles represent genes that are over-expressed in LN18 as compared to A172.

DNA repair genes that are up- or down-regulated in LN18 as compared to LN229 are shown in Figure 5.31. Two DNA repair genes were identified that are under-expressed in the LN18 cell line as compared with LN229, these are

DDB2 and EXO1. A total of 40 genes were up-regulated in LN18 as compared to LN229; the greatest level of up-regulation was seen for MGMT. All major DNA repair pathways are represented in the up-regulated genes. A list of all the up- and down-regulated genes (and the fold changes) in LN18 as compared to LN229 is shown in Appendix 11 and Appendix 12.

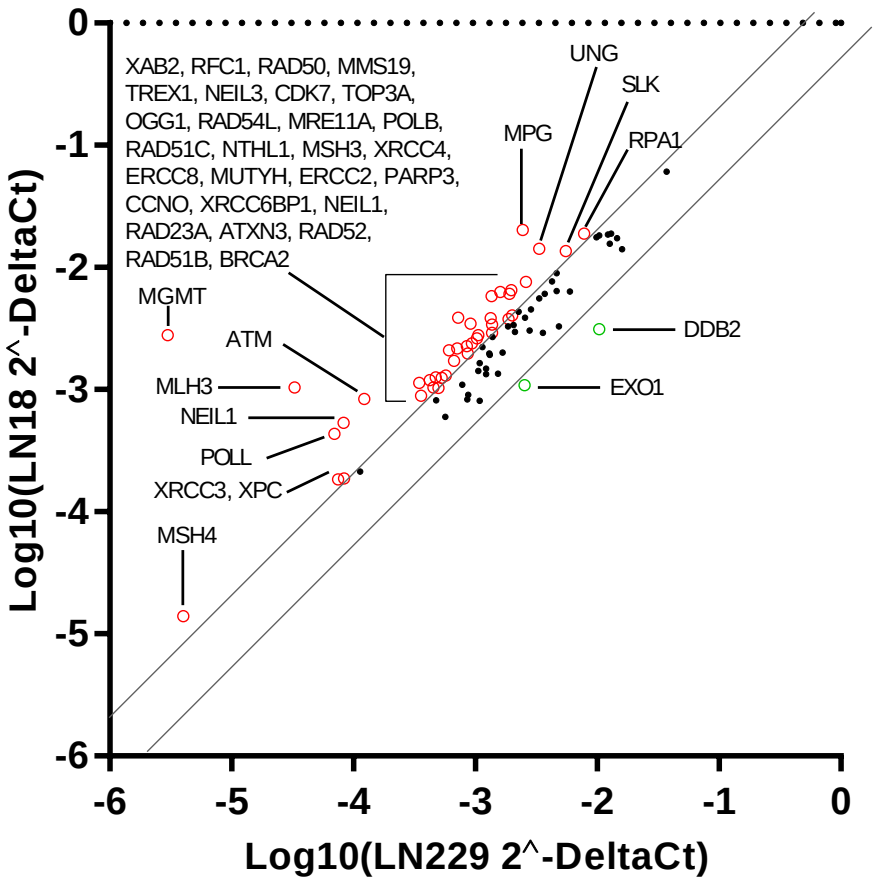


Figure 5.31. Scatterplot comparing DNA repair gene expression in LN229 (PTEN proficient) (x axis) against LN18 (PTEN proficient) (y axis) glioblastoma cell lines.

The area outside of the diagonal grey lines represents DNA repair genes with a greater than two-fold difference in expression between the two cell lines (n=2). Green circles represent genes that are under-expressed in LN18 as compared to LN229. Red circles represent genes that are over-expressed in LN18 as compared to LN229.

5.2.2.7.4 DNA repair genes commonly up-regulated in the LN18 cell line are identified

In an attempt to explain the sensitivity of the LN18 cell line to ATM inhibitor treatment, genes that were up-regulated in LN18 as compared to all three non-sensitive cell lines were identified. A total of fourteen genes were at least two-fold up-regulated in the LN18 cell line as compared to A172 and U87MG and LN229 cell lines. A list of these genes and the DNA repair pathway to which they belong is shown in Table 5.4.

Table 5.4. DNA repair genes at least two-fold up-regulated in LN18 as compared to A172 and U87MG and LN229.

	A172	**U87MG**	**LN229**	**Pathway**
CCNO	2.0907	2.4674	2.2768	BER
MGMT	724.2468*	1186.6372*	1056.748*	Direct repair
MMS19	2.1261	2.0405	3.9205	NER
MPG	3.3571	3.3185	8.2866	BER
MRE11A	3.5364	2.3116	5.3602	DSB repair
MSH4	3.7505	2.8092	3.717#	MMR
MSH5	3.3482#	4.2476#	2.8489	MMR
NEIL1	9.951*	22.6796*	6.4924*	BER
POLL	2.7642	2.4225	6.2276*	NER
RAD51C	2.2586	2.6043	2.5482	DSB repair
TOP3A	2.234	2.0497	2.8934	DSB repair
TREX1	3.4319	2.2489	4.2843	MMR
UNG	3.2228	2.7471	4.258	BER
XRCC6BP1	4.4856	3.0372	2.3554	DSB repair

The fold change for each gene (and cell line) is shown as compared to the LN18 cell line; the DNA repair pathway to which the gene belongs is also shown. #This gene's CT value is relatively high (>30) meaning that its expression level is low, in both control and test samples. The fold change result may therefore be subject to greater variation. *This gene's average CT value is relatively high (>30) in one sample and reasonably low in the other sample (<30). This means that the gene's expression is relatively low in one sample and reasonably detected in the other. This suggests that the actual fold-change is at least as high as reported. BER=base excision repair; NER=nucleotide excision repair; DSB repair=double strand break repair; MMR=mismatch repair.

5.2.2.7.5 No DNA repair genes are identified that are commonly down-regulated in the LN18 cell line

Analysis of the DNA repair genes at least two-fold down-regulated in LN18 as compared to A172, U87MG and LN229 reveals than no genes are down-regulated in LN18 as compared to all three of the other cell lines. Ligase 4 is down-regulated in LN18 as compared to both U87MG (-3.463 fold) and A172 (-3.6074 fold). DDB2 is down-regulated in LN18 as compared to both A172 (-2.5584 fold) and LN229 (-3.3392 fold).

5.2.2.7.6 Determination of p53 protein expression by Western blot analysis

In view of the previously described association between p53 status and sensitivity to ATM inhibition (Jiang et al., 2009), Western blot analysis was performed to assess p53 protein status in the panel of glioblastoma cell lines. A mouse monoclonal antibody was used that detects both wild-type and mutant p53. As shown in Figure 5.32, the highest levels of p53 protein expression are seen in the LN18 cell line. LN18 has significantly higher levels of p53 protein expression than LN229 ($p<0.001$), A172 ($p<0.0001$) and U87MG ($p<0.0001$).

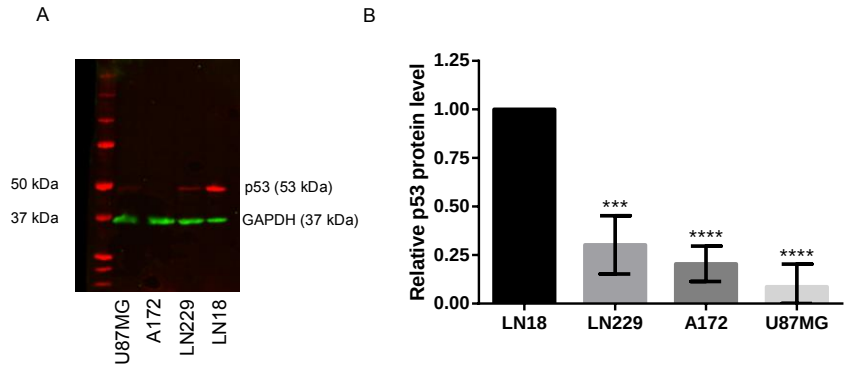


Figure 5.32. Western blot for p53 protein expression in glioblastoma cell lines.

A. A representative Western blot for p53 protein expression in four glioblastoma cell lines. GAPDH has been used as a housekeeper (molecular weight=37 kDa). A single p53 band of approximately 53 kDa is demonstrated when expressed. B. Graphical representation of relative p53 protein expression in four glioblastoma cell lines. p53 expression was first normalised to GAPDH expression for each cell line and then normalised expression stated relative to expression in the LN18 cell line. Relative p53 expression was significantly higher in LN18 as compared to LN229, A172 and U87MG cell lines. ***$p<0.001$, ****$p<0.0001$, calculated by one-way ANOVA. Data presented as mean ± SEM (n=3).

5.2.2.7.7 Determination of PI3K protein expression by Western blot analysis

PI3K protein expression was assessed by Western blot analysis using a mouse monoclonal antibody against the p85 regulatory subunit of PI3K. A representative Western blot of PI3K (p85) protein expression is shown in Figure 5.33. The highest levels of PI3K (p85) expression are seen in the U87MG cell line, which has significantly higher expression levels than A172 ($p<0.01$), LN229 ($p<0.01$) and LN18 ($p<0.001$). Expression of PI3K (p85) is essentially absent in the LN18 cell line.

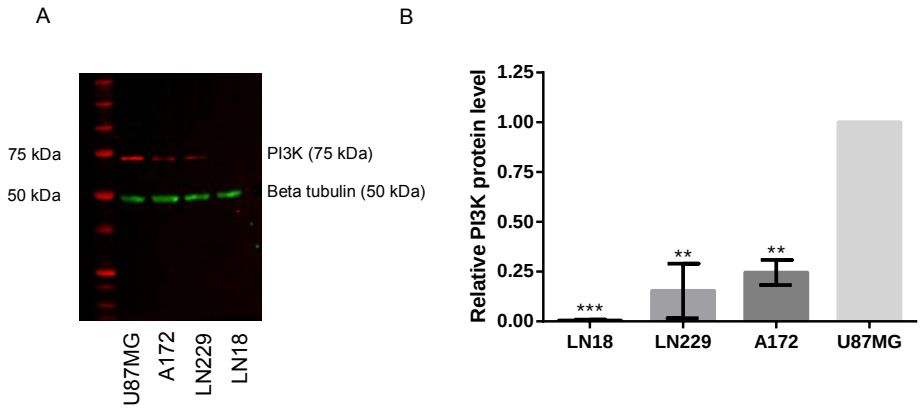


Figure 5.33. Western blot for PI3K (p85 subunit) protein expression in glioblastoma cell lines.

A. A representative Western blot for PI3K (p85 subunit) protein expression in four glioblastoma cell lines. Beta tubulin has been used as a housekeeper (molecular weight=50 kDa). A single p85 band of approximately 75 kDa is demonstrated, when expressed. B. Graphical representation of relative PI3K (p85 subunit) protein expression in four glioblastoma cell lines. p85 expression was first normalised to beta tubulin expression for each cell line and then normalised expression stated relative to expression in the U87MG cell line. Relative p85 expression was significantly higher in U87MG as compared to A172, LN229 and LN18 cell lines. **$p<0.01$, ***$p<0.001$, calculated by one-way ANOVA. Data presented as mean ± SEM (n=3).

5.2.3 Development of PTEN knock-down LN18 and LN229 cells using lentiviral transduction

As discussed previously, TRIPZ Inducible Lentiviral shRNA (Thermo Fisher Scientific) has been used to knock-down PTEN in both LN18 and LN229 PTEN proficient glioblastoma cell lines. This technique has previously been used in the group to knock-down PTEN in the MeWo PTEN proficient melanoma cell line. This work was performed by Dr R Abbotts as part of her PhD project (Abbotts, 2013). The initial steps of the knock-down process of shRNA plasmid preparation and recombinant lentiviral production (as described in Section 2.4.7.3) were performed by Dr Abbotts for the purposes of her thesis.

5.2.3.1 Determination of optimal puromycin concentration to select for transfected cells

The TRIPZ vector contains a puromycin drug resistance marker. The administration of puromycin to transfected cells allows selection of cells that have been successfully transfected. In order to establish the optimal dose for puromycin selection, non-transfected LN18 and LN229 cell lines were exposed to various doses of puromycin for 5 days. After 5 days exposure the dose of puromycin resulting in 100% cell death was identified. For both LN18 (Figure 5.34) and LN229 (Figure 5.35) 2µg/ml was found to be the optimal dose for puromycin selection.

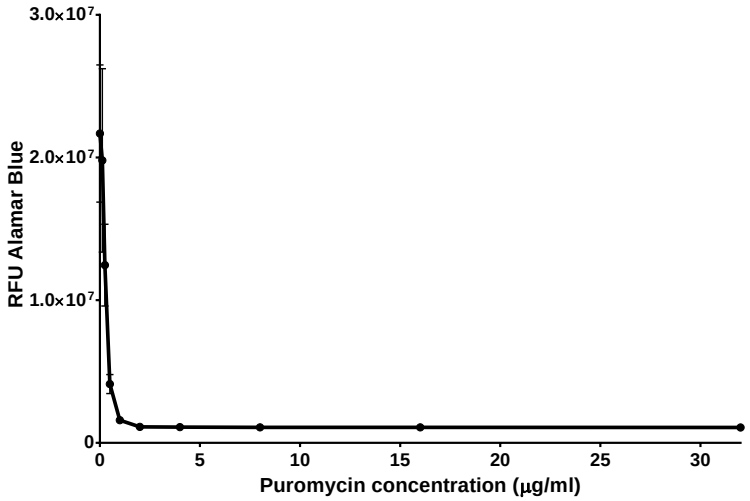


Figure 5.34. Puromycin selection curve for the LN18 cell line.

2µg/ml is the lowest concentration of puromycin resulting in 100% cell death. Data presented as mean±SD (n=1).

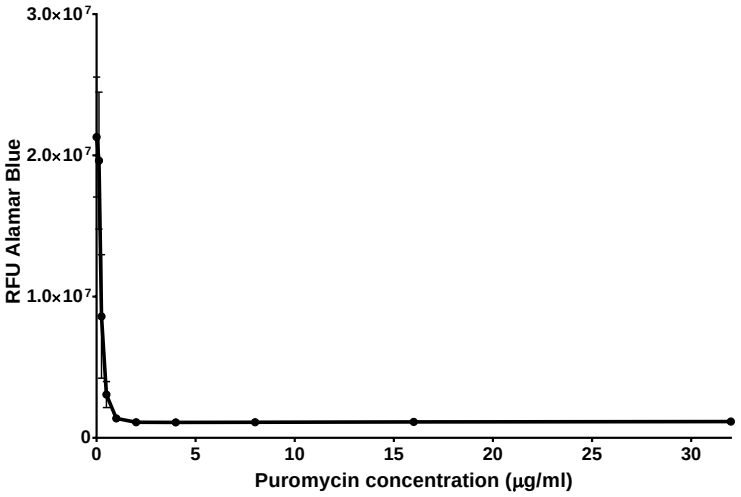


Figure 5.35. Puromycin selection curve for the LN229 cell line.

2µg/ml is the lowest concentration of puromycin resulting in 100% cell death. Data presented as mean±SD (n=1).

5.2.3.2 TurboRFP tag monitoring as confirmation of shRNA induction

An additional feature of the TRIPZ vector is that it contains a tetracycline-inducible promoter (TRE) that drives expression of both the shRNAmir and a TurboRFP reporter. Transduced cells can therefore be visualised by fluorescence microscopy after induction.

Doxycycline was added to transduced cells at a concentration of 0.5µg/ml once they were stable in puromycin selection. LN229-158 cells did not survive the addition of doxycycline, and are therefore excluded from all subsequent experiments. LN18 (158, 317 and 477) and LN229 (317 and 477) cell lines demonstrated expression of the TurboRFP tag after 72 hours exposure to doxycycline (Figure 5.36). Cells were inspected on a weekly basis for TurboRFP tag expression.

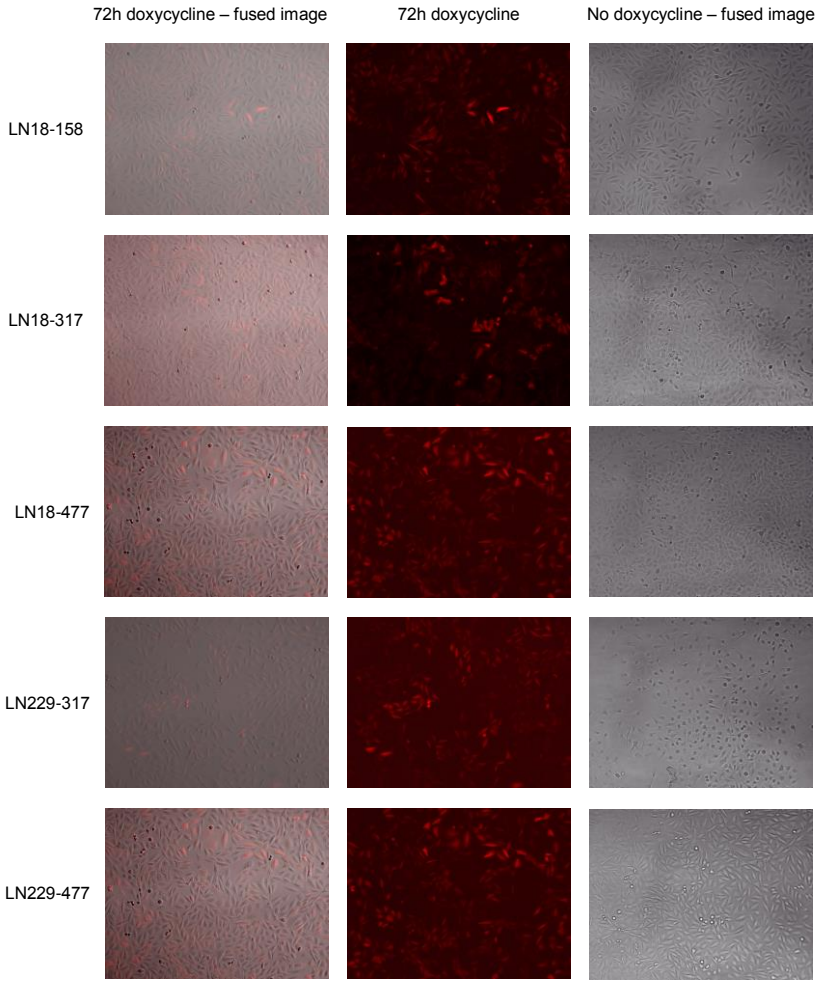


Figure 5.36. TurboRFP expression in transduced glioblastoma cell lines with and without doxycycline induction of shRNA expression.

Central panel demonstrates TurboRFP expression for all constructs, in both cell lines, after 72 hours doxycycline exposure. The proportion of cells expressing the shRNA can be visualised (left panel) in the merged brightfield/fluorescence image. In the absence of doxycycline, no TurboRFP is expressed (right panel – merged brightfield/ fluorescence image).

5.2.3.3 Confirmation of PTEN knock-down in transduced cells by Western blot analysis

Cell lysates were made for Western blot analysis at 7 and 14 days post doxycycline exposure. Cell lysates were also made for transduced cells not exposed to doxycycline. A representative Western blot for PTEN protein expression for all samples, with and without doxycycline, is shown in Figure 5.37.

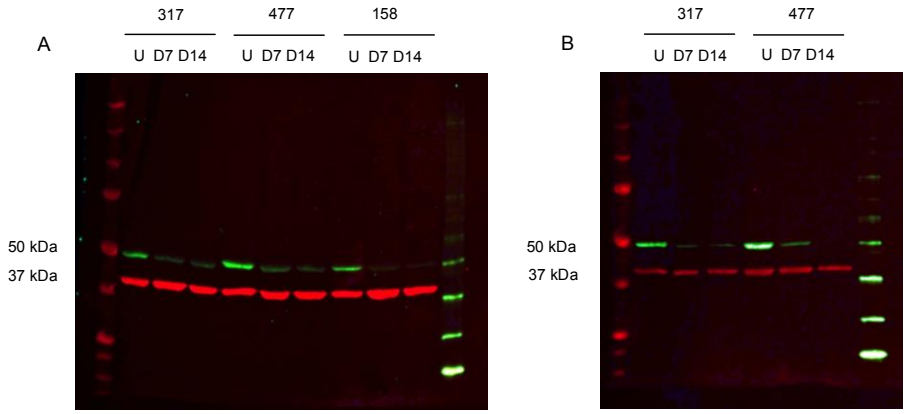


Figure 5.37. Western blot of PTEN expression in shRNA lentiviral transduced doxycycline inducible PTEN knock-down cell lines.

A. PTEN protein expression in the LN18 cell line transduced with 317, 477 and 158 constructs (untreated and after 7/14 days of doxycycline exposure). A single band is seen at approximately 50 kDa confirming PTEN protein expression in the samples not treated with doxycycline. Weaker bands or absent bands are demonstrated after 7 or 14 days of doxycycline exposure. B. PTEN protein expression in the LN229 cell line transduced with 317 and 477 constructs (untreated and after 7/14 days of doxycycline exposure). A single band is seen at approximately 50 kDa confirming PTEN protein expression in the samples not treated with doxycycline. Weaker bands or absent bands are demonstrated after 7 or 14 days of doxycycline exposure. U=untreated/not exposed to doxycycline, D7=7 days of doxycycline exposure, D14=14 days of doxycycline exposure.

PTEN knock-down is observed with all constructs at both time points. Greater levels of PTEN knock-down were demonstrated after 14 days exposure to doxycycline. The highest levels of PTEN knock-down are seen in the LN18-158 cell line (Figure 5.38). The LN229-158 cell line did not survive the

addition of doxycycline. The 477 construct demonstrated the greatest PTEN knock-down at day 14 across both the LN18 and LN229 cell lines (Figure 5.37) and has therefore been taken forward for further studies.

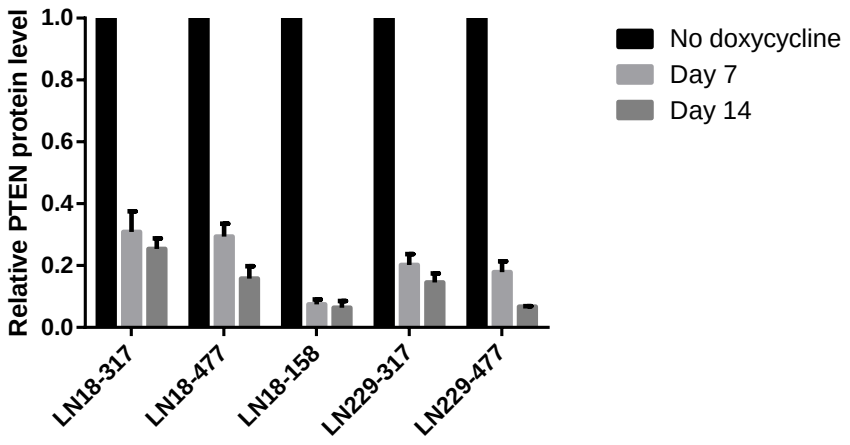


Figure 5.38. Quantification of PTEN protein expression from Western blot analysis in the LN18 and LN229 transduced doxycycline-inducible PTEN knock-down cell lines.

Graphical representation of relative PTEN protein expression for each of the PTEN knock-down cell lines (LN18-317, LN18-477, LN18-158, LN229-317 and LN229-477) is shown as no exposure to doxycycline, 7 or 14 days of doxycycline exposure. For each cell line/construct PTEN expression was first normalised to beta actin expression and then normalised expression stated relative to expression in the untreated (no doxycycline) sample. The greatest levels of PTEN knock-down were seen after 14 days of doxycycline exposure. Data presented as mean±SEM (n=3).

5.2.3.4 Western blot analysis demonstrates that ATM protein expression is preserved in the presence of PTEN knock-down

On Western blot analysis, ATM protein is expressed in both the LN18-477 and LN229-477 cell lines, in both the presence and absence of PTEN expression. No significant difference is demonstrated in ATM protein expression with PTEN knock-down (Figure 5.39).

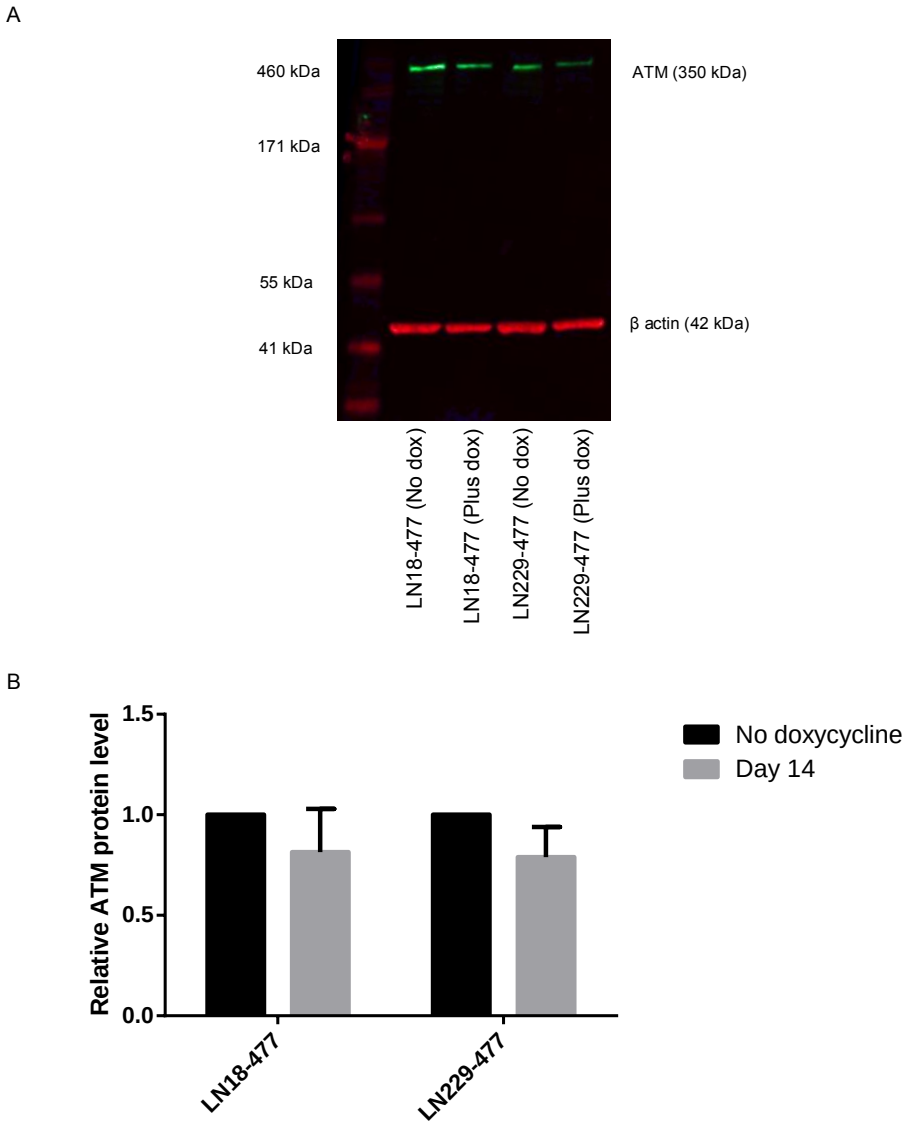


Figure 5.39. Western blot for ATM protein expression in the LN18-477 and LN229-477 lentiviral transduced PTEN knock-down cell lines.

A. Western blot analysis confirms expression of ATM in LN18-477 and LN229-477 doxycycline-inducible PTEN knock-down cell lines in the presence (plus doxycycline for 14 days) and absence (no doxycycline) of PTEN expression. A single band is seen at approximately 350 kDa confirming ATM protein expression. B. Quantification of ATM expression in LN18-477 and LN229-477 cell lines. No significant difference in ATM protein expression is observed between cells with PTEN expression (no doxycycline) or cells with absent PTEN expression (plus doxycycline for 14 days), calculated by one-way ANOVA. For each cell line/construct PTEN expression was first normalised to beta actin expression and then normalised expression stated relative to expression in the untreated (no doxycycline) sample. Data presented as mean±SEM (n=3).

5.2.3.5 Confirmation of PTEN knock-down and ATM expression in transduced cells by endpoint PCR

On endpoint PCR analysis, loss of PTEN mRNA expression is demonstrated in both the LN18-477 and LN229-477 cell lines after 14 days exposure to doxycycline (Figure 5.40C). ATM is expressed in the LN18-477 and LN229-477 cell lines, both treated and untreated with doxycycline (Figure 5.40D). PTEN expression status does not appear to alter expression of ATM.

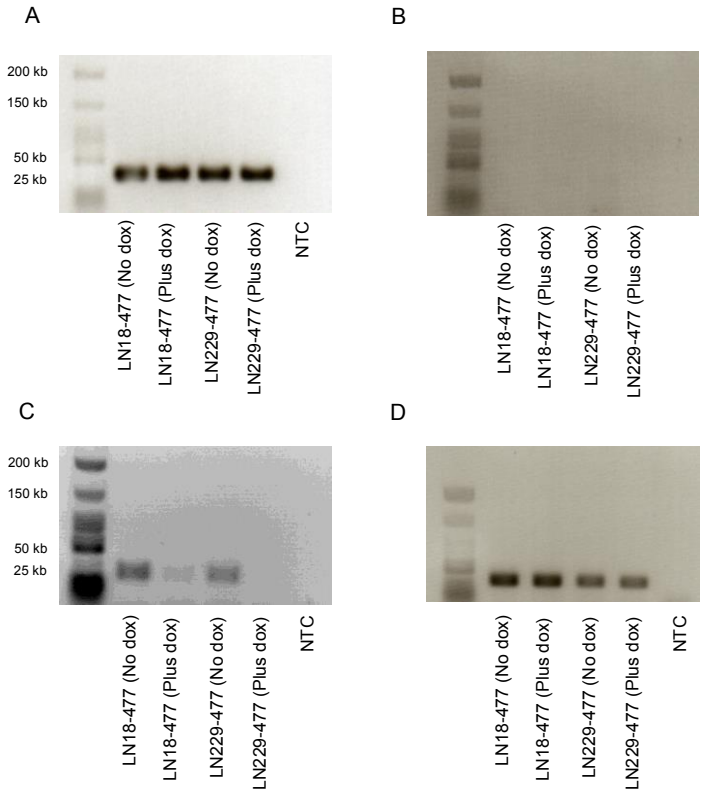


Figure 5.40. Endpoint PCR for beta actin, PTEN and ATM expression in LN18-477 and LN229-477 cell lines.

A. A representative endpoint PCR gel for beta actin expression in LN18-477 and LN229-477 cell lines. Beta actin is expressed to a similar level in all samples. B. No reverse transcriptase control samples. No bands are demonstrated for any of the cell lines indicating no genomic DNA contamination. C. A representative endpoint PCR gel for PTEN expression in LN18-477 and LN229-477 cell lines. PTEN expression is reduced/absent in LN18-477 (Plus dox) and LN229-477 (Plus dox) cell lines, as compared to the respective No dox controls. D. A representative endpoint PCR gel for ATM expression in LN18-477 and LN229-477 cell lines. ATM is expressed in No dox and Plus dox samples to a similar level in both cell lines. NTC=no template control. No dox=no previous exposure to doxycycline (control). Plus dox=14 days exposure to doxycycline (PTEN knock-down).

5.2.3.6 Doxycycline does not affect cellular proliferation in non-transduced cells

In order to assess the effects of doxycycline on non-transduced cells, LN18 and LN229 cell lines were seeded in to 96 well plates in media with and without doxycycline, as described in Section 2.4.7.9. Absorbance at 490nm was measured every 48 hours. Similar levels of absorbance were demonstrated for both cell lines (with and without doxycycline) for all time points (Figure 5.41).

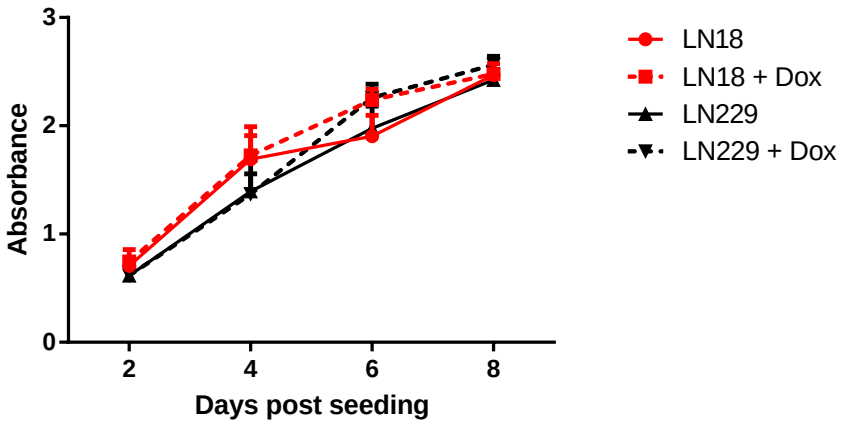


Figure 5.41. Proliferation of LN18 and LN229 non-transduced cell lines in response to doxycycline.

Absorbance (as a surrogate for proliferation) for LN18 and LN229 glioblastoma cell lines exposed to doxycycline was measured every 2 days (up to a total of 8 days). No difference is observed in the growth of LN18 and LN299 cell lines in response to doxycycline as measured by absorbance at 490nm. Data presented as mean±SD (n=1).

5.2.3.7 PTEN status does not influence cellular proliferation

In order to assess the effects of doxycycline withdrawal on PTEN expression and cellular proliferation, LN18-477 and LN229-477 cell lines were seeded in to 96 well plates (as described in Section 2.4.7.8). In summary, cells were seeded in triplicate that had a) never been exposed to doxycycline and were

maintained in doxycycline-free media (PTEN expressing), b) been growing in doxycycline-containing media but that were grown in doxycycline-free media for the duration of this assay, and c) been growing in doxycycline-containing media and were maintained in doxycycline-containing media (PTEN knock-down). Absorbance at 490nm was measured every 48 hours. Similar levels of absorbance, and hence similar rates of proliferation, were demonstrated for all three conditions with each cell line (Figure 5.42A). Western blot analysis demonstrated that for cells that had been grown in doxycycline-containing media (hence are PTEN knock-down), PTEN expression starts to return by 4 days post doxycycline removal (Figure 5.42 B-C).

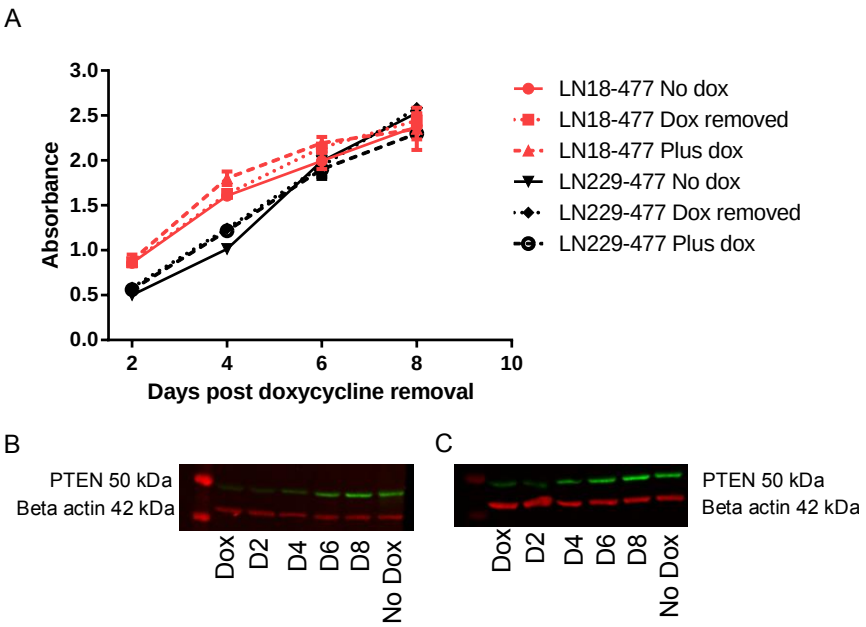


Figure 5.42. Proliferation of LN18-477 and LN229-477 glioblastoma cell lines in response to doxycycline removal.

A. Proliferation of LN18-477 and LN229-477 cell lines that: a) have never been exposed to doxycycline, b) have been exposed to doxycycline but this was removed on Day 0 or c) have been exposed to, and maintained in, doxycycline. Absorbance (used as a surrogate measure of proliferation) was recorded at 490nm every 48 hours. No difference in growth was demonstrated between the three conditions for either cell line. Data presented as mean±SD (n=1). B. Western blot analysis of LN18-477 cell line demonstrating PTEN protein expression in the presence (Dox) and absence (No Dox) of doxycycline and 2-8 days (D2-8) post removal of doxycycline. PTEN expression starts to recover approximately 4 days post doxycycline removal. C. Western blot analysis of LN229-477 cell line demonstrating PTEN expression in the presence (Dox) and absence (No Dox) of doxycycline and 2-8 days (D2-8) post removal of doxycycline. PTEN expression starts to recover approximately 4 days post doxycycline removal.

5.2.3.8 Sensitivity of lentiviral PTEN knock-down glioma cell lines to ATM inhibitor treatment

Two MTS protocols were followed to determine the sensitivity of LN18-477 and LN229-477 control and PTEN knock-down cell lines to the ATM inhibitor KU55933. In protocol one, the standard MTS protocol was followed, except that control cells were seeded in media without doxycycline whereas knock-down cells were seeded in media containing doxycycline at a concentration of 0.5µg/µl. On the addition of drug, old media was removed and replaced with drug containing fresh media, either with or without doxycycline, as appropriate. The results for protocol one are shown in Figure 5.43.

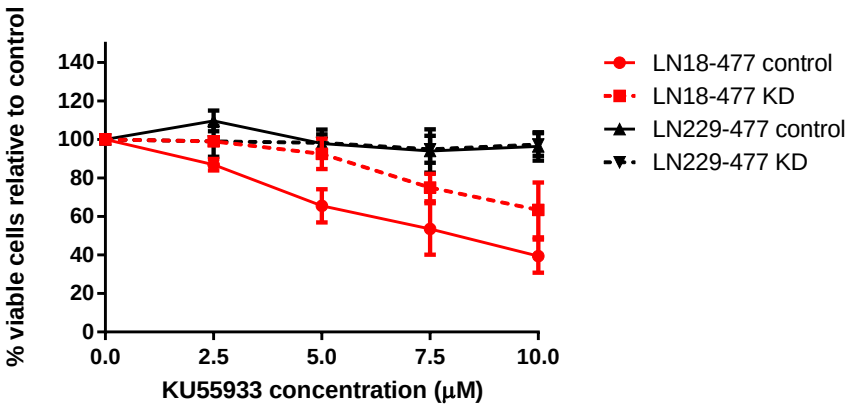


Figure 5.43. Protocol 1: MTS for LN18-477 and LN229-477 (control and PTEN knock-down) cell lines with KU55933.

LN18-477 control (PTEN expressing) cell line exhibited less viable cells after exposure to ATM inhibitor (10µM KU55933) than LN18-477 PTEN knock-down cell line; however this does not reach statistical significance. ATM inhibitor treatment had no effect on cell viability in the LN229-477 cell line, irrespective of PTEN status. Data presented as mean±SEM (n=3).

In the second protocol, the same methodology was followed as per protocol one except that fresh doxycycline was added every 48 hours to wells containing knock-down cells. The results for protocol 2 are shown in Figure 5.44. Similar results were demonstrated with both protocols. PTEN expression did not modify the sensitivity of the LN229-477 cell line to KU55933. PTEN expression, however, did modify sensitivity of the LN18-477 cell line to KU55933. LN18-477 control cells were more sensitive to ATM inhibitor treatment than LN18-477 PTEN knock-down cells; however this did not reach statistical significance.

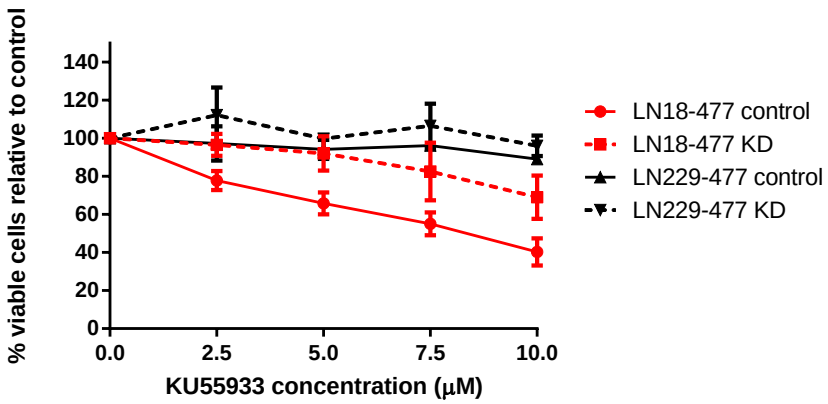


Figure 5.44. Protocol 2: MTS for LN18-477 and LN229-477 (control and PTEN knock-down) cell lines with KU55933.

LN18-477 control (PTEN expressing) cell line exhibited less viable cells after exposure to ATM inhibitor (10µM KU55933) than LN18-477 PTEN knock-down cell line; however this does not reach statistical significance. ATM inhibitor treatment had no effect on cell viability in LN229-477 cell line, irrespective of PTEN status. Data presented as mean±SEM (n=3).

5.2.3.9 DNA repair gene expression in PTEN knock-down cell lines

Previous studies have shown that PTEN loss is associated with DNA repair deficiency; however earlier work in this chapter has found conflicting results. While lower expression levels of a number of DNA repair genes was found in the A172 and U87MG (both PTEN deficient) cell lines as compared to the LN18 (PTEN proficient) cell line, these genes were also down-regulated in the LN229 (PTEN proficient) cell line. To further explore this concept isogenic PTEN knock-down LN18-477 and LN229-477 cell lines have been investigated using the same array. Genes that show a 2-fold difference in expression between the control and PTEN knock-down sample have been identified. RPLPO was selected as the house keeper gene with the least variation in C_T values across the samples from a panel of five commonly used housekeeper genes.

LN18-477 C_T values were normalised against the housekeeper gene, RPLPO. Multiple peaks were observed on the melting curves for the DMC1 primer in both the LN18-477 control and knock-down cell lines and it was therefore excluded from subsequent analyses. A graphical representation of the expression difference for the panel of DNA repair genes in the LN18-477 control cell line as compared to the same cell line with the induction of PTEN knock-down is shown in Figure 5.45.

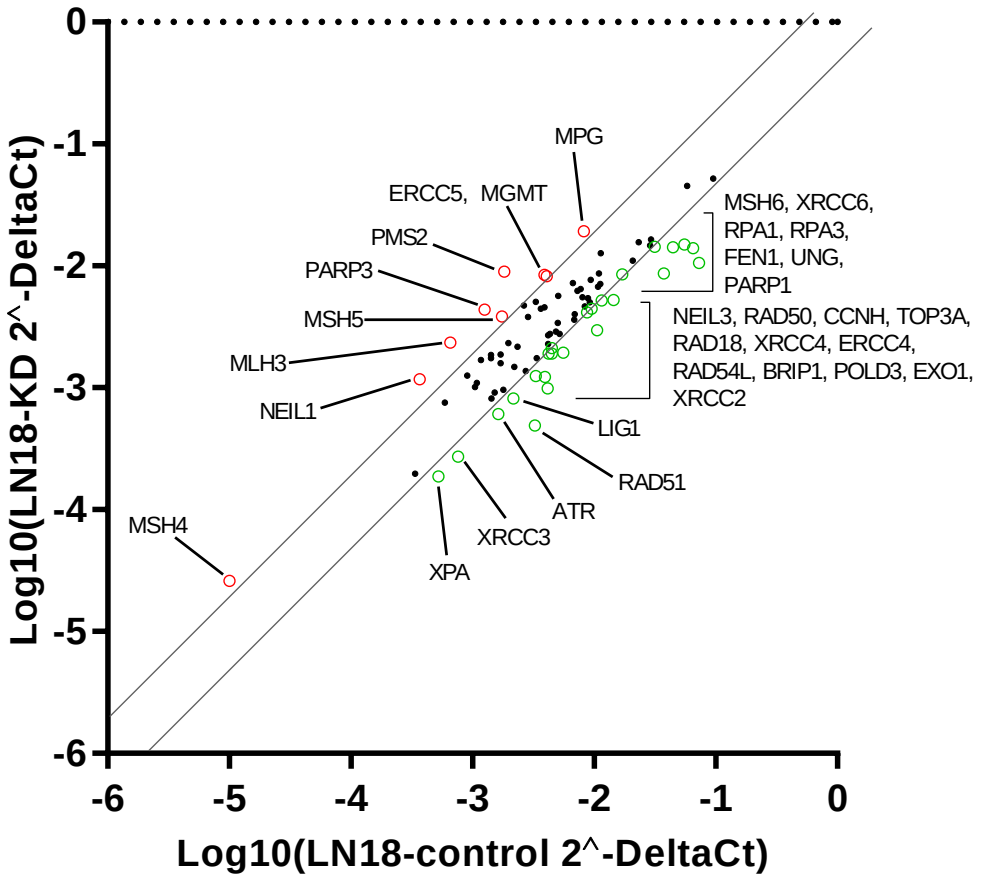


Figure 5.45. Scatterplot comparing DNA repair gene expression in the LN18-477 control (x axis) against the LN18-477 PTEN knock-down (y axis) glioblastoma cell lines.

The area outside of the diagonal grey lines represents DNA repair genes with a greater than two-fold difference in expression between the two cell lines (n=1). Green circles represent genes that are under-expressed in LN18-477 PTEN knock-down as compared to LN18-477 PTEN expressing control cells. Red circles represent genes that are over-expressed in LN18-477 PTEN knock-down as compared to LN18-477 PTEN expressing control cells.

Nine DNA repair genes were found to up-regulated in LN18-477 PTEN knock-down cell line as compared to the LN18-477 control cell line. PMS2 was the DNA repair gene showing the greatest up-regulation in the LN18-477 PTEN knock-down cell line as compared to the PTEN expressing LN18-477 control. Four of the nine up-regulated genes are components of the mismatch repair pathway (Table 5.5).

Table 5.5. DNA repair genes at least two-fold up-regulated in LN18-477 PTEN knock-down as compared to LN18-477 PTEN expressing control.

Gene	Fold change	Pathway
ERCC5	2.1647	NER
MGMT	2.0048	Direct repair
MLH3	3.573	MMR
MPG	2.3261	BER
MSH4	2.6614#	MMR
MSH5	2.2111	MMR
NEIL1	3.2022	BER
PARP3	3.4745	BER
PMS2	4.9398	MMR

The fold change for each gene (and cell line) is shown as compared to the LN18-477 control cell line; the DNA repair pathway to which the gene belongs is also shown. #This gene's CT value is relatively high (>30) meaning that its expression level is low, in both control and knock-down samples. The fold change result may therefore be subject to greater variation. NER=nucleotide excision repair; MMR=mismatch repair; BER=base excision repair.

In addition, 24 genes were found to be under-expressed in the LN18-477 PTEN knock-down cell line as compared to the LN18-477 control cell line. The greatest change in expression was seen for FEN1, a greater than 6-fold reduction in expression in the LN18-477 PTEN knock-down cell line was demonstrated as compared to the control. Genes from all major DNA repair pathways are represented in the down-regulated genes (Table 5.6).

Table 5.6. DNA repair genes at least two-fold down-regulated in LN18-477 PTEN knock-down as compared to LN18-477 PTEN expressing control.

Gene	Fold change	Pathway
ATR	-2.6767	DDR
BRIP1	-2.1217	NER
CCNH	-2.1386	NER
ERCC4	-2.372	DSB repair
EXO1	-3.2294	DSB repair
FEN1	-6.8957	BER
LIG1	-2.6545	NER
MSH6	-2.2064	MMR
NEIL3	-2.7604	BER
PARP1	-2.0005	BER
POLD3	-2.6617	MMR
RAD18	-3.5732	DSB repair
RAD50	-2.2201	DSB repair
RAD51	-6.6398	DBS repair
RAD54L	-2.2154	DSB repair
RPA1	-3.6983	NER
RPA3	-4.6517	NER
TOP3A	-2.0929	TOP3A
UNG	-4.3187	BER
XPA	-2.8064*	NER
XRCC2	-4.2048	DSB repair
XRCC3	-2.8161*	DSB repair
XRCC4	-2.8801	DSB repair
XRCC6	-3.1564	DSB repair

The fold change for each gene (and cell line) is shown as compared to the LN18-477 control cell line; the DNA repair pathway to which the gene belongs is also shown. *This gene's average CT value is relatively high (>30) in one sample and reasonably low in the other sample (<30). This means that the gene's expression is relatively low in one sample and reasonably detected in the other. This is suggests that the actual fold-change is at least as high as reported. DDR=DNA damage response; NER=nucleotide excision repair; DBS=double strand break; BER=base excision repair; MMR=mismatch repair.

LN229-477 C_T values were normalised against the housekeeper gene RPLPO. Multiple peaks were observed on the melting curves for the MGMT and MSH4 primers in the LN229-477 control cell lines and therefore these primers were excluded from subsequent analyses. A graphical representation of the expression difference for the panel of DNA repair genes in the LN229-477 control cell line as compared to the same cell line with the induction of PTEN knock-down is shown in Figure 5.46.

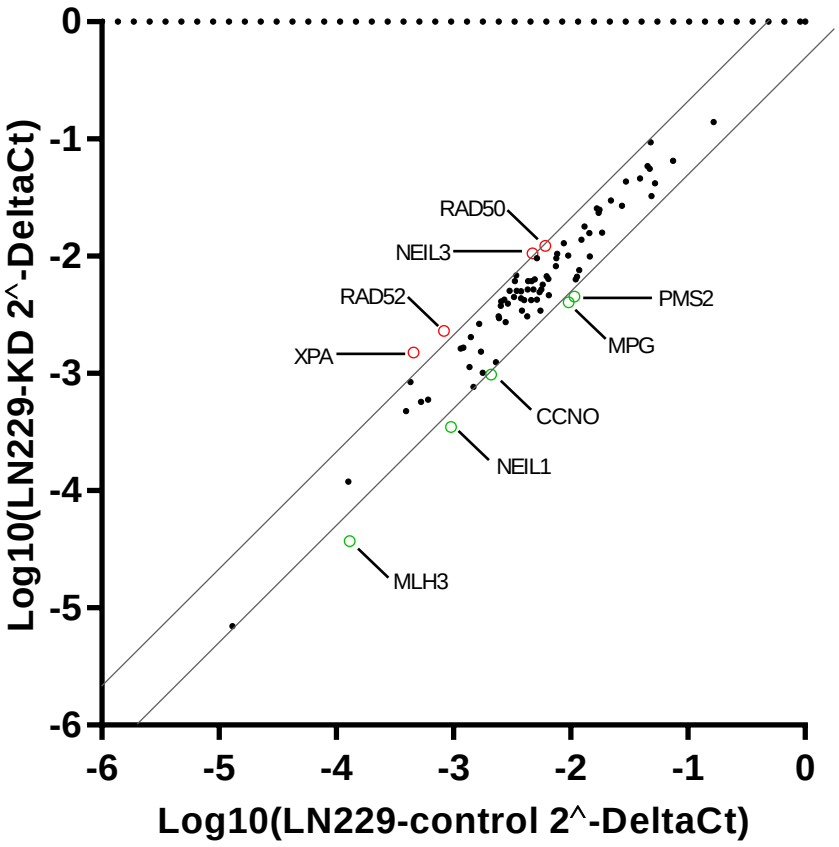


Figure 5.46. Scatterplot comparing DNA repair gene expression in LN229-477 control (x axis) against LN229-477 PTEN knock-down (y axis) glioblastoma cell lines.

The area outside of the diagonal grey lines represents DNA repair genes with a greater than two-fold difference in expression between the two cell lines (n=1). Green circles represent genes that are under-expressed in LN229-477 PTEN knock-down as compared to LN229-477 PTEN expressing control cells. Red circles represent genes that are over-expressed in LN229-477 PTEN knock-down as compared to LN229-477 PTEN expressing control cells.

Four DNA repair genes demonstrated a greater than 2-fold increase in expression in the LN229-477 PTEN knock-down cell line as compared to the control, PTEN expressing LN229-477 cell line. The greatest increase in expression was seen for XPA, which exhibited a greater than 3-fold up-regulation in the LN229 PTEN knock-down cell line as compared to the control. Interestingly, NEIL3, RAD50 and XPA were all at least 2-fold up-

regulated in the LN229-477 PTEN knock-down cell line as compared to the control (Table 5.7), but in the LN18-477 PTEN knock-down cell line (as compared to the LN18-477 control), the same genes were at least 2-fold down-regulated.

Table 5.7. DNA repair genes at least two-fold up-regulated in LN229-477 PTEN knock-down as compared to LN229-477 PTEN expressing control.

Gene	Fold change	Pathway
NEIL3	2.2322	BER
RAD50	2.002	DSB repair
RAD52	2.7648	DSB repair
XPA	3.3213	NER

The fold change for each gene (and cell line) is shown as compared to the LN229-477 control cell line; the DNA repair pathway to which the gene belongs is also shown. BER=base excision repair; DBS=double strand break; NER=nucleotide excision repair.

Five DNA repair genes were found to be down-regulated in the LN229-477 PTEN knock-down cell line as compared to the LN229-477 PTEN expressing control. MLH3 has the greatest decrease in expression in the LN229-477 PTEN knock-down cell line as compared to the control (Table 5.8). Four of these genes (MLH3, MPG, NEIL1 and PMS2) are down regulated in the LN299-477 knock-down PTEN cell line as compared to the LN299-477 control, but in the LN18-477 cell line are up-regulated in the PTEN knock-down as compared to the control.

Table 5.8. DNA repair genes at least two-fold down-regulated in LN229-477 PTEN knockdown as compared to LN229-477 PTEN expressing control.

Gene	Fold change	Pathway
CCNO	-2.142	BER
MLH3	-3.5334#	MMR
MPG	-2.3609	BER
NEIL1	-2.7172	BER
PMS2	-2.3729	MMR

The fold change for each gene (and cell line) is shown as compared to the LN229-477 control cell line; the DNA repair pathway to which the gene belongs is also shown. #This gene's Ct value is relatively high (>30) meaning that it's expression level is low, in both control and knock-down samples. The fold change result may therefore be subject to greater variation. BER=base excision repair; MMR=mismatch repair.

5.3 Discussion and conclusions

The aim of this chapter was to explore the effectiveness of targeting PTEN deficient glioblastoma cell lines with an ATM inhibitor. The rationale behind this is that PTEN deficiency has previously be shown to associated with down-regulation of DNA repair; hence, the addition of an ATM inhibitor would result in increased cell kill in PTEN deficient cell lines by a concept known as synthetic lethality. The results described in this chapter are interesting in that they provide no evidence to support a synthetic lethality relationship between PTEN deficiency and ATM inhibitor treatment. They do, however, demonstrate increased sensitivity to ATM inhibitor treatment in a PTEN proficient glioblastoma cell line. The potential mechanisms behind this will be discussed further.

5.3.1 High ATM expression is associated with poor survival in PTEN expressing high grade glioma

The results of this study demonstrate that at both the gene and protein level, high ATM expression is associated with poor survival in patients with high levels of PTEN expression. No association is seen between ATM expression and survival in patients with low levels of PTEN expression, in either the TCGA or Nottingham cohort.

Several earlier studies have investigated ATM expression as a marker of prognosis across a range of tumour sites. In the majority these studies, low ATM expression, either alone or in combination with other markers, was associated with poor survival (Kamphues et al., 2015, Ye et al., 2007, Kim et

al., 2014a, Kim et al., 2014b, Abdel-Fatah et al., 2014a). In a study of oesophageal squamous cell carcinoma patients treated with neoadjuvant chemo-radiotherapy, ATM expression was not found to be associated with overall survival (Sarbia et al., 2007). A further study by Guo *et al* investigated ATM expression in combination with oestrogen receptor status in a cohort of breast cancer patients and found that ATM is over-expressed in oestrogen receptor negative tumours. In the same study, elevated ATM levels were found to be associated with local recurrence after radiotherapy (Guo et al., 2013), which would be in keeping with the knowledge that low ATM expression is associated with radio-sensitivity (Hickson et al., 2004, Golding et al., 2009).

Seol *et al* also investigated ATM protein expression in a cohort of 69 glioblastoma patients by immunohistochemistry. Interestingly, they demonstrated that high ATM protein expression was associated with longer overall and progression-free survival on both univariate and multivariate analysis. This is in contrast to the results of the Nottingham cohort, where no association was seen between ATM protein expression and survival when all patients were included. A potential explanation for these differing results may be the different cut-offs used to define ATM expression; in the study by Seol *et al* ATM over-expression was defined as ≥25% of cells positive for ATM expression (Seol et al., 2011), whereas in the current study, high ATM expression was classified as >75% of cells positive. Also, the Nottingham cohort included both grade 3 and grade 4 gliomas, whereas the Seol study just included glioblastoma patients. Grade 3 glioma patients are known to have a better prognosis than patients with a glioblastoma; however as grade has not

been found to be associated with ATM expression in the Nottingham cohort this is unlikely to explain the variation in results seen between the two studies.

High ATM expression has been shown to be associated with poor survival, but only in tumours with high PTEN expression, at both the mRNA and protein level. ATM and PTEN are known to interact. ATM has previously been shown to phosphorylate PTEN and to promote PTEN translocation to the nucleus (Chen et al., 2015). One potential explanation for why high ATM expression is associated with poor survival, but only in tumours that have high nuclear PTEN expression, may be that elevated levels of ATM result in a re-distribution of cellular PTEN, resulting in increased nuclear translocation (Chen et al., 2015). Nuclear PTEN has previously been shown to promote DNA repair (Bassi et al., 2013, Lian and Di Cristofano, 2005) and hence an elevated level of nuclear PTEN could result in increased repair of damage induced by cytotoxic agents and hence increased cancer cell survival. Interestingly, in this current study nuclear PTEN expression was not associated with ATM expression.

The Nottingham cohort, as mentioned previously, includes both grade 3 and 4 glioma patients which could influence the results of the PTEN subgroup analysis. PTEN expression has been shown to be associated with grade, with higher levels of PTEN expression seen in lower grade glioma. Further review of the high PTEN expressing subgroup reveals that it does include more grade 3 tumours than GBMs (eight vs three); however, all three GBMs had low ATM protein expression levels. This would suggest that the association between high ATM expression and poor overall survival is not purely a reflection of

tumour grade. Obviously caution should be exercised in the interpretation of results from this small, mixed group of patients; however, the same association between high ATM expression and poor overall survival is also observed in patients with high PTEN expression at the mRNA level in the larger TCGA cohort.

5.3.2 ATM inhibitor treatment does not result in synthetic lethality in PTEN deficient glioblastoma cell lines

In keeping with the results from the mRNA and protein expression studies, no evidence of a synthetic lethality relationship between PTEN deficiency and ATM inhibitor treatment was demonstrated. Interestingly, the cell line most sensitive to an ATM inhibitor was a PTEN proficient cell line (LN18). The second PTEN proficient cell line (LN229) included in this study, however, did not show increased sensitivity to the ATM inhibitor treatment. This suggests that genetic aberrations other than, or in addition to PTEN status, are responsible for the observed results. Previous studies that have explored synthetic lethality in PTEN deficient systems have shown differing results. In colorectal and endometrial cancer cell lines synthetic lethality has been demonstrated with PARP inhibitors in PTEN deficient tumours (Mendes-Pereira et al., 2009). Similarly, PTEN-deficient astrocytes have been shown to have impaired DSB repair and to be sensitive to PARP inhibition (McEllin et al., 2010). In PTEN deficient prostate cancer cell lines, however, PARP inhibition was not able to induce synthetic lethality (Fraser et al., 2012). In melanoma cell lines, synthetic lethality was demonstrated between PTEN status and inhibition of APE1, a key component of the BER pathway (Abbotts

et al., 2014). Also, a recently published study has demonstrated that PTEN deficient colorectal cancer cell lines can be targeted by ATM inhibition to induce synthetic lethality. Interestingly, in the same study by McCabe *et al*, additional cell lines were also screened for sensitivity to ATM inhibition dependent on PTEN status, including the glioma cell lines T98G (PTEN proficient) and U251 (PTEN deficient). In contrast to the current study, the study by McCabe *et al* found that ATM inhibitor treatment was synthetically lethal in U251, the PTEN deficient glioma cell line (McCabe et al., 2015).

In order to further investigate the effect of PTEN status on sensitivity to ATM inhibitor treatment, the PTEN gene was successfully knocked-down using lentiviral techniques in both the LN18 and LN229 cell lines. Levels of PTEN knock-down of greater than 80% were achieved using this technique. The 477 construct was chosen for further evaluation as across both cell lines this demonstrated the highest average level of knock-down. PTEN status did not affect the sensitivity of the LN229 cell line to an ATM inhibitor. However, in the LN18 transduced cell line, knock-down of PTEN did decrease the effects of ATM inhibitor treatment on survival, although this did not reach statistical significance. Although these results should be interpreted with caution, they do provide some evidence that PTEN status may impact on the sensitivity of glioblastoma cell lines to ATM inhibitor treatment, but that other factors are likely to be implicated.

One potential limitation of this component of the study is that the MTS assay was used throughout as the method for determining cell survival. This type of assay is widely used in the scientific literature (O'Toole et al., 2003,

Thangapazham et al., 2007, Farabegoli et al., 2007) and produced consistent results in this study. There are potential disadvantages, including that this is an indirect measure of cellular viability achieved via measurement of mitochondrial metabolic rate; if metabolic rate is affected, then this can influence the results of the test (Plumb et al., 1989). Additionally, some chemicals may interact with MTS reagents or change the activity of succinate dehydrogenase, the enzyme that reduces the tetrazolium salt to formazan (Bruggisser et al., 2002, Wang et al., 2010b). MTS assays may therefore underestimate the cytotoxic effects of agents on cells (Wang et al., 2010b).

Ideally, the results seen in the MTS assay would have been further validated using a technique, such as a clonogenic assay. Due to the inability of these cell lines to grow at low density, clonogenic assays were not possible. This is in keeping with previous studies which have shown very low plating efficiencies for a range of glioma cell lines (Nilsson et al., 1980, Diserens et al., 1981). In order to provide additional evidence to support the results seen in the MTS assay, a growth curve experiment was performed for the LN18 and U87MG cell lines. LN18 was selected as this is the cell line sensitive to ATM inhibitor treatment in the MTS assay, and U87MG as this represents a PTEN deficient cell line. In the presence of 10μM KU55933, treated LN18 cells demonstrated reduced viability as compared to untreated LN18 cells; no difference in viability was found between treated and untreated U87MG cells. This is in keeping with the results seen in the MTS assay. Due the intensive nature of this assay, it was only performed once and ideally should be repeated, with the

inclusion of both the LN229 and A172 cell lines, and also for the PTEN knock-down cell lines.

Although ATM is constitutively expressed in cells it is maintained in an inactive state, only undergoing phosphorylation, and hence activation, in response to DNA damage (Bakkenist and Kastan, 2003). Potentially DNA damage may have been required prior to the administration of an ATM inhibitor in this study in order to activate ATM. In addition, confirmation of inhibition of ATM by KU55933 should be determined by assessing levels of phosphorylated ATM both pre- and post-inhibitor administration. Decreased levels of phosphorylated ATM would be expected if KU55933 were to be inducing ATM inhibition.

5.3.3 ATM inhibitor treatment does not result in an increase in DNA damage and double-strand breaks

As discussed in Section 1.4, ATM plays an important role in DSB recognition. Previous studies have shown ATM inhibitor treatment to be associated with an increase in DNA DSBs and the number of γH2AX foci (Sultana et al., 2013, Shaheen et al., 2011, Albarakati et al., 2015). Despite demonstrating increased sensitivity to ATM inhibitor treatment, the LN18 cell line did not show a significant increase in DNA damage. In the neutral comet assay mean tail moments were higher in LN18 cells treated with ATM inhibitor as compared to untreated cells; however, this did not reach statistical significance. In keeping with the results of the neutral comet assay, no significant difference was seen between untreated and KU55933 treated glioblastoma cell lines and the γH2AX levels observed. This included the LN18 cell line that in the MTS

assay showed significant sensitivity to KU55933. The LN18 cell line does, however, demonstrate increased levels of apoptosis in response to ATM inhibitor treatment, a finding not observed in the A172, U87MG or LN229 cell lines.

The finding that the LN18 cell line does not accumulate DNA DSBs in response to ATM inhibitor treatment was surprising. This may be explained by problems with the assays. Inter-experimental variation was seen in the neutral comet assay and is a recognised problem (Collins et al., 2014). While the number and extent of comets formed was limited, the inclusion of MMS as a positive control would suggest that the assay was successful. One potential explanation for the findings in this current study may be that the best time point to detect DSBs was not selected. During optimisation of the neutral comet assay 6, 24 and 48 hour time points were studied. 6 and 24 hour time points demonstrated the highest levels of damage that appeared to be resolving by 48 hours. Time-points less than 6 hours, however, were not investigated and could be incorporated into future studies. For the γH2AX assay, the 24 hour time point was studied and potentially may be either too early to detect DNA DSBs or too late, i.e. that by this time they had already been repaired. An earlier study that investigated the formation of γH2AX foci in response to treatment with UV light and alkylating agents found that peak values were seen between 8-12 hours post exposure (Staszewski et al., 2008). However, in the study by McCabe *et al*, γH2AX foci formation was assessed 72 hours after treatment with ATM inhibitor and found to be significantly increased in the PTEN deficient cell lines (McCabe et al., 2015). Albarakati *et al* investigated

synthetic lethality in BRCA1 deficiency with ATM inhibitors, and found an increase in γH2AX foci formation 48 hours after drug administration (Albarakati et al., 2015). Taken together this would suggest that the timing of γH2AX formation depends on both the drug administered and the growth characteristics of the cell line studied. Ideally both shorter and longer time points should be investigated before concluding that the glioma cell lines studied do not accumulate DSBs in response to ATM inhibitor treatment.

A potential explanation for the lack of increase in γH2AX levels detected in response to an ATM inhibitor may be explained by the fact that H2AX is a phosphorylation target for ATM; therefore a reduction in ATM levels would result in a reduction in γH2AX formation. This is supported by the findings of a study by Porcedda *et al* who describe the use of a flow cytometry assay for γH2AX detection after 2Gy irradiation for the diagnosis of Ataxia Telangiectasia (AT). Using this methodology, patients with AT were found to have lower levels of γH2AX after 2Gy irradiation than healthy controls. γH2AX was still detectable in AT patients (Porcedda et al., 2008) suggesting, in line with earlier studies, that DNA-PKcs can also phosphorylate H2AX (Stiff et al., 2004).

5.3.4 ATM inhibitor treatment results in G1/S phase cell cycle arrest in the LN18 cell line

While this study did not demonstrate an accumulation of DNA damage in response to an ATM inhibitor, a cell cycle effect was observed. Treatment of the LN18 cell line (PTEN proficient) with KU55933 resulted in an accumulation of cells in the G1 phase of cell cycle. This is in contrast to

results seen in several earlier studies which have demonstrated a G2/M arrest in response to an ATM inhibitor. McCabe *et al* found that 72 hours treatment with KU55933 resulted in increased G2/M cell cycle arrest in HCT116 PTEN deficient cells as compared to wild-type cells, and that more specifically these cells were arrested in mitosis (McCabe et al., 2015). A similar G2/M arrest was demonstrated by Albarakati *et al* in BRCA1 deficient cells treated with an ATM inhibitor (Albarakati et al., 2015).

Interestingly, in common with the present study, Li *et al* found that treatment with KU55933 can induce G1 cell cycle arrest. In this earlier study, KU55933 was found to reduce cellular proliferation and induce apoptosis by blocking AKT in cancer cells with over-activated AKT. The G1 arrest seen in the study by Li *et al* was found to be associated with down-regulation of cyclin D1 (Li and Yang, 2010). AKT has previously been shown to promote G1 cell cycle progression by upregulation of the G1 associated cyclins, such as cyclin D1, or the activity of cyclin-dependent kinases (Nicholson and Anderson, 2002). In order to establish whether the G1 cell cycle arrest seen in the current study is due to decreased levels of cyclin D1, western blot analysis should be performed for cyclin D1 expression in cell lines both untreated and treated with KU55933.

5.3.5 Potential explanations for the enhanced sensitivity of the LN18 cell line to ATM inhibitor treatment

The results in this study do not provide any evidence to support the hypothesis that PTEN deficient glioblastoma can be targeted by a synthetic lethality strategy using an ATM inhibitor. The study does, however, demonstrate that

the LN18 (PTEN proficient) cell line is sensitive to ATM inhibitor treatment and that in response to this it undergoes G1/S cell cycle arrest and apoptosis. ATM inhibitor treatment in the LN18 cell line does not appear to result in an accumulation of DNA damage. In an attempt to begin to unravel the potential mechanism(s) behind these findings, genetic differences between the LN18 cell line and the three non-sensitive glioblastoma cell lines have been explored. This includes differences in: DNA repair gene expression, p53 expression and expression of components of the PI3K signalling pathway.

5.3.5.1 DNA repair gene expression does not explain the sensitivity of the LN18 cell line to ATM inhibitor treatment

In the context of this study, the RT^2 DNA Repair PCR Array has been used to address two questions, firstly to determine if PTEN deficiency is associated with a down-regulation of DNA repair, especially DSB repair, and secondly to identify if changes in DNA repair gene expression may explain the unexpected sensitivity of the LN18 cell line to ATM inhibitor treatment.

Review of the data from the RT^2 DNA Repair PCR Array reveals a panel of fourteen DNA repair genes to be down-regulated in the PTEN deficient glioblastoma cell lines A172 and U87MG, as compared to the LN18 (PTEN proficient) cell line. However, these same DNA repair genes are also down-regulated in the LN229 (PTEN proficient) cell line when compared to LN18, suggesting that this is not a PTEN mediated effect. Similarly, analysis of the array plates for the LN18-477 and LN229-477 PTEN isogenic cell lines reveals

that while DNA repair gene expression does differ between the paired cell lines, this is not consistent between LN18-477 and LN229-477.

PTEN deficiency has previously been reported to be associated with down-regulation of RAD51 (Mendes-Pereira et al., 2009) but while there is some evidence of down-regulation of RAD51 and its paralogs in the PTEN deficient state in this current study, this again is not consistent across all cell lines studied. In the context of this current study, which does not demonstrate synthetic lethality between the PTEN deficient state and ATM inhibitor treatment, the lack of a consistent pattern of DNA repair gene down-regulation in PTEN deficiency is not surprising. This is also in keeping with previous studies, including the work by McCabe *et al*. In this earlier study, no association was demonstrated between RAD51 expression, or RAD51 foci formation, and PTEN expression in a panel of glioma cell lines including T98G (PTEN proficient) and U251 (PTEN deficient). This is despite the fact that the U251 cell line was sensitive to ATM inhibition (McCabe et al., 2015).

DNA repair gene expression does not appear to explain the sensitivity of the LN18 (PTEN proficient) cell line to ATM inhibitor treatment. While a small number of DNA repair genes are down-regulated in LN18 as compared to the other cell lines studied, none are consistently down-regulated. Ligase 4 expression is down-regulated in LN18 as compared to A172 and U87MG, and DDB1 is down-regulated in LN18 as compared to A172 and LN229. This lack of consistent down-regulation of either Ligase 4 or DDB1 makes a synthetic lethality association between either of these DNA repair genes and ATM unlikely. This would be in keeping with the results of the functional studies,

whereby no increase in DNA damage was demonstrated in response to an ATM inhibitor.

5.3.5.2 p53 status may influence response to ATM inhibitor treatment

As discussed previously, the sensitivity of the LN18 (PTEN proficient) glioblastoma cell line to an ATM inhibitor does not appear to be a DNA repair mediated process. The presence of G1 cell cycle arrest in response to ATM inhibitor treatment suggests that this may be having a cell cycle mediated effect.

The p53 tumour suppressor gene has numerous roles within the cell, including inducing G1/S phase cell cycle arrest (Khalil, 2012). Cells with mutant p53 demonstrate loss of the G1/S checkpoint. As discussed in Section 1.4.3, ATM is involved in the activation and stabilisation of p53 (Banin et al., 1998). This results in the expression of endogenous inhibitors of the cyclin-cdk complexes and G1 cell-cycle arrest (Khalil, 2012). ATM also activates the checkpoint kinases, Chk1 and Chk2, which are involved the stabilisation of p53 and also G2/M cell cycle arrest (Gatei et al., 2003). Previous studies have shown that in cells that have functional p53, inhibition of ATM prevents cells undergoing a p53-dependent apoptotic response, promoting a chemo-resistant phenotype. Conversely, in cells that are p53 deficient, ATM inhibition results in increased sensitivity to chemotherapy (Jiang et al., 2009).

Of the four glioma cell lines included in this study, the highest level of p53 protein expression was demonstrated in the LN18 cell line, using an antibody that detects both wild-type and mutant p53. The LN18 cell line is regarded in

the literature to be p53 mutant (Hossain et al., 2012, Pohl et al., 1999, Ishii et al., 1999) and is recorded as having a homozygous missense mutation in p53 at 343G>C and 713G>C in the Cosmic database (Forbes et al., 2015). The LN229 cell line is also recorded as having a missense mutation at 293C>T (Forbes et al., 2015), although there are conflicting reports in the literature as to whether this cell line should be regarded as p53 mutant or wild-type (Pohl et al., 1999, Ishii et al., 1999). Neither the A172 nor the U87MG cell lines have any documented mutations in the p53 gene and are regarded as p53 wild-type (Forbes et al., 2015, Ishii et al., 1999). Assessment of p21 expression may help to confirm the p53 mutational status of the cell lines, as this is activated by wild-type rather than mutant p53 (Elbendary et al., 1996). If the LN229 cell line is regarded as p53 mutant, then this would suggest that p53 status alone is not influencing sensitivity to ATM inhibitor treatment. If the LN229 cell line is regarded as p53 wild-type, then potentially the sensitivity of the LN18 cell line may be associated with p53 mutation. Additional cell lines, with varying p53 mutational status, would need to be assessed for sensitivity to ATM inhibitor treatment to explore this further.

5.3.5.3 PI3K pathway activity may influence response to ATM inhibitor treatment

PTEN is a key negative regulator of the PI3K signalling pathway. This is a multi-step, multi-molecular pathway known to be involved in both tumour development and outcome. Dysregulation of the PI3K signalling pathway is common in cancer, with mutations in various components recorded in a range of human tumours (Hafsi et al., 2012, Yuan and Cantley, 2008). Recently,

ATM has been shown to interact with several components of the PI3K pathway, including PTEN, PI3K and AKT (Irarrazabal et al., 2006, Chen et al., 2015, Viniegra et al., 2005). For example, ATM mediated phosphorylation of PTEN may promote nuclear translocation of PTEN in response to DNA damage (Chen et al., 2015). In addition, PI3K has been shown to activate ATM, which in turn can phosphorylate and activate AKT (Irarrazabal et al., 2006, Viniegra et al., 2005). It is therefore conceivable that expression of other components of the PI3K signalling pathway may influence response to ATM inhibitor treatment.

As discussed in Chapter 1, the PI3Ks are a family of receptor tyrosine kinases that have been found to be mutated in a variety of human tumours including glioblastoma (Liu et al., 2009, Samuels et al., 2004). The PI3Ks are responsible for converting PIP_2 to PIP_3 in response to the binding of extracellular ligands to cell surface receptors. It is this conversion of PIP_2 to PIP_3 that is antagonised by PTEN. The most extensively studied group of PI3Ks is the Class IA enzymes, which consist of a p110 catalytic subunit and a p85 regulatory subunit (Liu et al., 2009).

In this current study Western blot analysis for the p85 regulatory subunit has found expression to be absent in the LN18 cell line. The other three glioma cell lines express p85, with the highest levels seen in U87MG. The p85 subunit exists in three isoforms (α, β and γ) and in the steady state is bound to the catalytic subunit of PI3K known as p110. p85 acts to relocate the p110 catalytic subunit to the membrane where the membrane-associated phospholinositol lipids are located (Geering et al., 2007). It also binds to and

stabilises the p110 subunit, inhibiting its activation until receptor tyrosine kinase activation (Yu et al., 1998). Loss of p85 therefore results in constitutive PI3K activity (Miled et al., 2007) with expression of higher levels of PIP_3 and activation of AKT (Taniguchi et al., 2010). This results in an increased demand for both insulin and growth factors through the actions of the AKT pathway. On review of the Cosmic database (www.cancer.sanger.ac.uk) (Forbes et al., 2015), both AKT1 and AKT2 expression are documented to be elevated in the LN18 cell line but not in U87MG, A172 or LN229 cell lines. No mutations were recorded in either the AKT1 or AKT2 genes for any of the four cell lines included in the study (Forbes et al., 2015).

In addition to the role of ATM as a signal transducer of DNA damage, ATM has recently been shown to be present in the cytoplasm where it is responsible for the full activation of AKT in response to insulin. Li *et al* demonstrated that KU55933 can block the phosphorylation of AKT in cancer cells. This resulted in inhibition of cellular proliferation, induction of G1 cell cycle arrest and apoptosis. Interestingly, KU55933 was found to selectively target cancer cells with a high basal level of AKT and hence greater inhibition of cancer cell proliferation was seen in cell lines with over-activated AKT. In addition, KU55933 was found to down-regulate cyclin D1 by regulating translation, rather than decreasing transcription or facilitating degradation. KU55933 treatment resulted in dephosphorylation of the translational regulatory protein 4E-BP1. Hypophosphorylated 4E-BP1 binds more readily to the eIF-4E transcription factor, inhibiting cap-dependent translation (Li and Yang, 2010).

In this current study, decreased cellular proliferation, G1 cell cycle arrest and apoptosis were demonstrated in response to an ATM inhibitor in the LN18 cell line. As discussed earlier, if the LN18 cell line is confirmed to have higher levels of AKT, then the results seen in this study may be explained by KU55933 selectively targeting cells with over-active AKT. In the LN18 cell line the elevated levels of AKT would be predicted as a result of loss of expression of the p85 regulatory subunit of PI3K, resulting in constitutive activation of PI3K.

Clearly further work is required to substantiate this hypothesis. Firstly confirmation of p85 expression at the mRNA level is required. While p85 protein expression is essentially absent in the LN18 cell line, no mutations are recorded within the PI3R1 gene that encodes the p85 subunit (Forbes et al., 2015). This suggests that loss of expression of p85 is likely to be occurring at the post-transcriptional level.

Secondly, the Western blot analyses should be performed for cyclin D1, AKT and phosphorylated AKT expression for all 4 cell lines, both treated and untreated with KU55933. As PTEN is a negative regulator of the PI3K signalling pathway, PTEN loss has previously been shown to be associated with an increase in PI3K signalling and activation of AKT. If cells with over-activated AKT are sensitive to an ATM inhibitor then one could hypothesise that PTEN deficient cell lines should also be sensitive to an ATM inhibitor. Clearly this has not been demonstrated in this, and determination of AKT expression in the PTEN proficient A172 and U87MG cell lines may provide

further insight into the reasons behind this. Potentially, over-activated AKT alone may not be responsible for sensitivity to ATM inhibitor treatment.

Thirdly, further cell lines should be screened with varying AKT, PI3K, PTEN and p53 status for sensitivity to ATM inhibitor treatment. A full expression profile of components of the PI3K signalling pathway in these cell lines should be performed. This may confirm or refute the proposed hypothesis that p85 loss results in AKT activation and ATM inhibitor sensitivity and in addition identify other permutations of the PI3K signalling pathway that induce similar sensitivity.

5.3.6 Summary

This study does not demonstrate a synthetically lethal relationship between PTEN deficiency and ATM inhibitor treatment. It does, however, show that the PTEN proficient LN18 cell line is sensitive to ATM inhibitor treatment. This was evidenced by decreased cell survival, G1 cell cycle arrest and apoptosis in response to an ATM inhibitor. Interestingly, no increase in DNA damage was demonstrated. Neither the LN229 (PTEN proficient) or U87MG and A172 (both PTEN deficient) cell lines demonstrated sensitivity to ATM inhibitor treatment.

A previous study has shown that loss of p53 is associated with increased sensitivity to chemotherapy. The KU55933-sensitive LN18 cell line is known express mutant p53 and further work is required to determine the effects of p53 status on response to ATM inhibitor treatment. Another earlier study has shown that cells with over-activated AKT are more sensitive to an ATM

inhibitor and demonstrate G1 cell cycle arrest and apoptosis. Preliminary investigations demonstrated loss of expression of p85, the regulatory subunit of PI3K, in the LN18 cell line. Further work is required to determine whether this results in AKT activation and hence sensitivity to ATM inhibitor treatment with associated G1 cell cycle arrest and apoptosis.

CHAPTER 6

General discussion and conclusions

6 General discussion and conclusions

6.1 General discussion

DNA repair status has been shown to be involved in the development of cancer and to influence how a tumour responds to treatment. The most widely studied DNA repair gene in high grade glioma is MGMT, with high expression levels found to be associated with poor survival and resistance to treatment (Costa et al., 2010, Hegi et al., 2005, Rivera et al., 2010, Zhang et al., 2013). Few studies have investigated the wider effects of DNA repair gene expression in high grade glioma.

The first aim of this study was to investigate the expression of a panel of over 150 DNA repair genes on survival in high grade glioma (Chapter 3). By screening two large publically available gene expression datasets, 12 DNA repair genes have been identified that consistently influence survival in high grade glioma. After correction for multiple comparisons, multivariate analysis identified *APE1*, *MGMT, NBN*, *PMS2* and *PTEN* as five key DNA repair genes that influence prognosis in glioblastoma. Incorporation of the expression levels of these five genes into a prognostic index was found to separate glioblastoma patients into three prognostic groups, with the potential to identify subgroups of patients with either a particularly good or bad prognosis. While this prognostic index clearly needs to be validated in further large, prospective trials it may have the potential to identify a group of patients with a

particularly poor prognosis who could either be spared ineffective intensive treatment regimens or be targeted with new individualised treatment approaches.

In addition to interacting with each other, DNA repair genes also interact with genes from the wider genome. In order to explore further how *APE1*, *MGMT*, *NBN*, *PMS2* and *PTEN* interact with the wider genome, ANN analysis techniques were employed (Chapter 3). Using ANN analysis, further DNA repair genes were identified that may act as driver genes in glioblastoma and that may influence prognosis. Additionally, key hubs were identified within each network that may identify genes that interact with the DNA repair gene under investigation and may play a role in the malignant process. Of note *FOXG1*, *TGOLN2*, *HPRT*, *THRA* and *HIF1AN* expression were all found to be significantly associated with survival on Kaplan Meier analysis. All of these genes may warrant further investigation in future studies.

In Chapter 4 the prognostic significance of APE1, NBN, PMS2 and PTEN protein expression was explored in 61 high grade glioma patients by immunohistochemistry (Nottingham cohort). MGMT protein expression was not studied as this has been evaluated extensively in earlier studies. Both low APE1 and low nuclear PTEN expression were found to be associated with poor overall survival in high grade glioma. This raised the possibility of selectively targeting patients with low expression of one, or both, of these proteins with an individualised treatment approach. Neither NBN nor PMS2 protein expression were associated with survival in the Nottingham cohort.

PTEN deficiency has previously been linked to genomic instability and defective DSB repair in several, but not all, studies (Fraser et al., 2012, McEllin et al., 2010, Mendes-Pereira et al., 2009, Abbotts et al., 2014). This has highlighted the potential of targeting PTEN deficient tumours using a synthetic lethality approach. Both PTEN deficient colorectal and endometrial cell lines have be shown to demonstrate synthetic lethality when treated with PARP inhibitors (Mendes-Pereira et al., 2009), as have PTEN deficient melanoma cell lines treated with APE1 inhibitors (Abbotts et al., 2014). Similarly, PTEN deficient astrocytes have been shown to be deficient in HR and to be sensitive to PARP inhibition (McEllin et al., 2010). A more recent study has provided evidence that ATM inhibition in PTEN deficient colorectal cell lines can induce synthetic lethality (McCabe et al., 2015).

In Chapter 5, the potential of targeting PTEN deficient glioblastoma cell lines with the ATM inhibitor KU55933 was explored. This study was unable to demonstrate a synthetically lethal relationship between PTEN deficiency and ATM inhibitor treatment. Firstly, it was demonstrated that high ATM mRNA and protein expression was associated with poor overall survival, but only in patients with high levels of PTEN expression. Secondly, it was found that the LN18 PTEN proficient cell line was sensitive to treatment with an ATM inhibitor whereas another PTEN proficient (LN229) and two PTEN deficient (A172 and U87MG) cell lines were not sensitive. ATM inhibitor treatment in the LN18 cell line, however, did not result in significant amounts of DNA damage, but did cause an increase in G1/S cell cycle arrest and apoptosis. The generation of PTEN knock-down LN18 and LN229 cell lines did show that

knock-down of PTEN in the LN18 cell line resulted in a decrease in sensitivity to KU55933, but this did not reach statistical significance. PTEN status did not influence the response of the non-sensitive LN229 cell line to an ATM inhibitor. These findings would suggest that if PTEN status is influencing sensitivity of the LN18 cell line to ATM inhibitor treatment, it is not acting alone. While preliminary investigations have identified potential differences in p53 and PI3K status between the cell lines studied, further work is required to explore these associations further.

In conclusion, this study has identified key DNA repair genes that have prognostic significance in high grade glioma and in addition has explored the therapeutic potential to exploit a synthetically lethal strategy. While this study was unable to demonstrate a synthetically lethal relationship between PTEN deficiency and ATM inhibitor treatment, it did demonstrate that the PTEN proficient mutant LN18 cell line was sensitive to ATM inhibitor treatment. Further work is required to investigate the mechanisms behind this sensitivity.

6.2 Summary of key findings

1. After screening an extensive panel of DNA repair genes, expression levels of five (*APE1*, *MGMT*, *NBN*, *PMS2* and *PTEN*) genes are identified as having the greatest influence on survival in glioblastoma.
2. A DNA repair prognostic index that incorporates the mRNA expression levels of *APE1*, *MGMT*, *NBN*, *PMS2* and *PTEN* can separate glioblastoma patients into three prognostic groups.
3. Artificial neural network analysis for *APE1*, *MGMT*, *NBN*, *PMS2* and *PTEN* reveals potential driver genes in glioblastoma, including several which also have prognostic significance and may warrant further investigation in future studies.
4. At the protein level both low expression of PTEN and low expression of APE1 are associated with poor overall survival in high grade glioma, suggesting that tumours with low levels of these proteins may be amenable to targeting with a synthetic lethality approach.
5. PTEN deficient glioblastoma cell lines are not sensitive to ATM inhibitor treatment using a synthetic lethality approach; however, the LN18 (PTEN proficient) glioblastoma cell line is sensitive to ATM inhibitor treatment and raises the possibility that other synthetic lethality interactions may exist with ATM.

6.3 Future developments

6.3.1 Further evaluation of the DNA repair gene prognostic index

In Chapter 3 a DNA repair gene prognostic index has been developed that can separate glioblastoma patients into 3 prognostic groups. This was developed based on two large gene expression datasets analysed retrospectively. One problem with these datasets was a lack of easily interpretable treatment data. Ideally, this prognostic index should be tested in a large, prospective cohort with the collection of good quality clinical and treatment data. This would enable the index to be tested in different subgroups, such as those that received chemo-radiotherapy. It would also be possible to assess whether the index has predictive significance, for example to identify patients that are more likely to respond to chemo-radiotherapy.

6.3.2 Construction of a glioma TMA

One significant limitation of the protein expression analysis component of this study is the small size of the Nottingham cohort. Ideally, tissue should be collected (prior to any treatment) from a large cohort to glioma patients, enabling both mRNA and protein expression studies to be performed on the same individual. Additionally, inclusion of patients of all grades would enable differences in DNA repair status to be explored between low and high grade gliomas. Prospective collection of clinical and treatment data would again enable subgroup analysis and evaluation of a gene's predictive significance.

6.3.3 Further evaluation of selective genes identified from ANN

Artificial neural network analysis has generated a wealth of additional data, including the identification of additional genes (including DNA repair genes) that may play an important role in GBM. These genes could be further investigated at the protein level to evaluate their expression pattern in GBM and also their prognostic and predictive significance. If these proteins are found to have prognostic and/or predictive significance then this could lead to cell line studies examining the effect of gene knock-down or inhibition. Additionally, the potential to selectively target tumours, dependent on the status of the gene of interest, with chemotherapy or radiotherapy could also be explored.

FOXG1 was identified as one of the key hubs from the ANN analysis. This has already led on to a BMedSci project, and publication, investigating the prognostic significance of the FOX gene family in GBM. This additionally identified *FOXA2*, *FOXN2* and *FOXN3* to have prognostic significance in GBM (Robertson et al., 2015). As stated previously, the next step would be to investigate these genes at the protein level, leading on to potential cell line studies.

6.3.4 Unravelling the mechanism behind the sensitivity of the LN18 cell line to ATM inhibitor treatment

The LN18 (PTEN proficient) cell line has been shown to be sensitive to an ATM inhibitor and this does not appear to result in an increase in DNA damage. A G1/S cell cycle arrest does appear to be present in the LN18 cell

line in response to ATM inhibitor treatment. Assessment of cyclin D1 levels in cell lines both untreated and treated with KU55933 should be performed to provide evidence to support a G1/S cell cycle arrest. Cyclin D1 levels would be expected to be lower in PTEN proficient cell lines and in those demonstrating G1/S cell cycle arrest. To evaluate the hypothesis that sensitivity to ATM inhibitor treatment is as a consequence of over-activated AKT, levels of both total and phosphorylated AKT should be assessed in both untreated and treated cell lines. If decreased cyclin D1 levels are confirmed in the LN18 cell line in response to an ATM inhibitor, further work would be required to identify the mechanism behind this. In the study by Li *et al*, down-regulation of translation was found to be responsible for the lower levels of cyclin D1 in response to an ATM inhibitor by dephosphorylation of the translational regulatory protein 4E-BP1. To evaluate this, total 4E-BP1 levels should be determined in cells treated and untreated with KU55933. On western blot analysis, a decrease in 4E-BP1 isoforms would be predicted in the LN18 cell line treated with KU55933. eIF-4E/4E-BP1 complexes could be isolated from both untreated and treated cell lines and the amount of 4E-BP1 bound to eIF-4E determined. LN18 cells treated with KU55933 would be expected to have increased levels of 4E-BP1 bound to eIF-4E. Finally, to determine whether cyclin D1 is down-regulated by a decrease in translation, newly synthesised levels of cyclin D1 could be measured in cells both treated and untreated with KU55933.

The results obtained in this study should be confirmed in additional cell lines and with an alternative ATM inhibitor, such as KU60019 (Golding et al.,

2009). Inhibition of ATM should be confirmed by western blot analysis in treated cells. The cell lines used should be fully profiled for expression of components of the PI3K pathway and also p53 status. This would enable the identification of genetic changes in common between the cell lines sensitive to an ATM inhibitor. For example, if loss of p85 expression is found to be present in cell lines sensitive to ATM inhibitors, cell line studies could be performed to knock-down the p85 gene, or with the addition of a p85 inhibitor.

Although no DNA damage was demonstrated in this study in response to an ATM inhibitor, it would be interesting to assess the addition of an ATM inhibitor to conventional treatments for GBM. ATM inhibition has previously been shown to increase radiation sensitivity of glioblastoma cell lines (Biddlestone-Thorpe et al., 2013, Golding et al., 2009, Golding et al., 2012). Similarly, ATM inhibition has been shown to increase temozolomide response (Nadkarni et al., 2012). ATM inhibition, in combination with radiation and/or chemotherapy, may therefore have the potential to increase sensitivity to these agents further.

Finally, any significant findings in the cell line studies would need to be confirmed in xenograft studies, prior to early stage clinical trials.

CHAPTER 7

References

7 References

Abbotts, R. 2013. Therapeutic potential of human apurinic/apyrimidinic endomuclease 1 (APE1) modulation in PTEN-deficient melanoma. PhD, University of Nottingham.

Abbotts R, Fischer PM, Laughton CA, Maloney DJ, Rai G, Jadhav A, et al. Synthetic lethal targeting of homologous recombination repair defect in PTEN-deficienct melanoma cell lines using inhibitors of human apurinic/apyrimidinic endonuclease (APE1). Genomic Instability and DNA Repair Conference, 2013 Banff, Canada.

Abbotts, R., Jewell, R., Nsengimana, J., Maloney, D. J., Simeonov, A., Seedhouse, C., et al. 2014. Targeting human apurinic/apyrimidinic endonuclease 1 (APE1) in phosphatase and tensin homolog (PTEN) deficient melanoma cells for personalized therapy. *Oncotarget,* 5**,** 3273-86.

Abbotts, R. & Madhusudan, S. 2010. Human AP endonuclease 1 (APE1): from mechanistic insights to druggable target in cancer. *Cancer Treat Rev,* 36**,** 425-35.

Abdel-Fatah, T. M., Arora, A., Alsubhi, N., Agarwal, D., Moseley, P. M., Perry, C., et al. 2014a. Clinicopathological significance of ATM-Chk2 expression in sporadic breast cancers: a comprehensive analysis in large cohorts. *Neoplasia,* 16**,** 982-91.

Abdel-Fatah, T. M., Arora, A., Moseley, P. M., Perry, C., Rakha, E. A., Green, A. R., et al. 2015. DNA repair prognostic index modelling reveals an essential role for base excision repair in influencing clinical outcomes in ER negative and triple negative breast cancers. *Oncotarget*.

Abdel-Fatah, T. M., Perry, C., Moseley, P., Johnson, K., Arora, A., Chan, S., et al. 2014b. Clinicopathological significance of human apurinic/apyrimidinic endonuclease 1 (APE1) expression in oestrogen-receptor-positive breast cancer. *Breast cancer research and treatment,* 143**,** 411-21.

Abdel-Fatah, T. M., Perry, C., Moseley, P., Johnson, K., Arora, A., Chan, S., et al. 2014c. Clinicopathological significance of human apurinic/apyrimidinic endonuclease 1 (APE1) expression in oestrogen-receptor-positive breast cancer. *Breast Cancer Res Treat,* 143**,** 411-21.

Abraham, R. T. 2001. Cell cycle checkpoint signaling through the ATM and ATR kinases. *Genes Dev,* 15**,** 2177-96.

Adel Fahmideh, M., Schwartzbaum, J., Frumento, P. & Feychting, M. 2014. Association between DNA repair gene polymorphisms and risk of glioma: a systematic review and meta-analysis. *Neuro Oncol,* 16**,** 807-14.

Agachan, B., Kucukhuseyin, O., Aksoy, P., Turna, A., Yaylim, I., Gormus, U., et al. 2009. Apurinic/apyrimidinic endonuclease (APE1) gene polymorphisms and lung cancer risk in relation to tobacco smoking. *Anticancer Res,* 29**,** 2417-20.

Ai, L., Vo, Q. N., Zuo, C., Li, L., Ling, W., Suen, J. Y., et al. 2004. Ataxia-Telangiectasia-Mutated (ATM) Gene in Head and Neck Squamous Cell Carcinoma: Promoter Hypermethylation with Clinical Correlation in 100 Cases. *Cancer Epidemiology Biomarkers & Prevention,* 13**,** 150-156.

Akamatsu, Y., Ohno, T., Hirota, K., Kagoshima, H., Yodoi, J. & Shigesada, K. 1997. Redox regulation of the DNA binding activity in transcription factor PEBP2. The roles of two conserved cysteine residues. *J Biol Chem,* 272**,** 14497-500.

Al-Attar, A., Gossage, L., Fareed, K. R., Shehata, M., Mohammed, M., Zaitoun, A. M., et al. 2010. Human apurinic/apyrimidinic endonuclease (APE1) is a prognostic factor in ovarian, gastro-oesophageal and pancreatico-biliary cancers. *Br J Cancer,* 102**,** 704-9.

Albarakati, N., Abdel-Fatah, T. M., Doherty, R., Russell, R., Agarwal, D., Moseley, P., et al. 2015. Targeting BRCA1-BER deficient breast cancer by ATM or DNA-PKcs blockade either alone or in combination with cisplatin for personalized therapy. *Molecular Oncology,* 9**,** 204-17.

Altaner, C. 2008. Glioblastoma and stem cells. *Neoplasma,* 55**,** 369-74.

Andrew, A. S., Karagas, M. R., Nelson, H. H., Guarrera, S., Polidoro, S., Gamberini, S., et al. 2008. DNA repair polymorphisms modify bladder cancer risk: a multi-factor analytic strategy. *Human Heredity,* 65**,** 105-18.

Angele, S., Falconer, A., Edwards, S. M., Dork, T., Bremer, M., Moullan, N., et al. 2004. ATM polymorphisms as risk factors for prostate cancer development. *Br J Cancer,* 91**,** 783-7.

Arora, A., Abdel-Fatah, T. M., Agarwal, D., Doherty, R., Moseley, P. M., Aleskandarany, M. A., et al. 2015. Transcriptomic and Protein Expression Analysis Reveals Clinicopathological Significance of Bloom Syndrome Helicase (BLM) in Breast Cancer. *Mol Cancer Ther,* 14**,** 1057-65.

Arriagada, R., Dunant, A., Pignon, J. P., Bergman, B., Chabowski, M., Grunenwald, D., et al. 2010. Long-term results of the international adjuvant lung cancer trial evaluating adjuvant Cisplatin-based chemotherapy in resected lung cancer. *J Clin Oncol,* 28**,** 35-42.

Ashworth, A. 2008. Drug resistance caused by reversion mutation. *Cancer Res,* 68**,** 10021-3.

Austen, B., Powell, J. E., Alvi, A., Edwards, I., Hooper, L., Starczynski, J., et al. 2005. Mutations in the ATM gene lead to impaired overall and treatment-free survival that is independent of IGVH mutation status in patients with B-CLL. *Blood,* 106**,** 3175-82.

Auvergne, R. M., Sim, F. J., Wang, S., Chandler-Militello, D., Burch, J., Al Fanek, Y., et al. 2013. Transcriptional differences between normal and glioma-derived glial progenitor cells identify a core set of dysregulated genes. *Cell Rep,* 3**,** 2127-41.

Bader, A. G., Kang, S., Zhao, L. & Vogt, P. K. 2005. Oncogenic PI3K deregulates transcription and translation. *Nat Rev Cancer,* 5**,** 921-9.

Bakkenist, C. J. & Kastan, M. B. 2003. DNA damage activates ATM through intermolecular autophosphorylation and dimer dissociation. *Nature,* 421**,** 499-506.

Banin, S., Moyal, L., Shieh, S., Taya, Y., Anderson, C. W., Chessa, L., et al. 1998. Enhanced phosphorylation of p53 by ATM in response to DNA damage. *Science,* 281**,** 1674-7.

Bao, S., Wu, Q., McLendon, R. E., Hao, Y., Shi, Q., Hjelmeland, A. B., et al. 2006. Glioma stem cells promote radioresistance by preferential activation of the DNA damage response. *Nature,* 444**,** 756-60.

Barvaux, V. A., Lorigan, P., Ranson, M., Gillum, A. M., McElhinney, R. S., McMurry, T. B. & Margison, G. P. 2004. Sensitization of a human ovarian cancer cell line to temozolomide by simultaneous attenuation of the Bcl-2 antiapoptotic protein and DNA repair by O6-alkylguanine-DNA alkyltransferase. *Mol Cancer Ther,* 3**,** 1215-20.

Bassi, C., Ho, J., Srikumar, T., Dowling, R. J., Gorrini, C., Miller, S. J., et al. 2013. Nuclear PTEN controls DNA repair and sensitivity to genotoxic stress. *Science,* 341**,** 395-9.

Bau, D. T., Chang, C. H., Tsai, M. H., Chiu, C. F., Tsou, Y. A., Wang, R. F., et al. 2010. Association between DNA repair gene ATM polymorphisms and oral cancer susceptibility. *Laryngoscope,* 120**,** 2417-22.

Bauman, G. S., Ino, Y., Ueki, K., Zlatescu, M. C., Fisher, B. J., Macdonald, D. R., et al. 2000. Allelic loss of chromosome 1p and radiotherapy plus chemotherapy in patients with oligodendrogliomas. *Int J Radiat Oncol Biol Phys,* 48**,** 825-30.

Behin, A., Hoang-Xuan, K., Carpentier, A. F. & Delattre, J.-Y. 2003. Primary brain tumours in adults. *The Lancet,* 361**,** 323-331.

Bellacosa, A. 2001. Role of MED1 (MBD4) Gene in DNA repair and human cancer. *J Cell Physiol,* 187**,** 137-44.

Benjamini, Y., Hochberg, Y. 1995. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological),* 57**,** 289-300.

Bertolini, F., Zunarelli, E., Baraldi, C., Valentini, A., Del Giovane, C., Depenni, R., et al. 2012. Survival in patients with newly diagnosed conventional glioblastoma: a modified prognostic score based on a single-institution series. *Tumori,* 98**,** 756-61.

Bhat, S. S., Schmidt, K. R., Ladd, S., Kim, K. C., Schwartz, C. E., Simensen, R. J., et al. 2006. Disruption of DMD and deletion of ACSL4 causing developmental delay, hypotonia, and multiple congenital anomalies. *Cytogenetic and Genome Research,* 112**,** 170-5.

Biddlestone-Thorpe, L., Sajjad, M., Rosenberg, E., Beckta, J. M., Valerie, N. C., Tokarz, M., et al. 2013. ATM kinase inhibition preferentially sensitizes p53-mutant glioma to ionizing radiation. *Clin Cancer Res,* 19**,** 3189-200.

Bobola, M. S., Blank, A., Berger, M. S., Stevens, B. A. & Silber, J. R. 2001. Apurinic/Apyrimidinic endonuclease activity is elevated in human adult gliomas. *Clin Cancer Res,* 7**,** 3510-8.

Bobola, M. S., Emond, M. J., Blank, A., Meade, E. H., Kolstoe, D. D., Berger, M. S., et al. 2004. Apurinic endonuclease activity in adult gliomas and time to tumor progression after alkylating agent-based chemotherapy and after radiotherapy. *Clin Cancer Res,* 10**,** 7875-83.

Boland, C. R. & Goel, A. 2010. Microsatellite Instability in Colorectal Cancer. *Gastroenterology,* 138**,** 2073-2087 e3.

Boxem, M. 2006. Cyclin-dependent kinases in C. elegans. *Cell Div,* 1**,** 6.
Brown, M. & Wittwer, C. 2000. Flow Cytometry: Principles and Clinical Applications in Hematology. *Clinical chemistry,* 46**,** 1221-1229.
Bruggisser, R., von Daeniken, K., Jundt, G., Schaffner, W. & Tullberg-Reinert, H. 2002. Interference of plant extracts, phytoestrogens and antioxidants with the MTT tetrazolium assay. *Planta Med,* 68**,** 445-8.
Bryant, H. E., Schultz, N., Thomas, H. D., Parker, K. M., Flower, D., Lopez, E., et al. 2005. Specific killing of BRCA2-deficient tumours with inhibitors of poly(ADP-ribose) polymerase. *Nature,* 434**,** 913-7.
Burma, S., Chen, B. P., Murphy, M., Kurimasa, A. & Chen, D. J. 2001. ATM phosphorylates histone H2AX in response to DNA double-strand breaks. *J Biol Chem,* 276**,** 42462-7.
Cairncross, G., Berkey, B., Shaw, E., Jenkins, R., Scheithauer, B., Brachman, D., et al. 2006. Phase III trial of chemotherapy plus radiotherapy compared with radiotherapy alone for pure and mixed anaplastic oligodendroglioma: Intergroup Radiation Therapy Oncology Group Trial 9402. *J Clin Oncol,* 24**,** 2707-14.
Cairncross, J. G., Ueki, K., Zlatescu, M. C., Lisle, D. K., Finkelstein, D. M., Hammond, R. R., et al. 1998. Specific genetic predictors of chemotherapeutic response and survival in patients with anaplastic oligodendrogliomas. *J Natl Cancer Inst,* 90**,** 1473-9.
Cairns, P., Okami, K., Halachmi, S., Halachmi, N., Esteller, M., Herman, J. G., et al. 1997. Frequent Inactivation of PTEN/MMAC1 in Primary Prostate Cancer. *Cancer Research,* 57**,** 4997-5000.
Caldecott, K. W. 2008. Single-strand break repair and genetic disease. *Nat Rev Genet,* 9**,** 619-631.
Canbay, E., Agachan, B., Gulluoglu, M., Isbir, T., Balik, E., Yamaner, S., et al. 2010. Possible associations of APE1 polymorphism with susceptibility and HOGG1 polymorphism with prognosis in gastric cancer. *Anticancer Res,* 30**,** 1359-64.
Canbay, E., Cakmakoglu, B., Zeybek, U., Sozen, S., Cacina, C., Gulluoglu, M., et al. 2011. Association of APE1 and hOGG1 polymorphisms with colorectal cancer risk in a Turkish population. *Curr Med Res Opin,* 27**,** 1295-302.
Cao, X., Kambe, F., Ohmori, S. & Seo, H. 2002. Oxidoreductive modification of two cysteine residues in paired domain by Ref-1 regulates DNA-binding activity of Pax-8. *Biochem Biophys Res Commun,* 297**,** 288-93.
Carney, J. P., Maser, R. S., Olivares, H., Davis, E. M., Le Beau, M., Yates, J. R., 3rd, et al. 1998. The hMre11/hRad50 protein complex and Nijmegen breakage syndrome: linkage of double-strand break repair to the cellular DNA damage response. *Cell,* 93**,** 477-86.
Chalmers, A., Jackson A, Swaisland H, Stewart W, Halford S, Molife R, Hargrave D, McCormick A. Olaparib penetrates glioblastoma at therapeutic levels: results of stage 1 of the OPARATIC trial; a phase I study of olaparib in combination with temozolomide in patients with relapsed glioblastoma. NCRI Conference, 2014 Liverpool.
Chang, C. J., Mulholland, D. J., Valamehr, B., Mosessian, S., Sellers, W. R. & Wu, H. 2008. PTEN nuclear localization is regulated by oxidative stress

and mediates p53-dependent tumor suppression. *Mol Cell Biol,* 28**,** 3281-9.
Chen, J. H., Zhang, P., Chen, W. D., Li, D. D., Wu, X. Q., Deng, R., et al. 2015. ATM-mediated PTEN phosphorylation promotes PTEN nuclear translocation and autophagy in response to DNA-damaging agents in cancer cells. *Autophagy,* 11**,** 239-52.
Chen, P. C., Lavin, M. F., Kidson, C. & Moss, D. 1978. Identification of ataxia telangiectasia heterozygotes, a cancer prone population. *Nature,* 274**,** 484-6.
Chin, L., Andersen, J. N. & Futreal, P. A. 2011a. Cancer genomics: from discovery science to personalized medicine. *Nat Med,* 17**,** 297-303.
Chin, L., Hahn, W. C., Getz, G. & Meyerson, M. 2011b. Making sense of cancer genomic data. *Genes Dev,* 25**,** 534-55.
Chung, J. H., Ginn-Pease, M. E. & Eng, C. 2005. Phosphatase and tensin homologue deleted on chromosome 10 (PTEN) has nuclear localization signal-like sequences for nuclear import mediated by major vault protein. *Cancer Res,* 65**,** 4108-16.
Cleary, S. P., Zhang, W., Di Nicola, N., Aronson, M., Aube, J., Steinman, A., et al. 2003. Heterozygosity for the BLM(Ash) mutation and cancer risk. *Cancer Res,* 63**,** 1769-71.
Collins, A. R., El Yamani, N., Lorenzo, Y., Shaposhnikov, S., Brunborg, G. & Azqueta, A. 2014. Controlling variation in the comet assay. *Front Genet,* 5**,** 359.
Collins, A. R., Oscoz, A. A., Brunborg, G., Gaivao, I., Giovannelli, L., Kruszewski, M., et al. 2008. The comet assay: topical issues. *Mutagenesis,* 23**,** 143-51.
Coons, A. H., Creech, H. J., Jones, R. N. & Berliner, E. 1942. The Demonstration of Pneumococcal Antigen in Tissues by the Use of Fluorescent Antibody. *The Journal of Immunology,* 45**,** 159-170.
Costa, B. M., Caeiro, C., Guimaraes, I., Martinho, O., Jaraquemada, T., Augusto, I., et al. 2010. Prognostic value of MGMT promoter methylation in glioblastoma patients treated with temozolomide-based chemoradiation: a Portuguese multicentre study. *Oncol Rep,* 23**,** 1655-62.
Croteau, D., Bohr, VA 2013. Overview of DNA repair pathways. *In:* Madhusudan, S. & Wilson, D. (eds.) *DNA Repair and Cancer - From Bench to Clinic.* 1st ed. Boca Raton: CRC Press.
Cancer Research UK. 2011. *Breast cancer statistics - UK* [Online]. Available at: http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/breast-cancer [Accessed November 8th 2015].
Cancer Research UK. 2012. *Statistics and outlook for brain tumours* [Online]. Available at: www.cancerresearchuk.org/cancer-help/type/brain-tumour/treatment/statistics-and-outlook-for-brain-tumours [Accessed 17th April 2013].
Cancer Research UK. 2013. *Brain, other CNS and intracranial tumours incidence statistics* [Online]. Available at: www.cancerresearchuk.org/cancer-info/cancerstats/types/brain/incidence [Accessed 14th April 2013].

Cully, M., You, H., Levine, A. J. & Mak, T. W. 2006. Beyond PTEN mutations: the PI3K pathway as an integrator of multiple inputs during tumorigenesis. *Nat Rev Cancer,* 6**,** 184-92.

Curtin, N. 2013. PARP and PARP Inhibitor Therapeutics. *In:* Madhusudan, S. & Wilson, D. (eds.) *DNA Repair and Cancer: From Bench to Clinic.* 1st ed. Boca Raton: CRC Press.

Curtin, N. J., Mukhopadhyay, A., Drew, Y., Plummer, R. 2012. The role of PARP in DNA repair and its therapeutic exploitation. *In:* Kelley, M. R. (ed.) *DNA Repair in Cancer Therapy: Molecular Targets and Clinical Applications.* 1st ed. London: Academic Press.

Danesin, C. & Houart, C. 2012. A Fox stops the Wnt: implications for forebrain development and diseases. *Current Opinion in Genetics & Development,* 22**,** 323-30.

Datta, S. R., Dudek, H., Tao, X., Masters, S., Fu, H., Gotoh, Y. & Greenberg, M. E. 1997. Akt phosphorylation of BAD couples survival signals to the cell-intrinsic death machinery. *Cell,* 91**,** 231-41.

Davar, D., Beumer, J. H., Hamieh, L. & Tawbi, H. 2012. Role of PARP inhibitors in cancer biology and therapy. *Curr Med Chem,* 19**,** 3907-21.

Davis, F. B., Tang, H. Y., Shih, A., Keating, T., Lansing, L., Hercbergs, A., et al. 2006. Acting via a cell surface receptor, thyroid hormone is a growth factor for glioma cells. *Cancer research,* 66**,** 7270-5.

Deb, S., Xu, H., Tuynman, J., George, J., Yan, Y., Li, J., et al. 2014. RAD21 cohesin overexpression is a prognostic and predictive marker exacerbating poor prognosis in KRAS mutant colorectal carcinomas. *Br J Cancer,* 110**,** 1606-1613.

Debiec-Rychter, M., Dumez, H., Judson, I., Wasag, B., Verweij, J., Brown, M., et al. 2004. Use of c-KIT/PDGFRA mutational analysis to predict the clinical response to imatinib in patients with advanced gastrointestinal stromal tumours entered on phase I and II studies of the EORTC Soft Tissue and Bone Sarcoma Group. *European Journal of Cancer,* 40**,** 689-95.

Deng, Q., Sheng, L., Su, D., Zhang, L., Liu, P., Lu, K. & Ma, S. 2011. Genetic polymorphisms in ATM, ERCC1, APE1 and iASPP genes and lung cancer risk in a population of southeast China. *Med Oncol,* 28**,** 667-72.

Deng, S., Zhang, P., Zeng, H., Wang, W., Jin, T., Wang, J. & Dong, Q. 2014. Factor-inhibiting hypoxia-inducible factor expression in patients with high-risk locally advanced renal cell carcinoma and its relationship with tumor progression. *Kaohsiung J Med Sci,* 30**,** 12-9.

Devarakonda, S., Morgensztern, D. & Govindan, R. 2013. Clinical applications of The Cancer Genome Atlas project (TCGA) for squamous cell lung carcinoma. *Oncology (Williston Park),* 27**,** 899-906.

Di Cristofano, A., Pesce, B., Cordon-Cardo, C. & Pandolfi, P. P. 1998. Pten is essential for embryonic development and tumour suppression. *Nat Genet,* 19**,** 348-55.

Diao, L. & Chen, Y.-G. 2007. PTEN, a general negative regulator of cyclin D expression. *Cell Research,* 17**,** 291-292.

Digweed, M. & Sperling, K. 2004. Nijmegen breakage syndrome: clinical manifestation of defective response to DNA double-strand breaks. *DNA Repair,* 3**,** 1207-1217.

Diserens, A. C., de Tribolet, N., Martin-Achard, A., Gaide, A. C., Schnegg, J. F. & Carrel, S. 1981. Characterization of an established human malignant glioma cell line: LN-18. *Acta Neuropathol,* 53**,** 21-8.

Doucette, T., Rao, G., Rao, A., Shen, L., Aldape, K., Wei, J., et al. 2013. Immune heterogeneity of glioblastoma subtypes: extrapolation from the cancer genome atlas. *Cancer Immunol Res,* 1**,** 112-22.

Durrant, L. G., Chapman, M. A., Buckley, D. J., Spendlove, I., Robins, R. A. & Armitage, N. C. 2003. Enhanced expression of the complement regulatory protein CD55 predicts a poor prognosis in colorectal cancer patients. *Cancer Immunol Immunother,* 52**,** 638-42.

EBCTCG 2005. Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials. *The Lancet,* 365**,** 1687-1717.

Edwards, S. L., Brough, R., Lord, C. J., Natrajan, R., Vatcheva, R., Levine, D. A., et al. 2008. Resistance to therapy caused by intragenic deletion in BRCA2. *Nature,* 451**,** 1111-U8.

Eker, A. P., Quayle, C., Chaves, I. & van der Horst, G. T. 2009. DNA repair in mammalian cells: Direct DNA damage reversal: elegant solutions for nasty problems. *Cell Mol Life Sci,* 66**,** 968-80.

Elbendary, A. A., Cirisano, F. D., Evans, A. C., Jr., Davis, P. L., Iglehart, J. D., Marks, J. R. & Berchuck, A. 1996. Relationship between p21 expression and mutation of the p53 tumor suppressor gene in normal and malignant ovarian epithelial cells. *Clin Cancer Res,* 2**,** 1571-5.

Ema, M., Hirota, K., Mimura, J., Abe, H., Yodoi, J., Sogawa, K., et al. 1999. Molecular mechanisms of transcription activation by HLF and HIF1alpha in response to hypoxia: their stabilization and redox signal-induced interaction with CBP/p300. *EMBO J,* 18**,** 1905-14.

Eng, C. 2003. PTEN: one gene, many syndromes. *Hum Mutat,* 22**,** 183-98.

Engelman, J. A., Luo, J. & Cantley, L. C. 2006. The evolution of phosphatidylinositol 3-kinases as regulators of growth and metabolism. *Nat Rev Genet,* 7**,** 606-19.

Ermoian, R. P., Furniss, C. S., Lamborn, K. R., Basila, D., Berger, M. S., Gottschalk, A. R., et al. 2002a. Dysregulation of PTEN and Protein Kinase B Is Associated with Glioma Histology and Patient Survival. *Clinical Cancer Research,* 8**,** 1100-1106.

Ermoian, R. P., Furniss, C. S., Lamborn, K. R., Basila, D., Berger, M. S., Gottschalk, A. R., et al. 2002b. Dysregulation of PTEN and protein kinase B is associated with glioma histology and patient survival. *Clin Cancer Res,* 8**,** 1100-6.

Fantini, D., Vascotto, C., Deganuto, M., Bivi, N., Gustincich, S., Marcon, G., et al. 2008. APE1/Ref-1 regulates PTEN expression mediated by Egr-1. *Free radical research,* 42**,** 20-9.

Farabegoli, F., Barbi, C., Lambertini, E. & Piva, R. 2007. (-)-Epigallocatechin-3-gallate downregulates estrogen receptor alpha function in MCF-7 breast carcinoma cells. *Cancer Detection and Prevention,* 31**,** 499-504.

Farmer, H., McCabe, N., Lord, C. J., Tutt, A. N., Johnson, D. A., Richardson, T. B., et al. 2005. Targeting the DNA repair defect in BRCA mutant cells as a therapeutic strategy. *Nature,* 434**,** 917-21.

Felsberg, J., Thon, N., Eigenbrod, S., Hentschel, B., Sabel, M. C., Westphal, M., et al. 2011. Promoter methylation and expression of MGMT and the DNA mismatch repair genes MLH1, MSH2, MSH6 and PMS2 in paired primary and recurrent glioblastomas. *International Journal of Cancer,* 129**,** 659-670.

Fernandez-Marcelo, T., Frias, C., Pascua, I., de Juan, C., Head, J., Gomez, A., et al. 2014. Poly (ADP-ribose) polymerase 3 (PARP3), a potential repressor of telomerase activity. *J Exp Clin Cancer Res,* 33**,** 19.

Fernet, M., Chiker, S, Hall, J 2013. ATM/ATR cell cycle checkpoints: Mechanisms and manipulation in cancer therapy. *In:* Madhusudan, S. & Wilson, D. (eds.) *DNA Repair and Cancer: From Bench to Clinic.* 1st ed. Boca Raton: CRC Press.

Fingar, D. C., Salama, S., Tsou, C., Harlow, E. & Blenis, J. 2002. Mammalian cell size is controlled by mTOR and its downstream targets S6K1 and 4EBP1/eIF4E. *Genes Dev,* 16**,** 1472-87.

Fischer, U. & Meese, E. 2007. Glioblastoma multiforme: the role of DSB repair between genotype and phenotype. *Oncogene,* 26**,** 7809-15.

Fishel, M., Vascotto, C., Kelley, M. R. 2013. DNA Base Excision Repair Therapeutics: Summary of Targets with a focus on APE1. *In:* Madhusudan S & Wilson, D. (eds.) *DNA Repair and Cancer: From Bench to Clinic.* 1st ed. Boca Raton: CRC Press.

Forbes, S. A., Beare, D., Gunasekaran, P., Leung, K., Bindal, N., Boutselakis, H., et al. 2015. COSMIC: exploring the world's knowledge of somatic mutations in human cancer. *Nucleic Acids Res,* 43**,** D805-11.

Fountzilas, E., Kotoula, V., Angouridakis, N., Karasmanis, I., Wirtz, R. M., Eleftheraki, A. G., et al. 2013. Identification and validation of a multigene predictor of recurrence in primary laryngeal cancer. *PLoS One,* 8**,** e70429.

Fragkos, M., Jurvansuu, J. & Beard, P. 2009. H2AX is required for cell cycle arrest via the p53/p21 pathway. *Mol Cell Biol,* 29**,** 2828-40.

Fraser, M., Zhao, H., Luoto, K. R., Lundin, C., Coackley, C., Chan, N., et al. 2012. PTEN deletion in prostate cancer cells does not associate with loss of RAD51 function: implications for radiotherapy and chemotherapy. *Clin Cancer Res,* 18**,** 1015-27.

Freeman, D. J., Li, A. G., Wei, G., Li, H. H., Kertesz, N., Lesche, R., et al. 2003. PTEN tumor suppressor regulates p53 protein levels and activity through phosphatase-dependent and -independent mechanisms. *Cancer Cell,* 3**,** 117-30.

Fruman, D. A., Meyers, R. E. & Cantley, L. C. 1998. Phosphoinositide kinases. *Annual Review of Biochemistry,* 67**,** 481-507.

Fruman, D. A. & Rommel, C. 2014. PI3K and cancer: lessons, challenges and opportunities. *Nat Rev Drug Discov,* 13**,** 140-56.

Gangwar, R., Ahirwar, D., Mandhani, A. & Mittal, R. D. 2009. Influence of XPD and APE1 DNA Repair Gene Polymorphism on Bladder Cancer Susceptibility in North India. *Urology,* 73**,** 675-80.

Gannon, H. S., Woda, B. A. & Jones, S. N. 2012. ATM phosphorylation of Mdm2 Ser394 regulates the amplitude and duration of the DNA damage response in mice. *Cancer Cell,* 21**,** 668-79.

Gatei, M., Sloper, K., Sorensen, C., Syljuasen, R., Falck, J., Hobson, K., et al. 2003. Ataxia-telangiectasia-mutated (ATM) and NBS1-dependent phosphorylation of Chk1 on Ser-317 in response to ionizing radiation. *J Biol Chem,* 278**,** 14806-11.

Gatti, R. A., Berkel, I., Boder, E., Braedt, G., Charmley, P., Concannon, P., et al. 1988. Localization of an ataxia-telangiectasia gene to chromosome 11q22-23. *Nature,* 336**,** 577-80.

Geering, B., Cutillas, P. R., Nock, G., Gharbi, S. I. & Vanhaesebroeck, B. 2007. Class IA phosphoinositide 3-kinases are obligate p85-p110 heterodimers. *Proc Natl Acad Sci U S A,* 104**,** 7809-14.

GeneCardsSuite. 2015. *PTER Gene (Protein Coding)* [Online]. Available at: http://www.genecards.org/cgi-bin/carddisp.pl?gene=PTER [Accessed 31st August 2015].

Genschel, J. & Modrich, P. 2003. Mechanism of 5'-directed excision in human mismatch repair. *Molecular Cell,* 12**,** 1077-86.

Georgescu, M.-M., Kirsch, K. H., Akagi, T., Shishido, T. & Hanafusa, H. 1999. The tumor-suppressor activity of PTEN is regulated by its carboxyl-terminal region. *Proceedings of the National Academy of Sciences,* 96**,** 10182-10187.

Georgiadis, M. M., Luo, M., Gaur, R. K., Delaplane, S., Li, X. & Kelley, M. R. 2008. Evolution of the redox function in mammalian apurinic/apyrimidinic endonuclease. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis,* 643**,** 54-63.

Glaab, W. E. & Tindall, K. R. 1997. Mutation rate at the hprt locus in human cancer cell lines with specific mismatch repair-gene defects. *Carcinogenesis,* 18**,** 1-8.

Golding, S. E., Rosenberg, E., Adams, B. R., Wignarajah, S., Beckta, J. M., O'Connor, M. J. & Valerie, K. 2012. Dynamic inhibition of ATM kinase provides a strategy for glioblastoma multiforme radiosensitization and growth control. *Cell Cycle,* 11**,** 1167-73.

Golding, S. E., Rosenberg, E., Valerie, N., Hussaini, I., Frigerio, M., Cockcroft, X. F., et al. 2009. Improved ATM kinase inhibitor KU-60019 radiosensitizes glioma cells, compromises insulin, AKT and ERK prosurvival signaling, and inhibits migration and invasion. *Molecular Cancer Therapeutics,* 8**,** 2894-2902.

Gorlia, T., van den Bent, M. J., Hegi, M. E., Mirimanoff, R. O., Weller, M., Cairncross, J. G., et al. 2008. Nomograms for predicting survival of patients with newly diagnosed glioblastoma: prognostic factor analysis of EORTC and NCIC trial 26981-22981/CE.3. *Lancet Oncol,* 9**,** 29-38.

Grabsch, H., Dattani, M., Barker, L., Maughan, N., Maude, K., Hansen, O., et al. 2006. Expression of DNA double-strand break repair proteins ATM and BRCA1 predicts survival in colorectal cancer. *Clin Cancer Res,* 12**,** 1494-500.

Group, T. I. N. B. S. S. 2000. Nijmegen breakage syndrome. *Archives of Disease in Childhood,* 82**,** 400-406.

Gu, D., Wang, M., Wang, S., Zhang, Z. & Chen, J. 2011. The DNA repair gene APE1 T1349G polymorphism and risk of gastric cancer in a Chinese population. *PLoS One,* 6**,** e28971.

Guastadisegni, C., Colafranceschi, M., Ottini, L. & Dogliotti, E. 2010. Microsatellite instability as a marker of prognosis and response to therapy: a meta-analysis of colorectal cancer survival data. *European Journal of Cancer,* 46**,** 2788-98.

Guldberg, P., thor Straten, P., Birck, A., Ahrenkiel, V., Kirkin, A. F. & Zeuthen, J. 1997. Disruption of the MMAC1/PTEN Gene by Deletion or Mutation Is a Frequent Event in Malignant Melanoma. *Cancer Research,* 57**,** 3660-3663.

Guo, G. S., Zhang, F. M., Gao, R. J., Delsite, R., Feng, Z. H. & Powell, S. N. 2011. DNA repair and synthetic lethality. *Int J Oral Sci,* 3**,** 176-9.

Guo, X., Yang, C., Qian, X., Lei, T., Li, Y., Shen, H., et al. 2013. Estrogen receptor alpha regulates ATM Expression through miRNAs in breast cancer. *Clin Cancer Res,* 19**,** 4994-5002.

Gupta, A., Yang, Q., Pandita, R. K., Hunt, C. R., Xiang, T., Misri, S., et al. 2009. Cell cycle checkpoint defects contribute to genomic instability in PTEN deficient cells independent of DNA DSB repair. *Cell Cycle,* 8**,** 2198-210.

Hafsi, S., Pezzino, F. M., Candido, S., Ligresti, G., Spandidos, D. A., Soua, Z., et al. 2012. Gene alterations in the PI3K/PTEN/AKT pathway as a mechanism of drug-resistance (review). *Int J Oncol,* 40**,** 639-44.

Harima, Y., Sawada, S., Nagata, K., Sougawa, M., Ostapenko, V. & Ohnishi, T. 2001. Mutation of the PTEN gene in advanced cervical cancer correlated with tumor progression and poor outcome after radiotherapy. *Int J Oncol,* 18**,** 493-7.

Hartmann, C., Hentschel, B., Simon, M., Westphal, M., Schackert, G., Tonn, J. C., et al. 2013. Long-term survival in primary glioblastoma with versus without isocitrate dehydrogenase mutations. *Clin Cancer Res,* 19**,** 5146-57.

Hay, N. & Sonenberg, N. 2004. Upstream and downstream of mTOR. *Genes Dev,* 18**,** 1926-45.

He, Z., Wu, H., Jiao, Y. & Zheng, J. 2015. Expression and prognostic value of CD97 and its ligand CD55 in pancreatic cancer. *Oncol Lett,* 9**,** 793-797.

Hegi, M. E., Diserens, A. C., Gorlia, T., Hamou, M. F., de Tribolet, N., Weller, M., et al. 2005. MGMT gene silencing and benefit from temozolomide in glioblastoma. *N Engl J Med,* 352**,** 997-1003.

Heikkinen, K., Rapakko, K., Karppinen, S. M., Erkko, H., Nieminen, P. & Winqvist, R. 2005. Association of common ATM polymorphism with bilateral breast cancer. *International Journal of Cancer,* 116**,** 69-72.

Helleday, T., Petermann, E., Lundin, C., Hodgson, B. & Sharma, R. A. 2008. DNA repair pathways as targets for cancer therapy. *Nat Rev Cancer,* 8**,** 193-204.

Hickson, I., Zhao, Y., Richardson, C. J., Green, S. J., Martin, N. M., Orr, A. I., et al. 2004. Identification and characterization of a novel and specific inhibitor of the ataxia-telangiectasia mutated kinase ATM. *Cancer Res,* 64**,** 9152-9.

Hill, C. E., Wickliffe, J. K., Guerin, A. T., Kinslow, C. J., Wolfe, K. J., Ammenheuser, M. M. & Abdel-Rahman, S. Z. 2007. The L84F polymorphism in the O6-Methylguanine-DNA-Methyltransferase (MGMT) gene is associated with increased hypoxanthine

phosphoribosyltransferase (HPRT) mutant frequency in lymphocytes of tobacco smokers. *Pharmacogenet Genomics,* 17**,** 743-53.
Hirakawa, M., Sato, Y., Ohnuma, H., Takayama, T., Sagawa, T., Nobuoka, T., et al. 2013. A phase II study of neoadjuvant combination chemotherapy with docetaxel, cisplatin, and S-1 for locally advanced resectable gastric cancer: nucleotide excision repair (NER) as potential chemoresistance marker. *Cancer Chemother Pharmacol,* 71**,** 789-97.
Hirota, K., Murata, M., Sachi, Y., Nakamura, H., Takeuchi, J., Mori, K. & Yodoi, J. 1999. Distinct roles of thioredoxin in the cytoplasm and in the nucleus. A two-step mechanism of redox regulation of transcription factor NF-kappaB. *J Biol Chem,* 274**,** 27891-7.
Hiyama, E., Yokoyama, T., Tatsumoto, N., Hiyama, K., Imamura, Y., Murakami, Y., et al. 1995. Telomerase Activity in Gastric Cancer. *Cancer Research,* 55**,** 3258-3262.
Hoeijmakers, J. H. 2001. DNA repair mechanisms. *Maturitas,* 38**,** 17-22.
Hopfner, K. P., Karcher, A., Craig, L., Woo, T. T., Carney, J. P. & Tainer, J. A. 2001. Structural biochemistry and interaction architecture of the DNA double-strand break repair Mre11 nuclease and Rad50-ATPase. *Cell,* 105**,** 473-85.
Hossain, M., Banik, N. L. & Ray, S. K. 2012. Synergistic anti-cancer mechanisms of curcumin and paclitaxel for growth inhibition of human brain tumor stem cells and LN18 and U138MG cells. *Neurochem Int,* 61**,** 1102-13.
Houillier, C., Wang, X., Kaloshi, G., Mokhtari, K., Guillevin, R., Laffaire, J., et al. 2010. IDH1 or IDH2 mutations predict longer survival and response to temozolomide in low-grade gliomas. *Neurology,* 75**,** 1560-6.
Hsieh, P. & Yamane, K. 2008. DNA mismatch repair: molecular mechanism, cancer, and ageing. *Mechanisms of ageing and development,* 129**,** 391-407.
Huang, R. P. & Adamson, E. D. 1993. Characterization of the DNA-binding properties of the early growth response-1 (Egr-1) transcription factor: evidence for modulation by a redox mechanism. *DNA and cell biology,* 12**,** 265-73.
Ichimura, K., Pearson, D. M., Kocialkowski, S., Backlund, L. M., Chan, R., Jones, D. T. & Collins, V. P. 2009. IDH1 mutations are present in the majority of common adult gliomas but rare in primary glioblastomas. *Neuro Oncol,* 11**,** 341-7.
Inoki, K., Zhu, T. & Guan, K. L. 2003. TSC2 mediates cellular energy response to control cell growth and survival. *Cell,* 115**,** 577-90.
Irarrazabal, C. E., Burg, M. B., Ward, S. G. & Ferraris, J. D. 2006. Phosphatidylinositol 3-kinase mediates activation of ATM by high NaCl and by ionizing radiation: Role in osmoprotective transcriptional regulation. *Proc Natl Acad Sci U S A,* 103**,** 8882-7.
Ishii, N., Maier, D., Merlo, A., Tada, M., Sawamura, Y., Diserens, A. C. & Van Meir, E. G. 1999. Frequent co-alterations of TP53, p16/CDKN2A, p14ARF, PTEN tumor suppressor genes in human glioma cell lines. *Brain Pathology,* 9**,** 469-79.

Jacobs, K. M., Bhave, S. R., Ferraro, D. J., Jaboin, J. J., Hallahan, D. E. & Thotala, D. 2012. GSK-3beta: A Bifunctional Role in Cell Death Pathways. *Int J Cell Biol,* 2012**,** 930710.

Jelonek, K., Gdowicz-Klosok, A., Pietrowska, M., Borkowska, M., Korfanty, J., Rzeszowska-Wolny, J. & Widlak, P. 2010. Association between single-nucleotide polymorphisms of selected genes involved in the response to DNA damage and risk of colon, head and neck, and breast cancers in a Polish population. *J Appl Genet,* 51**,** 343-52.

Ji, Y. N., Zhan, P., Wang, J., Qiu, L. X. & Yu, L. K. 2011. APE1 Asp148Glu gene polymorphism and lung cancer risk: a meta-analysis. *Molecular Biology Reports,* 38**,** 4537-43.

Jiang, H., Reinhardt, H. C., Bartkova, J., Tommiska, J., Blomqvist, C., Nevanlinna, H., et al. 2009. The combined status of ATM and p53 link tumor development with therapeutic response. *Genes Dev,* 23**,** 1895-909.

Jones, D. T., Mulholland, S. A., Pearson, D. M., Malley, D. S., Openshaw, S. W., Lambert, S. R., et al. 2011. Adult grade II diffuse astrocytomas are genetically distinct from and more aggressive than their paediatric counterparts. *Acta Neuropathol,* 121**,** 753-61.

Kadyrov, F. A., Dzantiev, L., Constantin, N. & Modrich, P. 2006. Endonucleolytic function of MutLalpha in human mismatch repair. *Cell,* 126**,** 297-308.

Kalitsis, P., Fowler, K. J., Earle, E., Hill, J. & Choo, K. H. 1998. Targeted disruption of mouse centromere protein C gene leads to mitotic disarray and early embryo death. *Proc Natl Acad Sci U S A,* 95**,** 1136-41.

Kaloshi, G., Benouaich-Amiel, A., Diakite, F., Taillibert, S., Lejeune, J., Laigle-Donadey, F., et al. 2007. Temozolomide for low-grade gliomas: predictive impact of 1p/19q loss on response and outcome. *Neurology,* 68**,** 1831-6.

Kaloshi, G., Everhard, S., Laigle-Donadey, F., Marie, Y., Navarro, S., Mokhtari, K., et al. 2008. Genetic markers predictive of chemosensitivity and outcome in gliomatosis cerebri. *Neurology,* 70**,** 590-5.

Kamphues, C., Bova, R., Bahra, M., Klauschen, F., Muckenhuber, A., Sinn, B. V., et al. 2015. Ataxia-telangiectasia-mutated protein kinase levels stratify patients with pancreatic adenocarcinoma into prognostic subgroups with loss being a strong indicator of poor survival. *Pancreas,* 44**,** 296-301.

Kanamori, Y., Kigawa, J., Itamochi, H., Sultana, H., Suzuki, M., Ohwada, M., et al. 2002. PTEN expression is associated with prognosis for patients with advanced endometrial carcinoma undergoing postoperative chemotherapy. *International Journal of Cancer,* 100**,** 686-9.

Kang, J., D'Andrea, A. D. & Kozono, D. 2012. A DNA repair pathway-focused score for prediction of outcomes in ovarian cancer treated with platinum-based chemotherapy. *J Natl Cancer Inst,* 104**,** 670-81.

Kappelle, A. C., Postma, T. J., Taphoorn, M. J., Groeneveld, G. J., van den Bent, M. J., van Groeningen, C. J., et al. 2001. PCV chemotherapy for recurrent glioblastoma multiforme. *Neurology,* 56**,** 118-20.

Kassambara, A., Gourzones-Dmitriev, C., Sahota, S., Reme, T., Moreaux, J., Goldschmidt, H., et al. 2014. A DNA repair pathway score predicts survival in human multiple myeloma: the potential for therapeutic strategy. *Oncotarget,* 5**,** 2487-98.

Katso, R., Okkenhaug, K., Ahmadi, K., White, S., Timms, J. & Waterfield, M. D. 2001. Cellular function of phosphoinositide 3-kinases: implications for development, homeostasis, and cancer. *Annu Rev Cell Dev Biol,* 17**,** 615-75.

Kefala, M., Papageorgiou, S. G., Kontos, C. K., Economopoulou, P., Tsanas, A., Pappa, V., et al. 2013. Increased expression of phosphorylated NBS1, a key molecule of the DNA damage response machinery, is an adverse prognostic factor in patients with de novo myelodysplastic syndromes. *Leuk Res,* 37**,** 1576-82.

Kefford, R. F., Thomas, N. P., Corrie, P. G., Palmer, C., Abdi, E., Kotasek, D., et al. 2009. A phase I study of extended dosing with lomeguatrib with temozolomide in patients with advanced melanoma. *Br J Cancer,* 100**,** 1245-9.

Kelley, M. 2012. Introduction and Overview of DNA repair Targets: From Bench to Clinic. *In:* Kelley, M. (ed.) *DNA Repair in Cancer Therapy: Molecular Targets and Clinical Applications.* 1st ed. London: Academic Press.

Kelley, M. R., Georgiadis, M. M. & Fishel, M. L. 2012. APE1/Ref-1Role in Redox Signaling: Translational Applications of Targeting the Redox Function of the DNA Repair/Redox Protein APE1/Ref-1. *Curr Mol Pharmacol,* 5**,** 36-53.

Keniry, M. & Parsons, R. 2008. The role of PTEN signaling perturbations in cancer and in targeted therapy. *Oncogene,* 27**,** 5477-85.

Kennedy, R. D., Chen, C. C., Stuckert, P., Archila, E. M., De la Vega, M. A., Moreau, L. A., et al. 2007. Fanconi anemia pathway-deficient tumor cells are hypersensitive to inhibition of ataxia telangiectasia mutated. *The Journal of clinical investigation,* 117**,** 1440-9.

Kennedy, R. D., Quinn, J. E., Mullan, P. B., Johnston, P. G. & Harkin, D. P. 2004. The Role of BRCA1 in the Cellular Response to Chemotherapy. *Journal of the National Cancer Institute,* 96**,** 1659-1668.

Khalil, H. S., Tummala H, Chakarov S, Zhelev N, Lane D. P. 2012. Targeting ATM pathway pathway for therapeutic intervention in cancer. *Biodiscovery* [Online], 3.

Khan, O. A., Ranson, M., Michael, M., Olver, I., Levitt, N. C., Mortimer, P., et al. 2008. A phase II trial of lomeguatrib and temozolomide in metastatic colorectal cancer. *Br J Cancer,* 98**,** 1614-8.

Khoronenkova, S. V. & Dianov, G. L. 2015. ATM prevents DSB formation by coordinating SSB repair and cell cycle progression. *Proc Natl Acad Sci U S A,* 112**,** 3997-4002.

Kim, H., Saka, B., Knight, S., Borges, M., Childs, E., Klein, A., et al. 2014a. Having pancreatic cancer with tumoral loss of ATM and normal TP53 protein expression is associated with a poorer prognosis. *Clin Cancer Res,* 20**,** 1865-72.

Kim, J. W., Im, S. A., Kim, M. A., Cho, H. J., Lee, D. W., Lee, K. H., et al. 2014b. Ataxia-telangiectasia-mutated protein expression with

microsatellite instability in gastric cancer as prognostic marker. *International Journal of Cancer,* 134**,** 72-80.

Kinner, A., Wu, W., Staudt, C. & Iliakis, G. 2008. γ-H2AX in recognition and signaling of DNA double-strand breaks in the context of chromatin. *Nucleic Acids Research,* 36**,** 5678-5694.

Kolesnikov, N., Hastings, E., Keays, M., Melnichuk, O., Tang, Y. A., Williams, E., et al. 2015. ArrayExpress update--simplifying data submissions. *Nucleic Acids Res,* 43**,** D1113-6.

Kortum, F., Das, S., Flindt, M., Morris-Rosendahl, D. J., Stefanova, I., Goldstein, A., et al. 2011. The core FOXG1 syndrome phenotype consists of postnatal microcephaly, severe mental retardation, absent language, dyskinesia, and corpus callosum hypogenesis. *Journal of Medical Genetics,* 48**,** 396-406.

Koschny, R., Koschny, T., Froster, U. G., Krupp, W. & Zuber, M. A. 2002. Comparative genomic hybridization in glioma: a meta-analysis of 509 cases. *Cancer Genet Cytogenet,* 135**,** 147-59.

Labianca, R., Nordlinger, B., Beretta, G. D., Brouquet, A., Cervantes, A. & Group, O. b. o. t. E. G. W. 2010. Primary colon cancer: ESMO Clinical Practice Guidelines for diagnosis, adjuvant treatment and follow-up. *Annals of Oncology,* 21**,** v70-v77.

Lamarche, B. J., Orazio, N. I. & Weitzman, M. D. 2010. The MRN complex in Double-Strand Break Repair and Telomere Maintenance. *Febs Letters,* 584**,** 3682-3695.

Lambeth, L. S. & Smith, C. A. 2013. Short hairpin RNA-mediated gene silencing. *Methods Mol Biol,* 942**,** 205-32.

Lancashire, L. J., Lemetre, C. & Ball, G. R. 2009. An introduction to artificial neural networks in bioinformatics--application to complex microarray and mass spectrometry datasets in cancer studies. *Brief Bioinform,* 10**,** 315-29.

Lancashire, L. J., Powe, D. G., Reis-Filho, J. S., Rakha, E., Lemetre, C., Weigelt, B., et al. 2010. A validated gene expression profile for detecting clinical outcome in breast cancer using artificial neural networks. *Breast Cancer Res Treat,* 120**,** 83-93.

Lando, D., Pongratz, I., Poellinger, L. & Whitelaw, M. L. 2000. A redox mechanism controls differential DNA binding activities of hypoxia-inducible factor (HIF) 1alpha and the HIF-like factor. *J Biol Chem,* 275**,** 4618-27.

Latif, A. Z., Signorini, D., Gregor, A., Grant, R., Ironside, J. W. & Whittle, I. R. 1998. Application of the MRC brain tumour prognostic index to patients with malignant glioma not managed in randomised control trial. *J Neurol Neurosurg Psychiatry,* 64**,** 747-50.

Lavin, M. F. 2008. Ataxia-telangiectasia: from a rare disorder to a paradigm for cell signalling and cancer. *Nat Rev Mol Cell Biol,* 9**,** 759-69.

Le, S., Zhu, J. J., Anthony, D. C., Greider, C. W. & Black, P. M. 1998. Telomerase Activity in Human Gliomas. *Neurosurgery,* 42**,** 1120-1124.

Lee, J. O., Yang, H., Georgescu, M. M., Di Cristofano, A., Maehama, T., Shi, Y., et al. 1999. Crystal structure of the PTEN tumor suppressor: implications for its phosphoinositide phosphatase activity and membrane association. *Cell,* 99**,** 323-34.

LeicaBiosystems. 2016. *Novolink™ Polymer Detection Systems* [Online]. Available at: http://www.leicabiosystems.com/fileadmin/img_uploads/novocastra_reagents/Novocastra_datasheets/re7280-ce.pdf [Accessed 24th January 2016].

Li, D., Frazier, M., Evans, D. B., Hess, K. R., Crane, C. H., Jiao, L. & Abbruzzese, J. L. 2006. Single Nucleotide Polymorphisms of RecQ1, RAD54L, and ATM Genes Are Associated With Reduced Survival of Pancreatic Cancer. *Journal of Clinical Oncology,* 24**,** 1720-1728.

Li, D. M. & Sun, H. 1998. PTEN/MMAC1/TEP1 suppresses the tumorigenicity and induces G1 cell cycle arrest in human glioblastoma cells. *Proc Natl Acad Sci U S A,* 95**,** 15406-11.

Li, J., Yen, C., Liaw, D., Podsypanina, K., Bose, S., Wang, S. I., et al. 1997. PTEN, a putative protein tyrosine phosphatase gene mutated in human brain, breast, and prostate cancer. *Science,* 275**,** 1943-1947.

Li, L., Spendlove, I., Morgan, J. & Durrant, L. G. 2001. CD55 is over-expressed in the tumour environment. *Br J Cancer,* 84**,** 80-6.

Li, Y., Corradetti, M. N., Inoki, K. & Guan, K. L. 2004. TSC2: filling the GAP in the mTOR signaling pathway. *Trends in biochemical sciences,* 29**,** 32-8.

Li, Y., Jin, G., Wang, H., Liu, H., Qian, J., Gu, S., et al. 2007. Polymorphisms of CAK genes and risk for lung cancer: a case-control study in Chinese population. *Lung Cancer,* 58**,** 171-83.

Li, Y. & Yang, D. Q. 2010. The ATM inhibitor KU-55933 suppresses cell proliferation and induces apoptosis by blocking Akt in cancer cells with overactivated Akt. *Mol Cancer Ther,* 9**,** 113-25.

Li, Z., Guan, W., Li, M. X., Zhong, Z. Y., Qian, C. Y., Yang, X. Q., et al. 2011. Genetic polymorphism of DNA base-excision repair genes (APE1, OGG1 and XRCC1) and their correlation with risk of lung cancer in a Chinese population. *Archives of Medical Research,* 42**,** 226-34.

Lian, Z. & Di Cristofano, A. 2005. Class reunion: PTEN joins the nuclear crew. *Oncogene,* 24**,** 7394-400.

Liang, J. & Slingerland, J. M. 2003. Multiple Roles of the PI3K/PKB (Akt) Pathway in Cell Cycle Progression. *Cell Cycle,* 2**,** 336-342.

Liang, J., Zubovitz, J., Petrocelli, T., Kotchetkov, R., Connor, M. K., Han, K., et al. 2002. PKB/Akt phosphorylates p27, impairs nuclear import of p27 and opposes p27-mediated G1 arrest. *Nat Med,* 8**,** 1153-60.

Liaw, D., Marsh, D. J., Li, J., Dahia, P. L., Wang, S. I., Zheng, Z., et al. 1997. Germline mutations of the PTEN gene in Cowden disease, an inherited breast and thyroid cancer syndrome. *Nat Genet,* 16**,** 64-7.

Life Technologies. 2012. *Real-time PCR handbook* [Online]. Available at: http://www.gene-quantification.com/real-time-pcr-handbook-life-technologies-update-flr.pdf [Accessed 30th August 2015].

Lindahl, T. 1993. Instability and decay of the primary structure of DNA. *Nature,* 362**,** 709-15.

Linge, A., Maurya, P., Friedrich, K., Baretton, G. B., Kelly, S., Henry, M., et al. 2014. Identification and Functional Validation of RAD23B as a

Potential Protein in Human Breast Cancer Progression. *Journal of Proteome Research,* 13**,** 3212-3222.
Lippert, C., Ghahramani, Z. & Borgwardt, K. M. 2010. Gene function prediction from synthetic lethality networks via ranking on demand. *Bioinformatics,* 26**,** 912-8.
Liu, G., Loraine, A. E., Shigeta, R., Cline, M., Cheng, J., Valmeekam, V., et al. 2003. NetAffx: Affymetrix probesets and annotations. *Nucleic Acids Res,* 31**,** 82-6.
Liu, J. L., Mao, Z., LaFortune, T. A., Alonso, M. M., Gallick, G. E., Fueyo, J. & Yung, W. K. 2007. Cell cycle-dependent nuclear export of phosphatase and tensin homologue tumor suppressor is regulated by the phosphoinositide-3-kinase signaling cascade. *Cancer Res,* 67**,** 11054-63.
Liu, P., Cheng, H., Roberts, T. M. & Zhao, J. J. 2009. Targeting the phosphoinositide 3-kinase pathway in cancer. *Nat Rev Drug Discov,* 8**,** 627-44.
Lo, Y. L., Hsiao, C. F., Jou, Y. S., Chang, G. C., Tsai, Y. H., Su, W. C., et al. 2010. ATM polymorphisms and risk of lung cancer among never smokers. *Lung Cancer,* 69**,** 148-54.
Lo, Y. L., Jou, Y. S., Hsiao, C. F., Chang, G. C., Tsai, Y. H., Su, W. C., et al. 2009. A polymorphism in the APE1 gene promoter is associated with lung cancer risk. *Cancer Epidemiol Biomarkers Prev,* 18**,** 223-9.
Loeb, L. A. 2001. A Mutator Phenotype in Cancer. *Cancer Research,* 61**,** 3230-3239.
Louis, D., Ohgaki, H., Wiestler, O., Cavenee, W., Burger, P., Jouvet, A., et al. 2007. The 2007 WHO Classification of Tumours of the Central Nervous System. *Acta Neuropathol,* 114**,** 97-109.
Lu, J., Zhang, S., Chen, D., Wang, H., Wu, W., Wang, X., et al. 2009. Functional characterization of a promoter polymorphism in APE1/Ref-1 that contributes to reduced lung cancer susceptibility. *FASEB J.*
Lukas, C., Falck, J., Bartkova, J., Bartek, J. & Lukas, J. 2003. Distinct spatiotemporal dynamics of mammalian checkpoint regulators induced by DNA damage. *Nat Cell Biol,* 5**,** 255-60.
Lynch, H. T. & Smyrk, T. 1996. Hereditary nonpolyposis colorectal cancer (Lynch syndrome). An updated review. *Cancer,* 78**,** 1149-67.
Ma, C. X. & Ellis, M. J. 2013. The Cancer Genome Atlas: clinical applications for breast cancer. *Oncology (Williston Park),* 27**,** 1263-9, 1274-9.
Ma, Z., Lin, M., Li, K., Fu, Y., Liu, X., Yang, D., et al. 2013. Knocking down SMC1A inhibits growth and leads to G2/M arrest in human glioma cells. *Int J Clin Exp Pathol,* 6**,** 862-9.
Madhusudan, S. & Hickson, I. D. 2005. DNA Repair Inhibition: a selective tumour targeting strategy. *Trends Mol Med,* 11**,** 503-511.
Madhusudan, S. & Middleton, M. R. 2005. The emerging role of DNA repair proteins as predictive, prognostic and therapeutic targets in cancer. *Cancer Treat Rev,* 31**,** 603-17.
Madjd, Z., Durrant, L. G., Bradley, R., Spendlove, I., Ellis, I. O. & Pinder, S. E. 2004. Loss of CD55 is associated with aggressive breast tumors. *Clin Cancer Res,* 10**,** 2797-803.

Maemondo, M., Inoue, A., Kobayashi, K., Sugawara, S., Oizumi, S., Isobe, H., et al. 2010. Gefitinib or Chemotherapy for Non–Small-Cell Lung Cancer with Mutated EGFR. *New England Journal of Medicine,* 362**,** 2380-2388.

Maher, E. A., Brennan, C., Wen, P. Y., Durso, L., Ligon, K. L., Richardson, A., et al. 2006. Marked genomic differences characterize primary and secondary glioblastoma subtypes and identify two distinct molecular and clinical secondary glioblastoma entities. *Cancer Res,* 66**,** 11502-13.

Mahmood, S. F., Gruel, N., Chapeaublanc, E., Lescure, A., Jones, T., Reyal, F., et al. 2014. A siRNA screen identifies RAD21, EIF3H, CHRAC1 and TANC2 as driver genes within the 8q23, 8q24.3 and 17q23 amplicons in breast cancer with effects on cell growth, survival and transformation. *Carcinogenesis,* 35**,** 670-82.

Malumbres, M. & Barbacid, M. 2005. Mammalian cyclin-dependent kinases. *Trends in biochemical sciences,* 30**,** 630-41.

Marsh, D. J., Dahia, P. L., Zheng, Z., Liaw, D., Parsons, R., Gorlin, R. J. & Eng, C. 1997. Germline mutations in PTEN are present in Bannayan-Zonana syndrome. *Nature Genetics,* 16**,** 333-4.

Marteijn, J. A., Lans, H., Vermeulen, W. & Hoeijmakers, J. H. J. 2014. Understanding nucleotide excision repair and its roles in cancer and ageing. *Nat Rev Mol Cell Biol,* 15**,** 465-481.

Martin, A. & Scharff, M. D. 2002. AID and mismatch repair in antibody diversification. *Nat Rev Immunol,* 2**,** 605-614.

Martin, S. A. & Ouchi, T. 2008. Cellular commitment to reentry into the cell cycle after stalled DNA is determined by site-specific phosphorylation of Chk1 and PTEN. *Mol Cancer Ther,* 7**,** 2509-16.

Matsumura, Y. & Ananthaswamy, H. N. 2002. Short-term and long-term cellular and molecular events following UV irradiation of skin: implications for molecular medicine. *Expert Reviews in Molecular Medicine,* 4**,** 1-22.

Matsushita, H., Uenaka, A., Ono, T., Hasegawa, K., Sato, S., Koizumi, F., et al. 2005. Identification of glioma-specific RFX4-E and -F isoforms and humoral immune response in patients. *Cancer science,* 96**,** 801-9.

McCabe, N., Hanna, C., Walker, S. M., Gonda, D., Li, J., Wikstrom, K., et al. 2015. Mechanistic Rationale to Target PTEN-Deficient Tumor Cells with Inhibitors of the DNA Damage Response Kinase ATM. *Cancer Res,* 75**,** 2159-65.

McCabe, N., Turner, N. C., Lord, C. J., Kluzek, K., Bialkowska, A., Swift, S., et al. 2006. Deficiency in the repair of DNA damage by homologous recombination and sensitivity to poly(ADP-ribose) polymerase inhibition. *Cancer Research,* 66**,** 8109-8115.

McEllin, B., Camacho, C. V., Mukherjee, B., Hahm, B., Tomimatsu, N., Bachoo, R. M. & Burma, S. 2010. PTEN loss compromises homologous recombination repair in astrocytes: implications for glioblastoma therapy with temozolomide or poly(ADP-ribose) polymerase inhibitors. *Cancer Res,* 70**,** 5457-64.

McManus, M. T. & Sharp, P. A. 2002. Gene silencing in mammals by small interfering RNAs. *Nat Rev Genet,* 3**,** 737-47.

Mehta, S., Shelling, A., Muthukaruppan, A., Lasham, A., Blenkiron, C., Laking, G. & Print, C. 2010. Predictive and prognostic molecular markers for cancer medicine. *Ther Adv Med Oncol,* 2**,** 125-48.

Mendes-Pereira, A. M., Martin, S. A., Brough, R., McCarthy, A., Taylor, J. R., Kim, J. S., et al. 2009. Synthetic lethal targeting of PTEN mutant cells with PARP inhibitors. *Embo Molecular Medicine,* 1**,** 315-322.

Mereniuk TR, Foley E, Hall DG, Lai R & Weinfield M. Synthetic lethality through targeted disruption of polynucleotide kinase/phosphatase. Genomic Instability and DNA Repair Conference, 2013 Banff, Canada.

Mi, H., Muruganujan, A. & Thomas, P. D. 2013. PANTHER in 2013: modeling the evolution of gene function, and other gene attributes, in the context of phylogenetic trees. *Nucleic Acids Res,* 41**,** D377-86.

Miled, N., Yan, Y., Hon, W. C., Perisic, O., Zvelebil, M., Inbar, Y., et al. 2007. Mechanism of two classes of cancer mutations in the phosphoinositide 3-kinase catalytic subunit. *Science,* 317**,** 239-42.

Minami, D., Takigawa, N., Takeda, H., Takata, M., Ochi, N., Ichihara, E., et al. 2012. Synergistic effect of olaparib with combination of cisplatin on PTEN deficient lung cancer cells. *Molecular Cancer Research*.

Ming, M., Feng, L., Shea, C. R., Soltani, K., Zhao, B., Han, W., et al. 2011. PTEN positively regulates UVB-induced DNA damage repair. *Cancer Res,* 71**,** 5287-95.

Ming, M. & He, Y. Y. 2012. PTEN in DNA damage repair. *Cancer Lett,* 319**,** 125-9.

Ming, M., Shea, C. R., Guo, X., Li, X., Soltani, K., Han, W. & He, Y. Y. 2010. Regulation of global genome nucleotide excision repair by SIRT1 through xeroderma pigmentosum C. *Proc Natl Acad Sci U S A,* 107**,** 22623-8.

Miron, M., Lasko, P. & Sonenberg, N. 2003. Signaling from Akt to FRAP/TOR targets both 4E-BP and S6K in Drosophila melanogaster. *Mol Cell Biol,* 23**,** 9117-26.

Mittal, R. D., Mandal, R. K. & Gangwar, R. 2012. Base excision repair pathway genes polymorphism in prostate and bladder cancer risk in North Indian population. *Mechanisms of ageing and development,* 133**,** 127-32.

Mladenov, E. & Iliakis, G. 2011. Induction and repair of DNA double strand breaks: the increasing spectrum of non-homologous end joining pathways. *Mutat Res,* 711**,** 61-72.

Modrich, P. 2006. Mechanisms in eukaryotic mismatch repair. *J Biol Chem,* 281**,** 30305-9.

Moore, C. B., Guthrie, E. H., Huang, M. T.-H. & Taxman, D. J. 2010. Short Hairpin RNA (shRNA): Design, Delivery, and Assessment of Gene Knockdown. *Methods in molecular biology (Clifton, N.J.),* 629**,** 141-158.

Moreno-Herrero, F., de Jager, M., Dekker, N. H., Kanaar, R., Wyman, C. & Dekker, C. 2005. Mesoscale conformational changes in the DNA-repair complex Rad50/Mre11/Nbs1 upon binding DNA. *Nature,* 437**,** 440-3.

Morgan, S. E., Lovly, C., Pandita, T. K., Shiloh, Y. & Kastan, M. B. 1997. Fragments of ATM which have dominant-negative or complementing activity. *Mol Cell Biol,* 17**,** 2020-9.

MRC 1990. Prognostic factors for high-grade malignant glioma: development of a prognostic index. A Report of the Medical Research Council Brain Tumour Working Party. *J Neurooncol,* 9**,** 47-55.
Munshi, A., Hobbs, M. & Meyn, R. E. 2005. Clonogenic cell survival assay. *Methods Mol Med,* 110**,** 21-8.
Murao, T., Shiotani, A., Fujita, Y., Yamanaka, Y., Kamada, T., Manabe, N., et al. 2015. Overexpression of CD55 from Barrett's esophagus is associated with esophageal adenocarcinoma risk. *J Gastroenterol Hepatol*.
Nadkarni, A., Shrivastav, M., Mladek, A. C., Schwingler, P. M., Grogan, P. T., Chen, J. & Sarkaria, J. N. 2012. ATM inhibitor KU-55933 increases the TMZ responsiveness of only inherently TMZ sensitive GBM cells. *J Neurooncol,* 110**,** 349-57.
Naidu, M. D., Mason, J. M., Pica, R. V., Fung, H. & Peña, L. A. 2010. Radiation Resistance in Glioma Cells Determined by DNA Damage Repair Activity of Ape1/Ref-1. *Journal of Radiation Research,* 51**,** 393-404.
Narter, K. F., Ergen, A., Agachan, B., Gormus, U., Timirci, O. & Isbir, T. 2009. Bladder cancer and polymorphisms of DNA repair genes (XRCC1, XRCC3, XPD, XPG, APE1, hOGG1). *Anticancer Res,* 29**,** 1389-93.
NCBI. 2015a. *DCLK2 doublecortin-like kinase 2 [Homo sapiens (human)]* [Online]. Available at: http://www.ncbi.nlm.nih.gov/gene/166614 [Accessed 31st August 2015].
NCBI. 2015b. *HPRT1 hypoxanthine phosphoribosyltransferase 1 [Homo sapiens (human)]* [Online]. Available at: http://www.ncbi.nlm.nih.gov/gene/3251 [Accessed 31st August 2015].
Nicholson, K. M. & Anderson, N. G. 2002. The protein kinase B/Akt signalling pathway in human malignancy. *Cellular Signalling,* 14**,** 381-95.
National Institutes for Health. 2015. *ClinicalTrials.gov* [Online]. Available at: https://clinicaltrials.gov/ [Accessed July 29th].
Nijman, S. M. 2011. Synthetic lethality: general principles, utility and detection using genetic screens in human cells. *Febs Letters,* 585**,** 1-6.
Nilsson, S., Carlsson, J., Larsson, B. & Ponten, J. 1980. Survival of irradiated glia and glioma cells studied with a new cloning technique. *Int J Radiat Biol Relat Stud Phys Chem Med,* 37**,** 267-79.
O'Toole, S. A., Sheppard, B. L., McGuinness, E. P., Gleeson, N. C., Yoneda, M. & Bonnar, J. 2003. The MTS assay as an indicator of chemosensitivity/resistance in malignant gynaecological tumours. *Cancer Detection and Prevention,* 27**,** 47-54.
Okazaki, T., Chung, U., Nishishita, T., Ebisu, S., Usuda, S., Mishiro, S., et al. 1994. A redox factor protein, ref1, is involved in negative gene regulation by extracellular calcium. *J Biol Chem,* 269**,** 27855-62.
Olive, P. L. & Banath, J. P. 2006. The comet assay: a method to measure DNA damage in individual cells. *Nat Protoc,* 1**,** 23-9.
Omberg, L., Ellrott, K., Yuan, Y., Kandoth, C., Wong, C., Kellen, M. R., et al. 2013. Enabling transparent and collaborative computational analysis of

12 tumor types within The Cancer Genome Atlas. *Nature Genetics,* 45**,** 1121-6.
Opitz, I., Soltermann, A., Abaecherli, M., Hinterberger, M., Probst-Hensch, N., Stahel, R., et al. 2008. PTEN expression is a strong predictor of survival in mesothelioma patients. *Eur J Cardiothorac Surg,* 33**,** 502-6.
Orloff, M. S. & Eng, C. 2008. Genetic and phenotypic heterogeneity in the PTEN hamartoma tumour syndrome. *Oncogene,* 27**,** 5387-5397.
Pardini, B., Naccarati, A., Novotny, J., Smerhovsky, Z., Vodickova, L., Polakova, V., et al. 2008. DNA repair genetic polymorphisms and risk of colorectal cancer in the Czech Republic. *Mutat Res,* 638**,** 146-53.
Parsons, D. W., Jones, S., Zhang, X., Lin, J. C., Leary, R. J., Angenendt, P., et al. 2008. An integrated genomic analysis of human glioblastoma multiforme. *Science,* 321**,** 1807-12.
Pastink, A., Eeken, J. C. & Lohman, P. H. 2001. Genomic integrity and the repair of double-strand DNA breaks. *Mutat Res,* 480-481**,** 37-50.
Pegg, A. E. 2000. Repair of O(6)-alkylguanine by alkyltransferases. *Mutat Res,* 462**,** 83-100.
Pegg, A. E. & Byers, T. L. 1992. Repair of DNA containing O6-alkylguanine. *FASEB J,* 6**,** 2302-10.
Pegg, A. E., Dolan, M. E. & Moschel, R. C. 1995. Structure, function, and inhibition of O6-alkylguanine-DNA alkyltransferase. *Prog Nucleic Acid Res Mol Biol,* 51**,** 167-223.
Peltomäki, P. 2001. Deficient DNA mismatch repair: a common etiologic factor for colon cancer. *Human Molecular Genetics,* 10**,** 735-740.
Peltomaki, P., Aaltonen, L. A., Sistonen, P., Pylkkanen, L., Mecklin, J. P., Jarvinen, H., et al. 1993. Genetic mapping of a locus predisposing to human colorectal cancer. *Science,* 260**,** 810-2.
Phillips, H. S., Kharbanda, S., Chen, R., Forrest, W. F., Soriano, R. H., Wu, T. D., et al. 2006. Molecular subclasses of high-grade glioma predict prognosis, delineate a pattern of disease progression, and resemble stages in neurogenesis. *Cancer Cell,* 9**,** 157-73.
Phuah, S. Y., Looi, L. M., Hassan, N., Rhodes, A., Dean, S., Taib, N. A., et al. 2012. Triple-negative breast cancer and PTEN (phosphatase and tensin homologue)loss are predictors of BRCA1 germline mutations in women with early-onset and familial breast cancer, but not in women with isolated late-onset breast cancer. *Breast Cancer Research,* 14**,** R142.
Pletsa, V., Valavanis, C., van Delft, J. H., Steenwinkel, M. J. & Kyrtopoulos, S. A. 1997. DNA damage and mutagenesis induced by procarbazine in lambda lacZ transgenic mice: evidence that bone marrow mutations do not arise primarily through miscoding by O6-methylguanine. *Carcinogenesis,* 18**,** 2191-6.
Plumb, J. A., Milroy, R. & Kaye, S. B. 1989. Effects of the pH dependence of 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyl-tetrazolium bromide-formazan absorption on chemosensitivity determined by a novel tetrazolium-based assay. *Cancer Res,* 49**,** 4435-40.
Pohl, U., Wagenknecht, B., Naumann, U. & Weller, M. 1999. p53 enhances BAK and CD95 expression in human malignant glioma cells but does not enhance CD95L-induced apoptosis. *Cell Physiol Biochem,* 9**,** 29-37.

Porcedda, P., Turinetto, V., Brusco, A., Cavalieri, S., Lantelme, E., Orlando, L., et al. 2008. A rapid flow cytometry test based on histone H2AX phosphorylation for the sensitive and specific diagnosis of ataxia telangiectasia. *Cytometry A,* 73**,** 508-16.

Pouratian, N. & Schiff, D. 2010. Management of low-grade glioma. *Curr Neurol Neurosci Rep,* 10**,** 224-31.

Promega. 2015a. *Apoptosis* [Online]. Available at: https://www.bdbiosciences.com/in/research/apoptosis/analysis/cell_death.jsp [Accessed 19th October 2015].

Promega. 2015b. *CellTiter 96® AQueous One Solution Cell Proliferation Assay (MTS)* [Online]. Available at: https://www.promega.co.uk/products/cell-health-and-metabolism/cell-viability-assays/celltiter-96-aqueous-one-solution-cell-proliferation-assay-_mts_/ [Accessed 10th October 2015].

Puc, J., Keniry, M., Li, H. S., Pandita, T. K., Choudhury, A. D., Memeo, L., et al. 2005. Lack of PTEN sequesters CHK1 and initiates genetic instability. *Cancer Cell,* 7**,** 193-204.

Puc, J. & Parsons, R. 2005. PTEN loss inhibits CHK1 to cause double stranded-DNA breaks in cells. *Cell Cycle,* 4**,** 927-9.

Qiu, Z. K., Shen, D., Chen, Y. S., Yang, Q. Y., Guo, C. C., Feng, B. H. & Chen, Z. P. 2014. Enhanced MGMT expression contributes to temozolomide resistance in glioma stem-like cells. *Chin J Cancer,* 33**,** 115-22.

Quiros, S., Roos, W. P. & Kaina, B. 2011. Rad51 and BRCA2--New molecular targets for sensitizing glioma cells to alkylating anticancer drugs. *PLoS One,* 6**,** e27183.

Radu, A., Neubauer, V., Akagi, T., Hanafusa, H. & Georgescu, M. M. 2003. PTEN induces cell cycle arrest by decreasing the level and nuclear localization of cyclin D1. *Mol Cell Biol,* 23**,** 6139-49.

Rainey, M. D., Charlton, M. E., Stanton, R. V. & Kastan, M. B. 2008. Transient Inhibition of ATM Kinase Is Sufficient to Enhance Cellular Sensitivity to Ionizing Radiation. *Cancer Research,* 68**,** 7466-7474.

Ranson, M., Hersey, P., Thompson, D., Beith, J., McArthur, G. A., Haydon, A., et al. 2007. Randomized trial of the combination of lomeguatrib and temozolomide compared with temozolomide alone in chemotherapy naive patients with metastatic cutaneous melanoma. *J Clin Oncol,* 25**,** 2540-5.

Ranson, M., Middleton, M. R., Bridgewater, J., Lee, S. M., Dawson, M., Jowle, D., et al. 2006. Lomeguatrib, a potent inhibitor of O6-alkylguanine-DNA-alkyltransferase: phase I safety, pharmacodynamic, and pharmacokinetic trial and evaluation in combination with temozolomide in patients with advanced solid tumors. *Clin Cancer Res,* 12**,** 1577-84.

Rasheed, B. K., Stenzel, T. T., McLendon, R. E., Parsons, R., Friedman, A. H., Friedman, H. S., et al. 1997. PTEN gene mutations are seen in high-grade but not in low-grade gliomas. *Cancer Res,* 57**,** 4187-90.

Regnier, V., Vagnarelli, P., Fukagawa, T., Zerjal, T., Burns, E., Trouche, D., et al. 2005. CENP-A is required for accurate chromosome segregation and

sustained kinetochore association of BubR1. *Mol Cell Biol,* 25**,** 3967-81.

Reifenberger, J., Reifenberger, G., Liu, L., James, C. D., Wechsler, W. & Collins, V. P. 1994. Molecular genetic analysis of oligodendroglial tumors shows preferential allelic deletions on 19q and 1p. *American Journal of Pathology,* 145**,** 1175-90.

Renwick, A., Thompson, D., Seal, S., Kelly, P., Chagtai, T., Ahmed, M., et al. 2006. ATM mutations that cause ataxia-telangiectasia are breast cancer susceptibility alleles. *Nature Genetics,* 38**,** 873-5.

Rivera, A. L., Pelloski, C. E., Gilbert, M. R., Colman, H., De La Cruz, C., Sulman, E. P., et al. 2010. MGMT promoter methylation is predictive of response to radiotherapy and prognostic in the absence of adjuvant alkylating chemotherapy for glioblastoma. *Neuro Oncol,* 12**,** 116-121.

Robertson, E., Perry, C., Doherty, R. & Madhusudan, S. 2015. Transcriptomic profiling of Forkhead box transcription factors in adult glioblastoma multiforme. *Cancer Genomics Proteomics,* 12**,** 103-12.

Rodriguez-Caballero, A., Torres-Lagares, D., Rodriguez-Perez, A., Serrera-Figallo, M. A., Hernandez-Guisado, J. M. & Machuca-Portillo, G. 2010. Cri du chat syndrome: a critical review. *Med Oral Patol Oral Cir Bucal,* 15**,** e473-8.

Rottenberg, S., Jaspers, J. E., Kersbergen, A., van der Burg, E., Nygren, A. O., Zander, S. A., et al. 2008. High sensitivity of BRCA1-deficient mammary tumors to the PARP inhibitor AZD2281 alone and in combination with platinum drugs. *Proc Natl Acad Sci U S A,* 105**,** 17079-84.

Ruano, Y., Mollejo, M., Ribalta, T., Fiano, C., Camacho, F. I., Gomez, E., et al. 2006. Identification of novel candidate target genes in amplicons of Glioblastoma multiforme tumors detected by expression and CGH microarray profiling. *Molecular Cancer,* 5**,** 39.

Russo, A. L., Kwon, H. C., Burgan, W. E., Carter, D., Beam, K., Weizheng, X., et al. 2009. In vitro and in vivo radiosensitization of glioblastoma cells by the poly (ADP-ribose) polymerase inhibitor E7016. *Clin Cancer Res,* 15**,** 607-12.

Samuels, Y., Wang, Z., Bardelli, A., Silliman, N., Ptak, J., Szabo, S., et al. 2004. High Frequency of Mutations of the PIK3CA Gene in Human Cancers. *Science,* 304**,** 554-554.

Sanal, O., Wei, S., Foroud, T., Malhotra, U., Concannon, P., Charmley, P., et al. 1990. Further mapping of an ataxia-telangiectasia locus to the chromosome 11q23 region. *American Journal of Human Genetics,* 47**,** 860-6.

Sano, T., Lin, H., Chen, X., Langford, L. A., Koul, D., Bondy, M. L., et al. 1999. Differential expression of MMAC/PTEN in glioblastoma multiforme: relationship to localization and prognosis. *Cancer Res,* 59**,** 1820-4.

Sarbia, M., Ott, N., Puhringer-Oppermann, F. & Brucher, B. L. D. M. 2007. The predictive value of molecular markers (p53, EGFR, ATM, CHK2) in multimodally treated squamous cell carcinoma of the oesophagus. *Br J Cancer,* 97**,** 1404-1408.

Sartori, A. A., Lukas, C., Coates, J., Mistrik, M., Fu, S., Bartek, J., et al. 2007. Human CtIP promotes DNA end resection. *Nature,* 450**,** 509-14.
Sassi, A., Popielarski, M., Synowiec, E., Morawiec, Z. & Wozniak, K. 2013. BLM and RAD51 genes polymorphism and susceptibility to breast cancer. *Pathol Oncol Res,* 19**,** 451-9.
Sawai, H., Yasuda, A., Ochi, N., Ma, J., Matsuo, Y., Wakasugi, T., et al. 2008. Loss of PTEN expression is associated with colorectal cancer liver metastasis and poor patient survival. *BMC Gastroenterol,* 8**,** 56.
Schafer, K. A. 1998. The Cell Cycle: A Review. *Veterinary Pathology Online,* 35**,** 461-478.
Schwartzbaum, J. A., Fisher, J. L., Aldape, K. D. & Wrensch, M. 2006a. Epidemiology and molecular pathology of glioma. *Nat Clin Pract Neuro,* 2**,** 494-503.
Schwartzbaum, J. A., Fisher, J. L., Aldape, K. D. & Wrensch, M. 2006b. Epidemiology and molecular pathology of glioma. *Nat Clin Pract Neurol,* 2**,** 494-503; quiz 1 p following 516.
Semenza, G. L. 2003. Targeting HIF-1 for cancer therapy. *Nat Rev Cancer,* 3**,** 721-32.
Seol, H. J., Yoo, H. Y., Jin, J., Joo, K. M., Kong, D. S., Yoon, S. J., et al. 2011. Prognostic implications of the DNA damage response pathway in glioblastoma. *Oncol Rep,* 26**,** 423-30.
Shaheen, F. S., Znojek, P., Fisher, A., Webster, M., Plummer, R., Gaughan, L., et al. 2011. Targeting the DNA double strand break repair machinery in prostate cancer. *PLoS One,* 6**,** e20311.
Shaw, A. T., Kim, D.-W., Nakagawa, K., Seto, T., Crinó, L., Ahn, M.-J., et al. 2013. Crizotinib versus Chemotherapy in Advanced ALK-Positive Lung Cancer. *New England Journal of Medicine,* 368**,** 2385-2394.
Shen, L., Yin, Z.-H., Wan, Y., Zhang, Y., Li, K. & Zhou, B.-S. 2012. Association between ATM polymorphisms and cancer risk: a meta-analysis. *Molecular Biology Reports,* 39**,** 5719-5725.
Shen, W. H., Balajee, A. S., Wang, J., Wu, H., Eng, C., Pandolfi, P. P. & Yin, Y. 2007. Essential Role for Nuclear PTEN in Maintaining Chromosomal Integrity. *Cell,* 128**,** 157-170.
Shin, I., Yakes, F. M., Rojo, F., Shin, N. Y., Bakin, A. V., Baselga, J. & Arteaga, C. L. 2002. PKB/Akt mediates cell-cycle progression by phosphorylation of p27(Kip1) at threonine 157 and modulation of its cellular localization. *Nat Med,* 8**,** 1145-52.
Singh, B., Reddy, P. G., Goberdhan, A., Walsh, C., Dao, S., Ngai, I., et al. 2002. p53 regulates cell survival by inhibiting PIK3CA in squamous cell carcinomas. *Genes Dev,* 16**,** 984-93.
Smiraldo, P. G., Gruver, A. M., Osborn, J. C. & Pittman, D. L. 2005. Extensive chromosomal instability in Rad51d-deficient mouse cells. *Cancer Res,* 65**,** 2089-96.
Smith, J. M., Kirk, E. P. E., Theodosopoulos, G., Marshall, G. M., Walker, J., Rogers, M., et al. 2002. Germline mutation of the tumour suppressor PTEN in Proteus syndrome. *Journal of Medical Genetics,* 39**,** 937-940.
Smith, J. S., Tachibana, I., Passe, S. M., Huntley, B. K., Borell, T. J., Iturria, N., et al. 2001. PTEN mutation, EGFR amplification, and outcome in

patients with anaplastic astrocytoma and glioblastoma multiforme. *J Natl Cancer Inst,* 93**,** 1246-56.
Smith, T. R., Levine, E. A., Freimanis, R. I., Akman, S. A., Allen, G. O., Hoang, K. N., et al. 2008. Polygenic model of DNA repair genetic polymorphisms in human breast cancer risk. *Carcinogenesis,* 29**,** 2132-8.
Smoot, M. E., Ono, K., Ruscheinski, J., Wang, P. L. & Ideker, T. 2011. Cytoscape 2.8: new features for data integration and network visualization. *Bioinformatics,* 27**,** 431-2.
Smyth, G. K. 2004. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat Appl Genet Mol Biol,* 3**,** Article3.
Snietura, M., Jaworska, M., Mlynarczyk-Liszka, J., Goraj-Zajac, A., Piglowski, W., Lange, D., et al. 2012. PTEN as a Prognostic and Predictive Marker in Postoperative Radiotherapy for Squamous Cell Cancer of the Head and Neck. *PLoS One,* 7.
Song, M. S., Carracedo, A., Salmena, L., Song, S. J., Egia, A., Malumbres, M. & Pandolfi, P. P. 2011. Nuclear PTEN regulates the APC-CDH1 tumor-suppressive complex in a phosphatase-independent manner. *Cell,* 144**,** 187-99.
Stambolic, V., MacPherson, D., Sas, D., Lin, Y., Snow, B., Jang, Y., et al. 2001. Regulation of PTEN transcription by p53. *Molecular Cell,* 8**,** 317-25.
Stambolic, V., Suzuki, A., de la Pompa, J. L., Brothers, G. M., Mirtsos, C., Sasaki, T., et al. 1998. Negative Regulation of PKB/Akt-Dependent Cell Survival by the Tumor Suppressor PTEN. *Cell,* 95**,** 29-39.
Staszewski, O., Nikolova, T. & Kaina, B. 2008. Kinetics of gamma-H2AX focus formation upon treatment of cells with UV light and alkylating agents. *Environmental and Molecular Mutagenesis,* 49**,** 734-40.
Steck, P. A., Pershouse, M. A., Jasser, S. A., Yung, W. K. A., Lin, H., Ligon, A. H., et al. 1997. Identification of a candidate tumour suppressor gene, MMAC1, at chromosome 10q23.3 that is mutated in multiple advanced cancers. *Nature Genetics,* 15**,** 356-362.
Stemke-Hale, K., Gonzalez-Angulo, A. M., Lluch, A., Neve, R. M., Kuo, W.-L., Davies, M., et al. 2008. An Integrative Genomic and Proteomic Analysis of PIK3CA, PTEN, and AKT Mutations in Breast Cancer. *Cancer Research,* 68**,** 6084-6091.
Stiff, T., O'Driscoll, M., Rief, N., Iwabuchi, K., Lobrich, M. & Jeggo, P. A. 2004. ATM and DNA-PK function redundantly to phosphorylate H2AX after exposure to ionizing radiation. *Cancer Res,* 64**,** 2390-6.
Stupp, R., Mason, W. P., van den Bent, M. J., Weller, M., Fisher, B., Taphoorn, M. J. B., et al. 2005. Radiotherapy plus Concomitant and Adjuvant Temozolomide for Glioblastoma. *New England Journal of Medicine,* 352**,** 987-996.
Stupp, R., Tonn, J. C., Brada, M. & Pentheroudakis, G. 2010. High-grade malignant glioma: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Annals of Oncology,* 21 Suppl 5**,** v190-3.
Sturm, D., Witt, H., Hovestadt, V., Khuong-Quang, D. A., Jones, D. T., Konermann, C., et al. 2012. Hotspot mutations in H3F3A and IDH1

define distinct epigenetic and biological subgroups of glioblastoma. *Cancer Cell,* 22**,** 425-37.

Sulli, G., Di Micco, R. & d'Adda di Fagagna, F. 2012. Crosstalk between chromatin state and DNA damage response in cellular senescence and cancer. *Nat Rev Cancer,* 12**,** 709-20.

Sultana, R., Abdel-Fatah, T., Perry, C., Moseley, P., Albarakti, N., Mohan, V., et al. 2013. Ataxia telangiectasia mutated and Rad3 related (ATR) protein kinase inhibition is synthetically lethal in XRCC1 deficient ovarian cancer cells. *PLoS One,* 8**,** e57098.

Sultana, R., McNeill, D. R., Abbotts, R., Mohammed, M. Z., Zdzienicka, M. Z., Qutob, H., et al. 2012. Synthetic lethal targeting of DNA double-strand break repair deficient cells by human apurinic/apyrimidinic endonuclease inhibitors. *International journal of cancer. Journal international du cancer*.

Sung, P. & Klein, H. 2006. Mechanism of homologous recombination: mediators and helicases take on regulatory functions. *Nat Rev Mol Cell Biol,* 7**,** 739-50.

Swift, M., Morrell, D., Massey, R. B. & Chase, C. L. 1991. Incidence of cancer in 161 families affected by ataxia-telangiectasia. *N Engl J Med,* 325**,** 1831-6.

Tachibana, M., Shibakita, M., Ohno, S., Kinugasa, S., Yoshimura, H., Ueda, S., et al. 2002. Expression and prognostic significance of PTEN product protein in patients with esophageal squamous cell carcinoma. *Cancer,* 94**,** 1955-60.

Tan, E. Y., Campo, L., Han, C., Turley, H., Pezzella, F., Gatter, K. C., et al. 2007. Cytoplasmic location of factor-inhibiting hypoxia-inducible factor is associated with an enhanced hypoxic response and a shorter survival in invasive breast cancer. *Breast Cancer Research,* 9**,** R89.

Taniguchi, C. M., Winnay, J., Kondo, T., Bronson, R. T., Guimaraes, A. R., Aleman Jé, O., et al. 2010. The PI3K regulatory subunit p85α can exert tumor suppressor properties through negative regulation of growth factor signalling. *Cancer Res,* 70**,** 5305-15.

Tatsumoto, N., Hiyama, E., Murakami, Y., Imamura, Y., Shay, J. W., Matsuura, Y. & Yokoyama, T. 2000. High Telomerase Activity Is an Independent Prognostic Indicator of Poor Outcome in Colorectal Cancer. *Clinical Cancer Research,* 6**,** 2696-2701.

TCGA 2008. Comprehensive genomic characterization defines human glioblastoma genes and core pathways. *Nature,* 455**,** 1061-8.

Tell, G., Zecca, A., Pellizzari, L., Spessotto, P., Colombatti, A., Kelley, M. R., et al. 2000. An 'environment to nucleus' signaling system operates in B lymphocytes: redox status modulates BSAP/Pax-5 activation through Ref-1 nuclear translocation. *Nucleic Acids Res,* 28**,** 1099-105.

Tentori, L. & Graziani, G. 2002. Pharmacological strategies to increase the antitumor activity of methylating agents. *Curr Med Chem,* 9**,** 1285-301.

Tentori, L., Leonetti, C., Scarsella, M., D'Amati, G., Vergati, M., Portarena, I., et al. 2003. Systemic administration of GPI 15427, a novel poly(ADP-ribose) polymerase-1 inhibitor, increases the antitumor activity of temozolomide against intracranial melanoma, glioma, lymphoma. *Clin Cancer Res,* 9**,** 5370-9.

Thangapazham, R. L., Singh, A. K., Sharma, A., Warren, J., Gaddipati, J. P. & Maheshwari, R. K. 2007. Green tea polyphenols and its constituent epigallocatechin gallate inhibits proliferation of human breast cancer cells in vitro and in vivo. *Cancer Letters,* 245**,** 232-41.

ThermoScientific. 2009. *Thermo Scientific Open Biosystems Expression Arrest TRIPZ Lentiviral shRNAmir* [Online]. Available at: http://www.health.umn.edu/sites/default/files/tripz-lentiviral-inducible-shrnamir-manual.pdf [Accessed 27th January 2016].

ThermoScientific. 2015. *Overview of Immunohistochemistry* [Online]. Available at: https://www.thermofisher.com/uk/en/home/life-science/protein-biology/protein-biology-learning-center/protein-biology-resource-library/pierce-protein-methods/overview-immunohistochemistry.html [Accessed 24th January 2016].

Thompson, L. H. & Schild, D. 2002. Recombinational DNA repair and human disease. *Mutat Res,* 509**,** 49-78.

Tibshirani, R., Hastie, T., Narasimhan, B. & Chu, G. 2002. Diagnosis of multiple cancer types by shrunken centroids of gene expression. *Proceedings of the National Academy of Sciences of the United States of America,* 99**,** 6567-72.

Tomkiel, J., Cooke, C. A., Saitoh, H., Bernat, R. L. & Earnshaw, W. C. 1994. CENP-C is required for maintaining proper kinetochore size and for a timely transition to anaphase. *J Cell Biol,* 125**,** 531-45.

Tong, W. M., Ohgaki, H., Huang, H., Granier, C., Kleihues, P. & Wang, Z. Q. 2003. Null mutation of DNA strand break-binding molecule poly(ADP-ribose) polymerase causes medulloblastomas in p53(-/-) mice. *American Journal of Pathology,* 162**,** 343-52.

Torgovnick, A. & Schumacher, B. 2015. DNA repair mechanisms in cancer development and therapy. *Front Genet,* 6**,** 157.

Trivedi, R. N., Almeida, K. H., Fornsaglio, J. L., Schamus, S. & Sobol, R. W. 2005. The Role of Base Excision Repair in the Sensitivity and Resistance to Temozolomide-Mediated Cell Death. *Cancer Research,* 65**,** 6394-6400.

Tutt, A. & Ashworth, A. 2002. The relationship between the roles of BRCA genes in DNA repair and cancer predisposition. *Trends Mol Med,* 8**,** 571-6.

Ueno, M., Masutani, H., Arai, R. J., Yamauchi, A., Hirota, K., Sakai, T., et al. 1999. Thioredoxin-dependent redox regulation of p53-mediated p21 activation. *J Biol Chem,* 274**,** 35809-15.

Underhill, C., Toulmonde, M. & Bonnefoi, H. 2011. A review of PARP inhibitors: from bench to bedside. *Annals of Oncology,* 22**,** 268-79.

UniProt. 2015a. *UniProtKB - O43493 (TGON2_HUMAN)* [Online]. Available at: http://www.uniprot.org/uniprot/O43493 [Accessed 31st August 2015].

UniProt. 2015b. *UniProtKB - O60488 (ACSL4_HUMAN)* [Online]. Available at: http://www.uniprot.org/uniprot/O60488 [Accessed 31st August 2015].

UniProt. 2015c. *UniProtKB - P08174 (DAF_HUMAN)* [Online]. Available at: http://www.uniprot.org/uniprot/P08174 [Accessed 31st August 2015].

UniProt. 2015d. *UniProtKB - P10827 (THA_HUMAN)* [Online]. Available at: http://www.uniprot.org/uniprot/P10827 [Accessed 31st August 2015].
UniProt. 2015e. *UniProtKB - Q4G170 (Q4G170_HUMAN)* [Online]. Available at: http://www.uniprot.org/uniprot/Q4G170) [Accessed 31st August 2015].
UniProt. 2015f. *UniProtKB - Q8NFX7 (STXB6_HUMAN)* [Online]. Available at: http://www.uniprot.org/uniprot/Q8NFX7 [Accessed 31st August 2015].
UniProt. 2015g. *UniProtKB - Q9HCD6 (TANC2_HUMAN)* [Online]. Available at: http://www.uniprot.org/uniprot/Q9HCD6 [Accessed 31st August 2015].
UniProt. 2015h. *UniProtKB - Q9NWT6 (HIF1N_HUMAN)* [Online]. Available at: http://www.uniprot.org/uniprot/Q9NWT6 [Accessed 31st August 2015].
UniProt. 2015i. *UniProtKB - Q33E94 (RFX4_HUMAN)* [Online]. Available at: http://www.uniprot.org/uniprot/Q33E94 [Accessed 31st August 2015].
Uziel, T., Savitsky, K., Platzer, M., Ziv, Y., Helbitz, T., Nehls, M., et al. 1996. Genomic Organization of the ATM gene. *Genomics,* 33**,** 317-20.
van den Bent, M. J., Carpentier, A. F., Brandes, A. A., Sanson, M., Taphoorn, M. J., Bernsen, H. J., et al. 2006. Adjuvant procarbazine, lomustine, and vincristine improves progression-free survival but not overall survival in newly diagnosed anaplastic oligodendrogliomas and oligoastrocytomas: a randomized European Organisation for Research and Treatment of Cancer phase III trial. *J Clin Oncol,* 24**,** 2715-22.
van Gent, D. C. & van der Burg, M. 2007. Non-homologous end-joining, a sticky affair. *Oncogene,* 26**,** 7731-40.
Vascotto, C., Fantini, D., Romanello, M., Cesaratto, L., Deganuto, M., Leonardi, A., et al. 2009. APE1/Ref-1 interacts with NPM1 within nucleoli and plays a role in the rRNA quality control process. *Mol Cell Biol*.
Vascotto, C. F., M. L. 2012. Blockade of Base Excision Repair: Inhibition of Small Lesions Results in Big Consequences to Cancer Cells. *In:* Kelley, M. (ed.) *DNA Repair in Cancer Therapy: Molecular Targets and Clinical Applications.* 1st ed. London: Academic Press.
Verginelli, F., Perin, A., Dali, R., Fung, K. H., Lo, R., Longatti, P., et al. 2013. Transcription factors FOXG1 and Groucho/TLE promote glioblastoma growth. *Nat Commun,* 4.
Verhaak, R. G., Hoadley, K. A., Purdom, E., Wang, V., Qi, Y., Wilkerson, M. D., et al. 2010. Integrated genomic analysis identifies clinically relevant subtypes of glioblastoma characterized by abnormalities in PDGFRA, IDH1, EGFR, and NF1. *Cancer Cell,* 17**,** 98-110.
Viniegra, J. G., Martinez, N., Modirassari, P., Hernandez Losa, J., Parada Cobo, C., Sanchez-Arevalo Lobo, V. J., et al. 2005. Full activation of PKB/Akt in response to insulin or ionizing radiation is mediated through ATM. *J Biol Chem,* 280**,** 4029-36.
Vital, A. L., Tabernero, M. D., Castrillo, A., Rebelo, O., Tao, H., Gomes, F., et al. 2010. Gene expression profiles of human glioblastomas are associated with both tumor cytogenetics and histopathology. *Neuro Oncol,* 12**,** 991-1003.

Vivanco, I. & Sawyers, C. L. 2002. The phosphatidylinositol 3-Kinase AKT pathway in human cancer. *Nat Rev Cancer,* 2**,** 489-501.
Wang, D., Luo, M. & Kelley, M. R. 2004. Human apurinic endonuclease 1 (APE1) expression and prognostic significance in osteosarcoma: enhanced sensitivity of osteosarcoma to DNA damaging agents using silencing RNA APE1 expression inhibition. *Mol Cancer Ther,* 3**,** 679-86.
Wang, E., Zhang, C., Polavaram, N., Liu, F., Wu, G., Schroeder, M. A., et al. 2014. The Role of Factor Inhibiting HIF (FIH-1) in Inhibiting HIF-1 Transcriptional Activity in Glioblastoma Multiforme. *PLoS One,* 9**,** e86102.
Wang, M., Qin, C., Zhu, J., Yuan, L., Fu, G., Zhang, Z. & Yin, C. 2010a. Genetic variants of XRCC1, APE1, and ADPRT genes and risk of bladder cancer. *DNA and cell biology,* 29**,** 303-11.
Wang, P., Henning, S. M. & Heber, D. 2010b. Limitations of MTT and MTS-Based Assays for Measurement of Antiproliferative Activity of Green Tea Polyphenols. *PLoS One,* 5.
Wang, Q., Zhang, H., Fishel, R. & Greene, M. I. 2000. BRCA1 and cell signaling. *Oncogene,* 19**,** 6152-8.
Watson, A. J., Middleton, M. R., McGown, G., Thorncroft, M., Ranson, M., Hersey, P., et al. 2009. O(6)-methylguanine-DNA methyltransferase depletion and DNA damage in patients with melanoma treated with temozolomide alone or with lomeguatrib. *Br J Cancer,* 100**,** 1250-6.
Weber, A. M. & Ryan, A. J. 2015. ATM and ATR as therapeutic targets in cancer. *Pharmacology & Therapeutics,* 149**,** 124-138.
Wechsler, T., Newman, S. & West, S. C. 2011. Aberrant chromosome morphology in human cells defective for Holliday junction resolution. *Nature,* 471**,** 642-6.
Weinstein, J. N., Collisson, E. A., Mills, G. B., Shaw, K. R., Ozenberger, B. A., Ellrott, K., et al. 2013. The Cancer Genome Atlas Pan-Cancer analysis project. *Nature Genetics,* 45**,** 1113-20.
Weng, L.-P., Brown, J. L. & Eng, C. 2001. PTEN coordinates G1 arrest by down-regulating cyclin D1 via its protein phosphatase activity and up-regulating p27 via its lipid phosphatase activity in a breast cancer model. *Human Molecular Genetics,* 10**,** 599-604.
Wick, W., Furnari, F. B., Naumann, U., Cavenee, W. K. & Weller, M. 1999. PTEN gene transfer in human malignant glioma: sensitization to irradiation and CD95L-induced apoptosis. *Oncogene,* 18**,** 3936-43.
Wick, W., Platten, M., Meisner, C., Felsberg, J., Tabatabai, G., Simon, M., et al. 2012. Temozolomide chemotherapy alone versus radiotherapy alone for malignant astrocytoma in the elderly: the NOA-08 randomised, phase 3 trial. *Lancet Oncol,* 13**,** 707-15.
Wiencke, J. K., Zheng, S., Jelluma, N., Tihan, T., Vandenberg, S., Tamguney, T., et al. 2007. Methylation of the PTEN promoter defines low-grade gliomas and secondary glioblastoma. *Neuro Oncol,* 9**,** 271-9.
Willers, H., Pfaeffle, HN, Zou, L 2012. Targeting Homologous Recombination Repair in Cancer. *In:* Kelley, M. R. (ed.) *DNA Repair in Cancer Therapy: Molecular Targets and Clinical Applications.* 1st ed. London: Academic Press.

Worby, C. A. & Dixon, J. E. 2014. Pten. *Annual Review of Biochemistry,* 83**,** 641-69.

Wu, B., Liu, H. L., Zhang, S., Dong, X. R. & Wu, G. 2012. Lack of an Association between Two BER Gene Polymorphisms and Breast Cancer Risk: A Meta-Analysis. *PLoS One,* 7**,** e50857.

Xanthoudakis, S., Miao, G., Wang, F., Pan, Y. C. & Curran, T. 1992. Redox activation of Fos-Jun DNA binding activity is mediated by a DNA repair enzyme. *EMBO J,* 11**,** 3323-35.

Xia, L., Huang, W., Tian, D., Chen, Z., Zhang, L., Li, Y., et al. 2014. ACP5, a direct transcriptional target of FoxM1, promotes tumor metastasis and indicates poor prognosis in hepatocellular carcinoma. *Oncogene,* 33**,** 1395-1406.

Yang, M. H., Chiang, W. C., Chou, T. Y., Chang, S. Y., Chen, P. M., Teng, S. C. & Wu, K. J. 2006. Increased NBS1 expression is a marker of aggressive head and neck cancer and overexpression of NBS1 contributes to transformation. *Clin Cancer Res,* 12**,** 507-15.

Ye, C., Cai, Q., Dai, Q., Shu, X. O., Shin, A., Gao, Y. T. & Zheng, W. 2007. Expression patterns of the ATM gene in mammary tissues and their associations with breast cancer survival. *Cancer,* 109**,** 1729-35.

Yelamos, J., Farres, J., Llacuna, L., Ampurdanes, C. & Martin-Caballero, J. 2011. PARP-1 and PARP-2: New players in tumour development. *American Journal of Cancer Research,* 1**,** 328-346.

Yu, J., Zhang, Y., McIlroy, J., Rordorf-Nikolic, T., Orr, G. A. & Backer, J. M. 1998. Regulation of the p85/p110 phosphatidylinositol 3'-kinase: stabilization and inhibition of the p110alpha catalytic subunit by the p85 regulatory subunit. *Mol Cell Biol,* 18**,** 1379-87.

Yuan, T. L. & Cantley, L. C. 2008. PI3K pathway alterations in cancer: variations on a theme. *Oncogene,* 27**,** 5497-5510.

Zaremba, T., Thomas, H. D., Cole, M., Coulthard, S. A., Plummer, E. R. & Curtin, N. J. 2011. Poly(ADP-ribose) polymerase-1 (PARP-1) pharmacogenetics, activity and expression analysis in cancer patients and healthy volunteers. *Biochem J,* 436**,** 671-9.

Zhang, J., Stevens, M. F., Laughton, C. A., Madhusudan, S. & Bradshaw, T. D. 2010. Acquired resistance to temozolomide in glioma cell lines: molecular mechanisms and potential translational applications. *Oncology,* 78**,** 103-14.

Zhang, K., Wang, X. Q., Zhou, B. & Zhang, L. 2013. The prognostic value of MGMT promoter methylation in Glioblastoma multiforme: a meta-analysis. *Fam Cancer,* 12**,** 449-58.

Zhang, L. 2000. Molecular alterations in nucleotide excision repair genes in malignant glioma and their relevance to malignant progression and drug resistance. University of Texas.

Zhang, X. P., Liu, F. & Wang, W. 2011. Two-phase dynamics of p53 in the DNA damage response. *Proc Natl Acad Sci U S A,* 108**,** 8990-5.

Zhao, P., Zou, P., Zhao, L., Yan, W., Kang, C., Jiang, T. & You, Y. 2013. Genetic polymorphisms of DNA double-strand break repair pathway genes and glioma susceptibility. *Bmc Cancer,* 13**,** 234.

Zhao, Q., Wang, W., Zhang, Z., Wang, S., Wang, M., Zhou, J., et al. 2011. A genetic variation in APE1 is associated with gastric cancer survival in a Chinese population. *Cancer Science,* 102**,** 1293-7.
Zhou, K., Hu, D., Lu, J., Fan, W., Liu, H., Chen, H., et al. 2011. A genetic variant in the APE1/Ref-1 gene promoter -141T/G may modulate risk of glioblastoma in a Chinese Han population. *Bmc Cancer,* 11**,** 104.
Zhou, X.-P., Gimm, O., Hampel, H., Niemann, T., Walker, M. J. & Eng, C. Epigenetic PTEN Silencing in Malignant Melanomas without PTEN Mutation. *The American Journal of Pathology,* 157**,** 1123-1128.

CHAPTER 8

Appendices

8 Appendices

Appendix 1: List of the 188 DNA repair genes studied and the DNA repair pathway in which they are involved.

DNA repair gene	Pathway
CCNO	Base excision repair
MBD4	Base excision repair
MPG	Base excision repair
MUTYH (MYH)	Base excision repair
NEIL1	Base excision repair
NEIL2	Base excision repair
NEIL3	Base excision repair
NTHL1 (NTH1)	Base excision repair
OGG1	Base excision repair
SMUG1	Base excision repair
TDG	Base excision repair
UNG	Base excision repair
APEX1 (APE1)	Base excision repair and strand break joining factors
APEX2	Base excision repair and strand break joining factors
APLF (C2ORF13)	Base excision repair and strand break joining factors
LIG3	Base excision repair and strand break joining factors
PNKP	Base excision repair and strand break joining factors
XRCC1	Base excision repair and strand break joining factors
PARP1 (ADPRT)	Poly(ADP-ribose) polymerase (PARP) enzymes that bind to DNA
PARP2 (ADPRTL2)	Poly(ADP-ribose) polymerase (PARP) enzymes that bind to DNA
PARP3 (ADPRTL3)	Poly(ADP-ribose) polymerase (PARP) enzymes that bind to DNA
ALKBH2 (ABH2)	Direct reversal of damage
ALKBH3 (DEPC1)	Direct reversal of damage
MGMT	Direct reversal of damage
TDP1	Repair of DNA-topoisomerase crosslinks
TDP2 (TTRAP)	Repair of DNA-topoisomerase crosslinks
MLH1	Mismatch excision repair
MLH3	Mismatch excision repair
MSH2	Mismatch excision repair
MSH3	Mismatch excision repair
MSH4	Mismatch excision repair
MSH5	Mismatch excision repair
MSH6	Mismatch excision repair

PMS1	Mismatch excision repair
PMS2	Mismatch excision repair
PMS2L3	Mismatch excision repair
CCNH	Nucleotide excision repair
CDK7	Nucleotide excision repair
CETN2	Nucleotide excision repair
DDB1	Nucleotide excision repair
DDB2 (XPE)	Nucleotide excision repair
ERCC1	Nucleotide excision repair
ERCC2 (XPD)	Nucleotide excision repair
ERCC3 (XPB)	Nucleotide excision repair
ERCC4 (XPF)	Nucleotide excision repair
ERCC5 (XPG)	Nucleotide excision repair
ERCC6 (CSB)	Nucleotide excision repair
ERCC8 (CSA)	Nucleotide excision repair
GTF2H1	Nucleotide excision repair
GTF2H2	Nucleotide excision repair
GTF2H3	Nucleotide excision repair
GTF2H4	Nucleotide excision repair
GTF2H5 (TTDA)	Nucleotide excision repair
LIG1	Nucleotide excision repair
MMS19	Nucleotide excision repair
MNAT1	Nucleotide excision repair
RAD23A	Nucleotide excision repair
RAD23B	Nucleotide excision repair
RPA1	Nucleotide excision repair
RPA2	Nucleotide excision repair
RPA3	Nucleotide excision repair
SLK	Nucleotide excision repair
TFIIH	Nucleotide excision repair
UVSSA (KIAA1530)	Nucleotide excision repair
XAB2 (HCNP)	Nucleotide excision repair
XPA	Nucleotide excision repair
XPC	Nucleotide excision repair
BRCA1	Homologous recombination
DMC1	Homologous recombination
EME1 (MMS4L)	Homologous recombination
EME2	Homologous recombination
GEN1	Homologous recombination
GIYD1 (SLX1A)	Homologous recombination
GIYD2 (SLX1B)	Homologous recombination
MRE11A	Homologous recombination
MUS81	Homologous recombination
NBN (NBS1)	Homologous recombination

RAD21	Homologous recombination
RAD50	Homologous recombination
RAD51	Homologous recombination
RAD51B	Homologous recombination
RAD51D	Homologous recombination
RAD52	Homologous recombination
RAD54B	Homologous recombination
RAD54L	Homologous recombination
RBBP8 (CtIP)	Homologous recombination
SHFM1 (DSS1)	Homologous recombination
XRCC2	Homologous recombination
XRCC3	Homologous recombination
BRCA2 (FANCD1)	Fanconi anaemia
BRIP1 (FANCJ)	Fanconi anaemia
BTBD12 (SLX4) (FANCP)	Fanconi anaemia
FAAP20 (C1orf86)	Fanconi anaemia
FAAP24 (C19orf40)	Fanconi anaemia
FANCA	Fanconi anaemia
FANCB	Fanconi anaemia
FANCC	Fanconi anaemia
FANCD2	Fanconi anaemia
FANCE	Fanconi anaemia
FANCF	Fanconi anaemia
FANCG (XRCC9)	Fanconi anaemia
FANCI (KIAA1794)	Fanconi anaemia
FANCL	Fanconi anaemia
FANCM	Fanconi anaemia
PALB2 (FANCN)	Fanconi anaemia
RAD51C (FANCO)	Fanconi anaemia
DCLRE1C (Artemis)	Non-homologous end-joining
DUT	Non-homologous end-joining
LIG4	Non-homologous end-joining
NHEJ1 (XLF, Cernunnos)	Non-homologous end-joining
NUDT1 (MTH1)	Non-homologous end-joining
PRKDC	Non-homologous end-joining
RRM2B (p53R2)	Non-homologous end-joining
XRCC4	Non-homologous end-joining
XRCC5 (Ku80)	Non-homologous end-joining

XRCC6 (Ku70)	Non-homologous end-joining
MAD2L2 (REV7)	DNA polymerases (catalytic subunits)
PCNA	DNA polymerases (catalytic subunits)
POLB	DNA polymerases (catalytic subunits)
POLD1	DNA polymerases (catalytic subunits)
POLD3	DNA polymerases (catalytic subunits)
POLE	DNA polymerases (catalytic subunits)
POLG	DNA polymerases (catalytic subunits)
POLH	DNA polymerases (catalytic subunits)
POLI (RAD30B)	DNA polymerases (catalytic subunits)
POLK (DINB1)	DNA polymerases (catalytic subunits)
POLL	DNA polymerases (catalytic subunits)
POLM	DNA polymerases (catalytic subunits)
POLN (POL4P)	DNA polymerases (catalytic subunits)
POLQ	DNA polymerases (catalytic subunits)
REV1L (REV1)	DNA polymerases (catalytic subunits)
REV3L (POLZ)	DNA polymerases (catalytic subunits)
APTX (aprataxin)	Editing and processing nucleases
ENDOV	Editing and processing nucleases
EXO1 (HEX1)	Editing and processing nucleases
FAN1 (MTMR15)	Editing and processing nucleases
FEN1 (DNase IV)	Editing and processing nucleases
SPO11	Editing and processing nucleases
TREX1 (DNase III)	Editing and processing nucleases
TREX2	Editing and processing nucleases
HLTF (SMARCA3)	Ubiquitination and modification
RAD18	Ubiquitination and modification
RNF168	Ubiquitination and modification
RNF4	Ubiquitination and modification
RNF8	Ubiquitination and modification
SHPRH	Ubiquitination and modification
SPRTN (c1orf124)	Ubiquitination and modification
UBE2A (RAD6A)	Ubiquitination and modification
UBE2B (RAD6B)	Ubiquitination and modification
UBE2N (UBC13)	Ubiquitination and modification
UBE2V2 (MMS2)	Ubiquitination and modification
CHAF1A (CAF1)	Chromatin Structure and Modification
H2AFX (H2AX)	Chromatin Structure and Modification
SETMAR (METNASE)	Chromatin Structure and Modification
ATM	Genes defective in diseases associated with sensitivity to DNA damaging agents
BLM	Genes defective in diseases associated with sensitivity to DNA damaging agents

RECQL4	Genes defective in diseases associated with sensitivity to DNA damaging agents
TTDN1 (C7orf11)	Genes defective in diseases associated with sensitivity to DNA damaging agents
WRN	Genes defective in diseases associated with sensitivity to DNA damaging agents
DCLRE1A (SNM1)	Other identified genes with known or suspected DNA repair function
DCLRE1B (SNM1B)	Other identified genes with known or suspected DNA repair function
HELQ (HEL308)	Other identified genes with known or suspected DNA repair function
OBFC2B (SSB1)	Other identified genes with known or suspected DNA repair function
PRPF19 (PSO4)	Other identified genes with known or suspected DNA repair function
RDM1 (RAD52B)	Other identified genes with known or suspected DNA repair function
RECQL (RECQ1)	Other identified genes with known or suspected DNA repair function
RECQL5	Other identified genes with known or suspected DNA repair function
RPA4	Other identified genes with known or suspected DNA repair function
ATR	Other conserved DNA damage response genes
ATRIP	Other conserved DNA damage response genes
CHEK1	Other conserved DNA damage response genes
CHEK2	Other conserved DNA damage response genes
CLK2	Other conserved DNA damage response genes
HUS1	Other conserved DNA damage response genes
MDC1	Other conserved DNA damage response genes
PER1	Other conserved DNA damage response genes
RAD1	Other conserved DNA damage response genes
RAD17 (RAD24)	Other conserved DNA damage response genes
RAD9A	Other conserved DNA damage response genes
RFC1	Other conserved DNA damage response genes
RIF1	Other conserved DNA damage response genes
TOP3A	Other conserved DNA damage response genes
TOP3B	Other conserved DNA damage response genes
TOPBP1	Other conserved DNA damage response genes
TP53	Other conserved DNA damage response genes
TP53BP1 (53BP1)	Other conserved DNA damage response genes
XRCC6BP1	Other conserved DNA damage response genes

Appendix 2: Univariate associations between DNA repair gene expression and survival in both the Test and TCGA Datasets

Base excision repair	Test Dataset		TCGA Dataset	
	Log rank p value	Expression level	Log rank p value	Expression level
APEX1				
210027_s_at	**0.003***	**Low**	**0.000018***	**Low**
APEX2				
204408_at	0.003	High	0.016	Low
CCNO				
210021_s_at	0.001	High	0.069	No association
FEN1				
204767_s_at	0.006	High	0.015	Low
204768_s_at	0.023	High	0.025	Low
LIG3				
204123_at	0.058	No association	0.115	No association
207348_s_at	0.035	High	0.010	Low
MBD4				
209579_s_at	0.165	No association	0.001	High
214047_s_at	0.071	No association	0.000326	High
209580_s_at	**0.018**	**High**	**0.005**	**High**
214048_s_at	0.150	No association	0.038	High
MPG				
203686_at	**0.023**	**High**	**0.007**	**High**
MUTYH				
207727_s_at	0.071	No association	0.088	No association
NEIL1				
219396_s_at	0.388	No association	0.046	High
NEIL3				
219502_at	0.068	No association	0.096	No association
NTHL1				
209731_at	0.166	No association	0.000256	High
OGG1				
205301_s_at	0.023	High	0.829	No association
205760_s_at	0.011	Low	0.338	No association
215272_at	0.034	High	0.058	No association
PARP1				
208644_at	0.025	High	0.140	No association
PARP2				
214086_s_at	**0.020**	**Low**	**0.002**	**Low**
215773_x_at	**0.007**	**Low**	**0.043**	**Low**
204752_x_at	**0.014**	**Low**	**0.014**	**Low**
PARP3				
209940_at	**0.002**	**High**	**0.004**	**High**

Base excision repair	Test Dataset		TCGA Dataset	
	Log rank p value	Expression level	Log rank p value	Expression level
SMUG1				
218685_s_at	0.148	No association	0.213	No association
TDG				
203743_s_at	0.151	No association	0.032	Low
203742_s_at	0.063	No association	0.151	No association
UNG				
202330_s_at	0.004	High	0.059	Low
XRCC1				
203655_at	0.028	High	0.225	No association

Significant associations after BH FDR correction are shown in bold

Nucleotide excision repair	Test Dataset		TCGA Dataset	
	Log rank p value	Expression level	Log rank P value	Expression level
BRIP1				
221703_at	0.000377	High	0.010	Low
CCNH				
204093_at	0.069	No association	0.016	Low
CDK7				
211297_s_at	0.224	No association	0.217	No association
CETN2				
209194_at	0.018	High	0.056	No association
DDB1				
208619_at	0.075	No association	0.005	Low
DDB2				
203409_at	**0.000097***	**High**	**0.00005***	**High**
ERCC1				
203719_at	0.030	Low	0.164	No association
203720_s_at	0.069	No association	0.016	Low
ERCC2				
213468_at	0.093	No association	0.077	No association
ERCC3				
202176_at	0.036	High	0.019	Low
ERCC4				
210158_at	0.061	No association	0.367	No association
ERCC5				
202414_at	0.201	No association	0.010	Low
ERCC6				
207347_at	**0.006**	**Low**	**0.010**	**Low**
ERCC8				
205162_at	0.084	No association	0.096	No association
GTF2H1				
202451_at	0.172	No association	0.076	No association
202453_s_at	0.174	No association	0.182	No association

Nucleotide excision repair	Test Dataset		TCGA Dataset	
	Log rank p value	Expression level	Log rank P value	Expression level
GTF2H2				
221540_x_at	0.032	Low	0.180	No association
GTF2H3				
222104_x_at	0.051	No association	0.015	High
GTF2H4				
203577_at	0.265	No association	0.001	Low
GTF2H5				
213357_at	0.125	No association	0.129	No association
LIG1				
202726_at	0.001	High	0.230	No association
MMS19				
202167_s_at	0.160	No association	0.000370	High
MNAT1				
203565_s_at	0.087	No association	0.003	Low
PNKP				
218961_s_at	0.166	No association	0.211	No association
POLL				
221049_s_at	0.056	No association	0.000223	Low
RAD23A				
201039_s_at	0.202	No association	0.159	No association
201046_s_at	0.132	No association	0.277	No association
RAD23B				
201222_s_at	**0.000254***	**High**	**0.006***	**High**
201223_s_at	0.096	No association	0.059	No association
214422_at	0.007	Low	0.154	No association
RPA1				
201528_at	0.044	High	0.051	No association
201529_s_at	**0.019**	**High**	**0.038**	**High**
RPA2				
201756_at	0.014	High	0.168	No association
RPA3				
209507_at	0.002	High	0.067	No association
SLK				
206875_s_at	0.087	No association	0.111	No association
XAB2				
218110_at	0.064	No association	0.024	Low

Nucleotide excision repair	Test Dataset		TCGA Dataset	
	Log rank p value	Expression level	Log rank p value	Expression level
XPC				
209375_at	**0.000067**	**High**	**0.041**	**High**
XPA				
205672_at	**0.027**	**Low**	**0.016**	**Low**

Significant associations after BH FDR correction are shown in bold

Mismatch repair	Test Dataset		TCGA Dataset	
	Log rank p value	Expression level	Log rank p value	Expression level
MLH1				
202520_s_at	0.012	High	0.182	No association
MLH3				
204838_s_at	0.259	No association	0.106	No association
214525_x_at	0.006	Low	0.009	High
217216_x_at	0.109	No association	0.000487	High
MSH2				
209421_at	0.132	No association	0.011	Low
MSH3				
205887_x_at	0.032	High	0.097	No association
210947_s_at	0.284	No association	0.175	No association
MSH4				
210533_at	0.084	No association	0.069	No association
MSH5				
221406_s_at	0.004	High	0.146	No association
210410_s_at	0.001	High	0.183	No association
212913_at	0.089	No association	0.066	No association
MSH6				
202911_at	0.030	High	0.017	Low
211449_at	0.160	No association	0.151	No association
211450_s_at	0.012	High	0.009	Low
PMS1				
213677_s_at	0.118	No association	0.000425	Low
PMS2				
209805_at	**0.007***	**High**	**0.011***	**High**
221206_at	0.053	No association	0.035	High
POLD3				
212836_at	0.002	High	0.052	No association
MLH1				
202520_s_at	0.012	High	0.182	No association
MLH3				
204838_s_at	0.259	No association	0.106	No association
214525_x_at	0.006	Low	0.009	High
217216_x_at	0.109	No association	0.000487	High

Mismatch repair	Test Dataset		TCGA Dataset	
	Log rank p value	Expression level	Log rank p value	Expression level
MSH2				
209421_at	0.132	No association	0.011	Low
MSH3				
205887_x_at	0.032	High	0.097	No association
210947_s_at	0.284	No association	0.175	No association
MSH4				
210533_at	0.084	No association	0.069	No association
MSH5				
221406_s_at	0.004	High	0.146	No association
210410_s_at	0.001	High	0.183	No association
212913_at	0.089	No association	0.066	No association
MSH6				
202911_at	0.030	High	0.017	Low
211449_at	0.160	No association	0.151	No association
211450_s_at	0.012	High	0.009	Low
PMS1				
213677_s_at	0.118	No association	0.000425	Low
PMS2				
209805_at	**0.007***	**High**	**0.011***	**High**
221206_at	0.053	No association	0.035	High
POLD3				
212836_at	0.002	High	0.052	No association

Significant associations after BH FDR correction are shown in bold

Homologous recombination	Test Dataset		TCGA Dataset	
	Log rank p value	Expression level	Log rank p value	Expression level
BRCA1				
204531_s_at	0.015	High	0.064	No association
211851_x_at	0.000469	High	0.080	No association
DMC1				
208382_s_at	0.125	No association	0.104	No association
208386_x_at	0.074	No association	0.147	No association
MRE11A				
205395_s_at	0.072	No association	0.126	No association
211334_at	0.059	No association	0.006	Low
RAD21				
200607_s_at	**0.001***	**Low**	**0.004***	**Low**
200608_s_at	**0.005**	**Low**	**0.028**	**Low**
RAD50				
209349_at	0.142	No association	0.132	No association
208393_s_at	0.073	No association	0.008	High

Homologous recombination	Test Dataset		TCGA Dataset	
	Log rank p value	**Expression level**	**Log rank p value**	**Expression level**
RAD51				
205023_at	0.073	No association	0.001	Low
205024_s_at	0.001	High	0.033	Low
RAD51C				
206066_s_at	0.063	No association	0.021	Low
209849_s_at	0.000329	High	0.063	Low
RAD51D				
37793_r_at	**0.005**	**High**	**0.036**	**High**
RAD52				
205647_at	0.009	High	0.003	Low
210630_s_at	0.004	High	0.067	No association
211904_x_at	0.120	No association	0.065	No association
RAD54L				
204558_at	0.010	High	0.005	Low
XRCC2				
207598_x_at	0.078	No association	0.376	No association
XRCC3				
216299_s_at	0.031	High	0.004	Low
MUS81				
218463_s_at	0.027	High	0.000006	Low
NBN (NBS1)				
202905_x_at	**0.000001***	**Low**	**0.001***	**Low**
202906_x_at	0.002	Low	0.160	No association
202907_x_at	**0.001***	**Low**	**0.000125***	**Low**
217299_s_at	**0.000006***	**Low**	**0.006***	**Low**
RAD54B				
219494_at	0.024	High	0.076	No association
220549_at	0.064	No association	0.011	Low
RBBP8 (CtIP)				
203344_s_at	0.019	High	0.025	Low
SHFM1 (DSS1)				
202276_at	0.020	High	0.117	No association
DCLRE1C (Artemis)				
222233_s_at	**0.049**	**Low**	**0.003**	**Low**
219678_x_at	**0.050**	**Low**	**0.004**	**Low**
LIG4				
206235_at	0.003	High	0.238	No association
NHEJ1 (XLF)				
219418_at	0.093	No association	0.010	Low
PRKDC				
208694_at	0.257	No association	0.401	No association
210543_s_at	0.263	No association	0.058	No association
215757_at	0.061	No association	0.070	No association

Non-Homologous end-joining	Test Dataset		TCGA Dataset	
	Log rank p value	**Expression level**	**Log rank p value**	**Expression level**
XRCC4				
210812_at	0.204	No association	0.025	Low
205071_x_at	0.072	No association	0.229	No association
205072_s_at	0.014	Low	0.069	No association
210813_s_at	**0.047**	**Low**	**0.008**	**Low**
XRCC5 (Ku80)				
208642_s_at	0.029	High	0.001	Low
208643_s_at	0.112	No association	0.000383	Low
XRCC6 (Ku70)				
200792_at	0.028	Low	0.172	No association
215308_at	0.025	Low	0.154	No association

Significant associations after BH FDR correction are shown in bold

Modulation of nucleotide pools	Test Dataset		TCGA Dataset	
	Log rank p value	**Expression level**	**Log rank p value**	**Expression level**
DUT				
208955_at	0.018	High	0.011	Low
208956_x_at	0.087	No association	0.125	No association
209932_s_at	0.103	No association	0.190	No association
NUDT1 (MTH1)				
204766_s_at	0.029	High	0.001	Low

Significant associations after BH FDR correction are shown in bold

DNA polymerases	Test Dataset		TCGA Dataset	
	Log rank p value	Expression level	Log rank p value	Expression level
PCNA 201202_at	0.012	High	0.061	No association
POLB 203616_at	0.764	No association	0.003	Low
POLD1 203422_at	**0.030**	**Low**	**0.011**	**Low**
POLE 216026_s_at	0.00027	High	0.092	No association
POLG 203366_at	0.030	High	0.068	No association
217635_s_at	0.013	High	0.005	Low
217636_at	0.008	High	0.119	No association
POLH 219380_x_at	0.201	No association	0.298	No association
POLI (RAD30B) 219317_at	0.189	No association	0.042	High
POLM 222238_s_at	0.068	No association	0.322	No association
POLQ 219510_at	0.019	High	0.157	No association
207746_at	0.196	No association	0.084	No association
REV1L (REV1) 218428_s_at	0.098	No association	0.044	Low
REV3L (POLZ) 208070_s_at	0.077	No association	0.040	Low

Significant associations after BH FDR correction are shown in bold

Fanconi anaemia	Test Dataset		TCGA Dataset	
	Log rank p value	**Expression level**	**Log rank p value**	**Expression level**
BRCA2				
208368_s_at	0.070	No association	0.027	Low
214727_at	**0.009**	**High**	**0.002**	**High**
FAAP24				
214816_x_at	0.220	No association	0.032	High
FANCA				
203806_s_at	0.001	High	0.229	No association
203805_s_at	0.000321	High	0.051	No association
215530_at	0.064	No association	0.004	Low
FANCC				
205189_s_at	0.058	No association	0.192	No association
FANCE				
220255_at	0.270	No association	0.014	Low
FANCF				
218689_at	0.076	No association	0.006	High
FANCG (XRCC9)				
203564_at	0.003	High	0.003	Low
FANCI				
213007_at	0.003	High	0.066	No association
213008_at	0.004	High	0.110	No association
FANCL				
218397_at	0.081	No association	0.017	Low
PALB2				
219530_at	0.030	High	0.005	Low

Significant associations after BH FDR correction are shown in bold

Repair of DNA-topoisomerase crosslinks	Test Dataset		TCGA Dataset	
	Log rank p value	**Expression level**	**Log rank p value**	**Expression level**
TDP1				
219715_s_at	0.021	High	0.036	Low
TDP2 (TTRAP)				
202266_at	0.046	Low	0.096	No association

Ubiquitination and modification	Test Dataset		TCGA Dataset	
	Log rank p value	Expression level	Log rank p value	Expression level
HLTF (SMARCA3)				
202983_at	0.187	No association	0.019	Low
RNF4				
212696_s_at	0.225	No association	0.100	No association
RNF8				
203160_s_at	0.062	No association	0.013	Low
203161_s_at	0.328	No association	0.163	No association
UBE2A (RAD6A)				
201898_s_at	**0.037**	**High**	**0.032**	**High**
201899_s_at	0.124	No association	0.039	High
UBE2B (RAD6B)				
202333_s_at	0.029	Low	0.083	No association
202334_s_at	0.033	Low	0.062	No association
202335_s_at	0.008	Low	0.090	No association
211763_s_at	0.190	No association	0.214	No association
UBE2N (UBC13)				
201523_x_at	0.117	No association	0.145	No association
201524_x_at	0.189	No association	0.062	No association
212751_at	0.065	No association	0.135	No association
UBE2V2 (MMS2)				
209096_at	0.125	No association	0.038	High

Significant associations after BH FDR correction are shown in bold

Chromatin structure and modification	Test Dataset		TCGA Dataset	
	Log rank p value	Expression level	Log rank p value	Expression level
CHAF1A (CAF1)				
214426_x_at	0.012	High	0.011	Low
203975_s_at	0.040	High	0.020	Low
203976_s_at	0.001	High	0.013	Low
H2AFX (H2AX)				
205436_s_at	0.014	High	0.007	Low
212525_s_at	0.024	High	0.014	Low
212524_x_at	0.106	No association	0.058	No association
213344_s_at	0.035	High	0.006	Low
SETMAR				
206554_x_at	**0.004**	**Low**	**0.020**	**Low**

Significant associations after BH FDR correction are shown in bold

DNA damage response	Test Dataset		TCGA Dataset	
	Log rank p value	Expression level	Log rank p value	Expression level
ATM				
208442_s_at	0.008	Low	0.052	No association
210858_x_at	0.009	Low	0.241	No association
212672_at	0.079	No association	0.129	No association
BLM				
205733_at	**0.019**	**High**	**0.014**	**High**
RECQL4				
213520_at	0.015	High	0.002	Low
WRN				
205667_at	0.268	No association	0.179	No association

Significant associations after BH FDR correction are shown in bold

Editing and processing nucleases	Test Dataset		TCGA Dataset	
	Log rank p value	Expression level	Log rank p value	Expression level
APTX (aprataxin)				
218527_at	0.015	High	0.110	Low
EXO1 (HEX1)				
204603_at	0.007	High	0.022	Low
FAN1 (MTMR15)				
203678_at	0.060	No association	0.230	No association
SPO11				
222259_s_at	0.000223	High	0.240	No association
TREX1 (DNase III)				
34689_at	0.041	High	0.218	No association
205875_s_at	0.025	High	0.152	No association
TREX2				
211788_s_at	0.005	High	0.096	No association
207891_s_at	0.117	No association	0.002	High

Significant associations after BH FDR correction are shown in bold

Other	Test Dataset		TCGA Dataset	
	Log rank P value	Expression level	Log rank P value	Expression level
ATR				
209902_at	**0.021**	**High**	**0.049**	**High**
209903_s_at	0.157	No association	0.176	No association
CHEK1				
205394_at	0.000317	High	0.039	Low
205393_s_at	0.001	High	0.020	Low
CHEK2				
210416_s_at	0.011	High	0.009	Low
CLK2				
203229_s_at	0.027	High	0.028	Low
DCLRE1A				
209804_at	**0.045**	**Low**	**0.001**	**Low**
DCLRE1B				
219490_s_at	0.009	High	0.012	Low
HUS1				
204883_s_at	0.056	No association	0.021	High
204884_s_at	0.063	No association	0.022	High
217618_x_at	0.025	Low	0.011	High
MDC1				
203061_s_at	0.043	High	0.034	Low
203062_s_at	0.003	High	0.005	Low

Other	Test Dataset		TCGA Dataset	
	Log rank P value	**Expression level**	**Log rank P value**	**Expression level**
MGMT				
204880_at	**0.003**	**High**	**0.001**	**High**
OBFC2B				
218903_s_at	0.021	Low	0.140	No association
PER1				
202861_at	0.032	High	0.136	No association
36829_at	0.037	Low	0.168	No association
PRPF19				
203103_s_at	0.061	No association	0.039	Low
RAD1				
204460_s_at	0.155	No association	0.073	No association
204461_x_at	0.005	Low	0.099	No association
210216_x_at	0.073	No association	0.122	No association
RAD9A				
204828_at	0.094	No association	0.028	Low
RAD17 (RAD24)				
207405_s_at	0.148	No association	0.285	No association
211228_s_at	0.082	No association	0.047	High
210826_x_at	0.027	High	0.059	No association
RECQL (RECQ1)				
205091_x_at	0.056	No association	0.119	No association
210568_s_at	0.039	Low	0.273	No association
212917_x_at	0.188	No association	0.055	No association
212918_at	0.151	No association	0.052	No association
RECQL5				
211468_s_at	0.011	High	0.332	No association
221686_s_at	0.049	High	0.096	No association
210309_at	0.011	High	0.054	No association
34063_at	0.001	High	0.293	No association
RFC1				
208021_s_at	0.085	No association	0.121	No association
RIF1				
214700_x_at	0.119	No association	0.053	No association
214872_at	0.088	No association	0.028	Low
RPA4				
221143_at	0.017	High	0.079	Low
TOP3A				
204946_s_at	0.112	No association	0.355	No association
TOP3B				
213660_s_at	0.061	No association	0.008	Low
TOPBP1				
202633_at	0.079	No association	0.001	Low

Other	Test Dataset		TCGA Dataset	
	Log rank P value	**Expression level**	**Log rank P value**	**Expression level**
TP53				
211300_s_at	0.004	High	0.205	No association
201746_at	0.000348	Low	0.173	No association
TP53BP1 (53BP1) 203050_at	0.079	No association	0.025	Low
PARP4				
202239_at	0.065	No association	0.042	High
PTEN				
204053_x_at	0.053	No association	0.001	Low
204054_at	**0.004**	**Low**	**0.001**	**Low**
211711_s_at	**0.022**	**Low**	**0.002**	**Low**
217492_s_at	0.074	No association	0.000188	Low

Significant associations after BH FDR correction are shown in bold

Appendix 3: Clinicopathological associations with DNA repair gene expression in the Test dataset

APE1 expression			
Variable	Low N=72	High N=119	X^2 **Adjusted p value**
MGMT expression			
Low	21 (29.2%)	40 (33.6%)	0.523
High	51 (70.8%)	79 (66.4%)	
NBN expression			
Low	50 (69.4%)	83 (69.7%)	0.965
High	22 (30.6%)	36 (30.3%)	
PMS2 expression			
Low	33 (45.8%)	47 (39.5%)	0.390
High	39 (54.2%)	72 (60.5%)	
PTEN expression			
Low	38 (52.8%)	58 (48.7%)	0.589
High	34 (47.2%)	61 (51.3%)	

MGMT expression			
Variable	Low N=61	High N=130	X^2 **Adjusted p value**
APE1 expression			
Low	21 (34.4%)	51 (39.2%)	0.523
High	40 (65.6%)	79 (60.8%)	
NBN expression			
Low	35 (57.4%)	98 (75.4%)	**0.012**
High	26 (42.6%)	32 (24.6%)	
PMS2 expression			
Low	20 (32.8%)	60 (46.2%)	0.081
High	41 (67.2%)	70 (53.8%)	
PTEN expression			
Low	26 (42.6%)	70 (53.8%)	0.148
High	35 (57.4%)	60 (46.2%)	

NBN expression			
Variable	Low N=133	High N=58	X^2 **Adjusted p value**
APE1 expression			
Low	50 (37.6%)	22 (37.9%)	0.965
High	83 (62.4%)	36 (62.1%)	
MGMT			
Low	35 (26.3%)	26 (44.8%)	**0.012**
High	98 (73.7%)	32 (55.2%)	
PMS2 expression			
Low	52 (39.1%)	28 (48.3%)	0.237
High	81 (60.9%)	30 (51.7%)	
PTEN expression			
Low	70 (52.6%)	26 (44.8%)	0.321
High	63 (47.4%)	32 (55.2%)	

PMS2 expression			
Variable	Low N=80	High N=111	X^2 **Adjusted p value**
APE1 expression			
Low	33 (41.3%)	39 (35.1%)	0.390
High	47 (58.8%)	72 (64.9%)	
MGMT expression			
Low	20 (25.0%)	41 (36.9%)	0.081
High	60 (75.0%)	70 (63.1%)	
NBN expression			
Low	52 (65.0%)	81 (73.0%)	0.237
High	28 (36.0%)	30 (27.0%)	
PTEN expression			
Low	47 (58.8%)	49 (44.1%)	**0.046**
High	33 (41.3%)	62 (55.9%)	

	PTEN expression		
Variable	Low N=96	High N=95	X^2 **Adjusted p value**
APE1 expression			
Low	38 (39.6%)	34 (35.8%)	0.589
High	58 (60.4%)	61 (64.2%)	
MGMT expression			
Low	26 (27.1%)	35 (36.8%)	0.148
High	70 (72.9%)	60 (63.2%)	
NBN expression			
Low	70 (72.9%)	63 (66.3%)	0.321
High	26 (27.1%)	32 (33.7%)	
PMS2 expression			
Low	47 (49.0%)	33 (34.7%)	**0.046**
High	49 (51.0%)	62 (65.3%)	

Appendix 4: Clinicopathological associations with DNA repair gene expression in the TCGA dataset

	APE1 expression		
Variable	Low N=443	High N=65	X^2 **Adjusted p value**
	Demographic details		
Age			
<59 years	205 (46.3%)	37 (56.9%)	0.108
≥59 years	238 (53.7%)	28 (43.1%)	
Gender			
Male	267 (60.3%)	41 (63.1%)	0.665
Female	176 (39.7%)	24 (36.9%)	
Performance status			
0-1	236 (72.6%)	49 (87.5%)	**0.018**
2-4	89 (27.4%)	7 (12.5%)	
	Treatment given		
Radiotherapy			
No	126 (28.4%)	9 (13.8%)	**0.013**
Yes	317 (71.6%)	56 (86.2%)	
Chemotherapy			
No	147 (33.2%)	10 (15.4%)	**0.004**
Yes	296 (66.8%)	55 (84.6%)	
	DNA repair gene expression		
MGMT expression			
Low	99 (22.3%)	21 (32.3%)	0.077
High	344 (77.7%)	44 (67.7%)	
NBN expression			
Low	93 (21.0%)	15 (23.1%)	0.701
High	350 (79.0%)	50 (76.9%)	
PMS2 expression			
Low	331 (74.7%)	51 (78.5%)	0.514
High	112 (25.3%)	14 (21.5)	
PTEN expression			
Low	353 (79.7%)	46 (70.8%)	0.102
High	90 (20.3%)	19 (29.2%)	

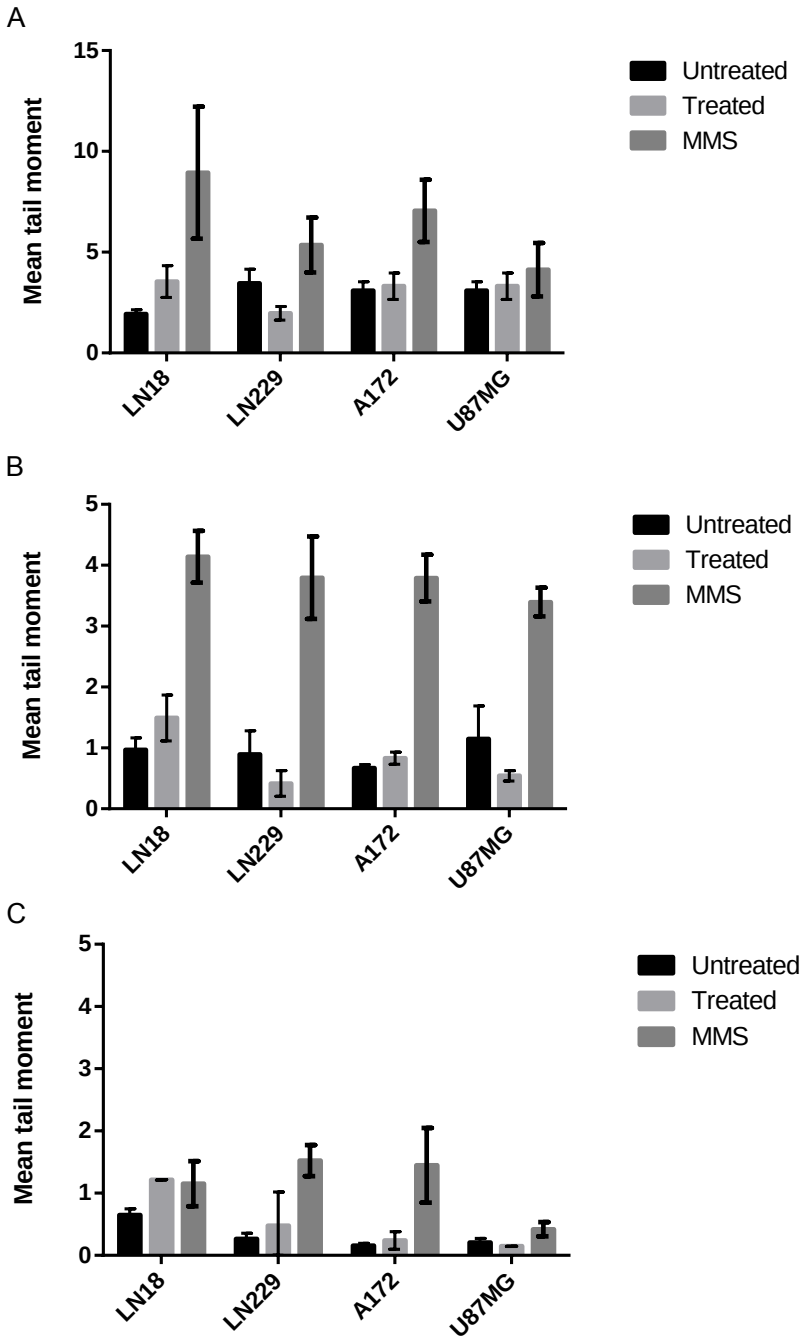
	MGMT expression		
Variable	Low N=120	High N=388	X^2 **Adjusted p value**
	Demographic details		
Age			
<59 years	52 (43.3%)	190 (49.0%)	0.280
≥59 years	68 (56.7%)	198 (51.0%)	
Gender			
Male	66 (55.0%)	242 (62.4%)	0.149
Female	54 (45.0%)	146 (37.6%)	
Performance status			
0-1	71 (71.7%)	214 (75.9%)	0.411
2-4	28 (28.3%)	68 (24.1%)	
	Treatment given		
Radiotherapy			
No	33 (27.5%)	102 (26.3%)	0.793
Yes	87 (72.5%)	286 (73.7%)	
Chemotherapy			
No	34 (28.3%)	123 (31.7%)	0.485
Yes	86 (71.7%)	265 (68.3%)	
	DNA repair gene expression		
APE1 expression			
Low	99 (82.5%)	344 (88.7%)	0.077
High	14 (17.5%)	44 (11.3%)	
NBN expression			
Low	14 (11.7%)	94 (24.2%)	**0.003**
High	106 (88.3%)	294 (75.8%)	
PMS2 expression			
Low	78 (65.0%)	304 (78.4%)	**0.003**
High	42 (35.0%)	84 (21.6%)	
PTEN expression			
Low	92 (76.7%)	307 (79.1%)	0.567
High	28 (23.3%)	81 (20.9%)	

	NBN expression		
Variable	Low N=108	High N=400	X^2 **Adjusted p value**
	Demographic details		
Age			
<59 years	48 (44.4%)	194 (48.5%)	0.454
≥59 years	60 (55.6%)	206 (51.5%)	
Gender			
Male	59 (54.6%)	249 (62.3%)	0.150
Female	49 (45.4%)	151 (37.8%)	
Performance status			
0-1	51 (71.8%)	234 (75.5%)	0.052
2-4	20 (28.3%)	76 (24.5%)	
	Treatment given		
Radiotherapy			
No	31 (28.7%)	104 (26.0%)	0.572
Yes	77 (71.3%)	296 (74.0%)	
Chemotherapy			
No	29 (26.9%)	128 (32.0%)	0.304
Yes	79 (73.1%)	272 (68.0%)	
	DNA repair gene expression		
APE1 expression			
Low	93 (86.1%)	350 (87.5%)	0.701
High	15 (13.9%)	50 (12.5%)	
MGMT expression			
Low	14 (13.0%)	106 (26.5%)	**0.003**
High	94 (87.0%)	294 (73.5%)	
PMS2 expression			
Low	95 (88.0%)	287 (71.8%)	**0.01**
High	13 (12.0%)	113 (28.2%)	
PTEN expression			
Low	74 (68.5%)	325 (81.2%)	**0.004**
High	34 (31.5%)	75 (18.8%)	

	PMS2 expression		
Variable	Low N=382	High N=126	X^2 **Adjusted p value**
	Demographic details		
Age			
<59 years	187 (49.0%)	55 (43.7%)	0.301
≥59 years	195 (51.0%)	71 (56.3%)	
Gender			
Male	237 (62.0%)	71 (56.3%)	0.257
Female	145 (38.0%)	55 (43.7%)	
Performance status			
0-1	213 (75.0%)	72 (74.2%)	0.880
2-4	71 (25.0%)	25 (25.8%)	
	Treatment given		
Radiotherapy			
No	98 (25.7%)	37 (29.4%)	0.414
Yes	284 (74.3%)	89 (70.6%)	
Chemotherapy			
No	115 (30.1%)	42 (33.3%)	0.496
Yes	267 (69.9%)	84 (66.7%)	
	DNA repair gene expression		
APE1 expression			
Low	331 (86.6%)	112 (88.9%)	0.514
High	51 (13.4%)	14 (11.1%)	
MGMT expression			
Low	78 (20.4%)	42 (33.3%)	**0.003**
High			
NBN expression			
Low	95 (24.9%)	13 (10.3%)	**0.001**
High			
PTEN expression			
Low	296 (77.5%)	103 (81.7%)	0.313
High	86 (22.5%)	23 (18.3%)	

	PTEN expression		
Variable	Low N=399	High N=109	X^2 **Adjusted p value**
	Demographic details		
Age			
<59 years	166 (41.6%)	76 (69.7%)	**0.000001**
≥59 years	233 (58.4%)	33 (30.3%)	
Gender			
Male	252 (63.2%)	56 (51.4%)	**0.026**
Female	147 (36.8%)	53 (48.6%)	
Performance status			
0-1	224 (73.4%)	61 (80.3%)	0.220
2-4	81 (26.6%)	15 (19.7%)	
	Treatment given		
Radiotherapy			
No	114 (28.6%)	21 (19.3%)	0.051
Yes	285 (71.4%)	88 (80.7%)	
Chemotherapy			
No	139 (32.3%)	28 (25.7%)	0.183
Yes	270 (67.7%)	81 (74.3%)	
	DNA repair gene expression		
APE1 expression			
Low	343 (88.5%)	90 (82.6%)	0.102
High	46 (11.5%)	19 (17.4%)	
MGMT expression			
Low	92 (23.1%)	28 (25.7%)	0.567
High	307 (76.9%)	81 (74.3%)	
NBN expression			
Low	74 (18.5%)	34 (31.2%)	**0.004**
High	325 (81.5%)	75 (68.8%)	
PMS2 expression			
Low	296 (74.2%)	86 (78.9%)	0.313
High	103 (25.8%)	23 (21.1%)	

Appendix 5. Determination of the optimal time point for the neutral comet assay.



Cells were exposed to 10µM KU55933 for (A) 6 hours, (B) 24 hours or (C) 48 hours prior to the assessment of DNA damage using the Neutral Comet assay. All four cell lines accumulate DNA damage (as assessed by the mean tail moment) in response to 100µM MMS, with the highest levels of damage seen in the LN18 cell line after 6 hours exposure to MMS. In the KU55933 treated samples the highest levels of DNA damage were seen at the 6 and 24 hour time points. Data presented as mean ± SD.

Appendix 6. List of the 84 DNA repair genes assessed on the RT^2 Profiler DNA repair PCR Array.

Base Excision Repair (BER)

APEX1, APEX2, CCNO, LIG3, MPG, MUTYH, NEIL1, NEIL2, NEIL3, NTHL1, OGG1, PARP1 (ADPRT1), PARP2, PARP3, POLB, SMUG1, TDG, UNG, XRCC1.

Nucleotide Excision Repair (NER)

ATXN3, BRIP1, CCNH, CDK7, DDB1, DDB2, ERCC1, ERCC2 (XPD), ERCC3 (XPB), ERCC4, ERCC5, ERCC6, ERCC8, LIG1, MMS19, PNKP, POLL, RAD23A, RAD23B, RPA1, RPA3, SLK, XAB2, XPA, XPC.

Mismatch Repair (MMR)

MLH1, MLH3, MSH2, MSH3, MSH4, MSH5, MSH6, PMS1, PMS2, POLD3, TREX1.

Double-Strand Break (DSB) Repair

BRCA1, BRCA2, DMC1, FEN1, LIG4, MRE11A, PRKDC, RAD21, RAD50, RAD51, RAD51C, RAD51B (RAD51L1), RAD51D (RAD51L3), RAD52, RAD54L, XRCC2, XRCC3, XRCC4, XRCC5, XRCC6 (G22P1).

Other DNA Repair Genes

ATM, ATR, EXO1, MGMT (AGT), RAD18, RFC1, TOP3A, TOP3B, XRCC6BP1 (KUB3).

Appendix 7. DNA repair genes down-regulated in the LN18 cell line as compared to the U87MG cell line.

Gene Symbol	Fold Regulation	Pathway
ATR	-2.2235	DDR
ERCC1	-2.2778	NER
ERCC5	-2.0137	NER
LIG4	-3.463	DSB repair
NEIL2	-2.0547	BER

DDR=DNA damage response.

Appendix 8. DNA repair genes up-regulated in the LN18 cell line as compared to the U87MG cell line.

Gene Symbol	Fold Regulation	Pathway
CCNO	2.4674	BER
MGMT	1186.6372*	Direct repair
MMS19	2.0405	NER
MPG	3.3185	BER
MRE11A	2.3116	DSB repair
MSH4	2.8092#	MMR
MSH5	4.2476	MMR
NEIL1	22.6796*	BER
NTHL1	4.6015	BER
PMS2	2.8687	MMR
POLB	2.5958	BER
POLL	2.4225	NER
RAD51C	2.6043	DSB repair
TDG	2.2819	BER
TOP3A	2.0497	DSB repair
TREX1	2.2489	MMR
UNG	2.7471	BER
XRCC6BP1	3.0372	DSB repair

The fold change for each gene (and cell line) is shown as compared to the LN18 cell line; the DNA repair pathway to which the gene belongs is also shown. #This gene's CT value is relatively high (>30) meaning that its expression level is low, in both control and test samples. The fold change result may therefore be subject to greater variation. *This gene's average CT value is relatively high (>30) in one sample and reasonably low in the other sample (<30). This means that the gene's expression is relatively low in one sample and reasonably detected in the other. This suggests that the actual fold-change is at least as high as reported. BER=base excision repair; NER=nucleotide excision repair; DSB repair=double strand break repair; MMR=mismatch repair.

Appendix 9. DNA repair genes up-regulated in the LN18 cell line as compared to the A172 cell line.

Gene Symbol	Fold Regulation	Pathway
ATM	3.4377	DDR
ATXN3	2.9691	NER
CCNO	2.0907	BER
ERCC2	2.852	NER
ERCC6	2.0699	NER
LIG1	2.7311	NER
MGMT	724.2468*	Direct repair
MLH3	2.4534	MMR
MMS19	2.1261	NER
MPG	3.3571	BER
MRE11A	3.5364	DSB repair
MSH4	3.7505#	MMR
MSH5	3.3482	MMR
MUTYH	3.9074	BER
NEIL1	9.951*	BER
NEIL3	4.3034	BER
OGG1	2.3428	BER
PARP2	3.0072	BER
POLL	2.7642	NER
PRKDC	2.6106	DSB repair
RAD51B	2.6415	DSB repair
RAD51C	2.2586	DSB repair
RAD52	2.4752	DSB repair
RAD54L	7.6479	DSB repair
RPA3	2.6655	NER
SMUG1	5.3683	BER
TOP3A	3.234	DSB repair
TOP3B	2.3704	DSB repair
TREX1	3.4319	MMR
UNG	3.2228	BER
XRCC2	2.9291	DSB repair
XRCC3	5.0714#	DSB repair
XRCC6	3.0493	DSB repair
XRCC6BP1	4.4856	DSB repair

The fold change for each gene (and cell line) is shown as compared to the LN18 cell line; the DNA repair pathway to which the gene belongs is also shown. #This gene's CT value is relatively high (>30) meaning that its expression level is low, in both control and test samples. The fold change result may therefore be subject to greater variation. *This gene's average CT value is relatively high (>30) in one sample and reasonably low in the other sample (<30). This means that the gene's expression is relatively low in one sample and reasonably detected in the other. This suggests that the actual fold-change is at least as high as reported. DDR=DNA damage response; BER=base excision repair; NER=nucleotide excision repair; DSB repair=double strand break repair; MMR=mismatch repair.

Appendix 10. DNA repair genes down-regulated in the LN18 cell line as compared to the A172 cell line.

Gene Symbol	Fold Regulation	Pathway
DDB2	-2.5584	NER
LIG4	-3.6074	DSB repair

Appendix 11. DNA repair genes up-regulated in the LN18 cell line as compared to the LN229 cell line.

Gene Symbol	Fold Regulation	Pathway
ATM	6.7938	DDR
ATXN3	2.8705	NER
BRCA2	2.4733	DSB repair
CCNO	2.2768	BER
CDK7	2.0115	NER
ERCC2	3.4674	NER
ERCC8	2.288	NER
MGMT	1056.7478*	Direct repair
MLH3	31.5604*	MMR
MMS19	3.9205	NER
MPG	8.2866	BER
MRE11A	5.3602	DSB repair
MSH3	2.5351	MMR
MSH4	3.717#	MMR
MSH5	2.8489	MMR
MUTYH	3.0357	BER
NEIL1	6.4924*	BER
NEIL2	2.0802	BER
NEIL3	2.0306	BER
NTHL1	2.6282	BER
OGG1	2.4773	BER
PARP3	2.5731	BER
POLB	3.8038	BER
POLL	6.2276*	NER
RAD23A	2.6587	NER
RAD50	3.2074	DSB repair
RAD51B	3.2919	DSB repair
RAD51C	2.5482	DSB repair
RAD52	2.3152	DSB repair
RAD54L	2.1248	DSB repair
RFC1	3.2917	Other
RPA1	2.4215	NER
SLK	2.4743	NER
TOP3A	2.8934	DSB repair
TREX1	4.2843	MMR
UNG	4.258	BER
XAB2	2.9061	NER

Gene Symbol	Fold Regulation	Pathway
XPC	2.2254*	NER
XRCC3	2.4388#	DSB repair
XRCC4	2.6749	DSB repair
XRCC6BP1	2.3554	DSB repair

The fold change for each gene (and cell line) is shown as compared to the LN18 cell line; the DNA repair pathway to which the gene belongs is also shown. #This gene's CT value is relatively high (>30) meaning that its expression level is low, in both control and test samples. The fold change result may therefore be subject to greater variation. *This gene's average CT value is relatively high (>30) in one sample and reasonably low in the other sample (<30). This means that the gene's expression is relatively low in one sample and reasonably detected in the other. This suggests that the actual fold-change is at least as high as reported. BER=base excision repair; NER=nucleotide excision repair; DSB repair=double strand break repair; MMR=mismatch repair.

Appendix 12. DNA repair genes down-regulated in the LN18 cell line as compared to the LN229 cell line.

Gene Symbol	Fold Regulation	Pathway
DDB2	-3.3392	NER
EXO1	-2.3323	MMR