

**The role of G-quadruplex DNA in
the transcriptional
reprogramming of drug-resistant
ovarian cancer cells**

By

Gem Flint

A thesis presented for the degree of Doctor of
Philosophy

Imperial College London

Department of Chemistry

October 2025

Abstract

Acquired drug resistance remains a major challenge in the treatment of late-stage ovarian cancer. There are many molecular mechanisms that contribute to the acquisition of resistance, that enable the cancer cells to reduce drug efficacy and evade apoptosis. Of note, epigenetic adaptation plays a crucial role in the development of drug resistance, through mediating changes in the expression of genes implicated in signalling pathways associated with resistance. One emerging epigenetic regulator that has not yet been considered in the context of ovarian cancer drug resistance is the non-canonical DNA structure called G-quadruplexes (G4s). These structures are implicated in transcriptional regulation, either proximally through the enrichment of G4s at oncogene promoters, or distally through the enrichment of G4s at the boundaries of chromatin loops, bringing genomic regulatory elements into proximity. Therefore, it was hypothesised that G4 formation could facilitate epigenetic adaptation, and promote drug resistance, in ovarian cancer.

To investigate this, this project employed a multi-layered epigenomics approach in both drug-sensitive and drug-resistant paired ovarian cancer cell lines. The aim was to evaluate the impact of G4 formation on gene expression in the drug-resistant cell lines, relative to changes in chromatin accessibility which is an established regulatory mechanism. It was observed that G4s forming at introns and intergenic regions had a greatly significant transcriptomic impact, compared to a gain in chromatin accessibility at these sites. However, G4s gained in all regions of the drug-resistant cell lines were enriched in pathways associated with acquired drug resistance. Perturbation of the G4s resulted in the significant downregulation of pathways associated with drug resistance and resulted in a partial re-sensitisation to cisplatin. This suggests that G4s are implicated with the epigenetic regulation of drug resistance in ovarian cancer and may represent a valuable therapeutic target to overcome therapy resistance.

Table of Contents

List of figures	9
List of tables	12
Statement of originality	13
Copyright Declaration	13
List of publications	14
Acknowledgements	15
1. Introduction	16
1.1. High Grade Serous Ovarian Carcinoma	16
1.1.1. HGSOC treatment	17
1.1.2. Drug resistance in HGSOC.....	18
1.1.2.1. Drug efflux and influx	20
1.1.2.2. DNA damage repair	20
1.1.2.3. Cancer stem cells and epithelial to mesenchymal transition	21
1.1.2.4. Autophagy and metabolism	21
1.2. Epigenetic regulation in HGSOC	23
1.2.1. DNA methylation and drug resistance	23
1.2.2. Histone modifications and chromatin accessibility	24
1.2.2.1. Histone modifications and chromatin accessibility in HGSOC drug resistance	

1.3. G-quadruplexes.....	28
1.3.1. Cellular detection of G4s	30
1.3.1.1. Computational prediction.....	30
1.3.1.2. <i>In vitro</i> mapping.....	31
1.3.1.3. Endogenous G4 mapping	33
1.3.2. Biological relevance of G4s	38
1.3.2.1. Telomere maintenance and DNA replication	39
1.3.2.2. Phase separation.....	40
1.3.2.3. Transcriptional regulation and chromatin architecture	41
1.3.3. G4s in cancer	48
1.3.3.1. Telomeric G4s and cancer	49
1.3.3.2. G4s and genomic instability in cancer.....	50
1.3.3.3. G4s and acquired drug resistance in cancer	53
1.3.3.4. G4s and epigenetic regulation in cancer	54
1.4. Project rationale and aims.....	55
2. The optimisation of genomic G4 mapping in HGSOC cells	58
2.1. Chapter overview	58
2.2. PEO1 and PEO4 paired ovarian cancer cell lines.....	58
2.3. BG4 generation and validation	61
2.3.1. BG4 expression and purification	61
2.3.2. BG4 selectivity validation	63

2.4. Chromatin immunoprecipitation and quantitative polymerase chain reaction (ChIP-qPCR)	65
2.4.1. ChIP-qPCR in U2OS cells	66
2.4.2. Optimisation of ChIP-qPCR in PEO1 and PEO4 cells.....	68
2.4.3. Successful generation of PEO1 and PEO4 ChIP-qPCR samples	75
2.5. Bioinformatic analysis of G4 ChIP-seq data	77
2.5.1. Assessment of raw data quality	78
2.5.2. Read alignment	80
2.5.3. Post alignment read sorting and filtering.....	81
2.5.4. Sample clustering.....	82
2.5.5. Peak calling	84
2.5.6. Genomic distribution of G4 peaks.....	89
2.5.7. Pathway enrichment of G4 distribution	90
2.6. Chapter summary and conclusions	92
3. Intersection with transcriptomic and accessibility data.....	94
3.1. Chapter overview	94
3.2. The transcriptional profiles of PEO1 and PEO4 cells.....	95
3.2.1. Assessment of RNA integrity.....	95
3.2.2. Analysis of RNA-seq data	96
3.2.3. Differential expression.....	98
3.3. Profiling of chromatin accessibility.....	103

3.3.1.	Assessment of ATAC profiles.....	103
3.3.2.	Analysis of ATAC-seq data.....	104
3.4.	Integration of G4 ChIP, ATAC and RNA-seq data.....	109
3.4.1.	Global gain and loss of G4s and ATAC in PEO4 cells.....	110
3.4.2.	Changes in gene expression at different genomic loci	111
3.4.3.	G4-related effect not observed with peak loss	116
3.5.	G4s transcriptional association is independent of BRCA status	119
3.6.	Pathway enrichment of G4 associated genes.....	122
3.6.1.	Global pathway enrichment	122
3.6.2.	Regional pathway enrichment.....	125
3.7.	Chapter summary and conclusions.....	129
4.	Perturbation of G4s transcriptional impact	131
4.1.	Chapter overview	131
4.1.1.	G4 ligand dose optimisation	131
4.1.2.	Transcriptional effect of targeting G4s over time	132
4.2.	Pathway enrichment associated with G4 targeting.....	137
4.2.1.	Global pathway enrichment of significantly downregulated genes	137
4.2.2.	Regional pathway enrichment in PEO4.....	140
4.3.	Re-sensitisation of PEO4 to cisplatin.....	143
4.4.	Chapter summary and conclusions.....	146
5.	Validation in an alternative paired HGSOC cell line.....	147

5.1. Chapter overview	147
5.2. PEA1 and PEA2 paired ovarian cancer cell lines	147
5.3. Generation of genomic G4 maps in PEA1 and PEA2	148
5.3.1. Evaluation of the transcriptional effect of a gain or loss of G4 in PEA2	155
5.3.2. Pathway enrichment of gained PEA2 G4s	160
5.4. Comparison of acquired G4s in drug-resistant ovarian cancer cells	162
5.4.1. Acquired PEA2 G4s are associated with a greater transcriptional upregulation 162	
5.4.2. Conservation of G4s between PEA1/2 and PEO1/4	164
5.4.3. Pathway enrichment of conserved PEO4 and PEA2 G4s	167
5.5. Chapter summary and conclusions	169
6. Conclusions and future outlook	172
7. Materials and methods	182
7.1. Cell Culture	182
7.1.1. General	182
7.1.2. MTS Cytotoxicity assay	182
7.2. BG4 Expression and Validation	183
7.2.1. BG4 expression	183
7.2.2. BG4 purification and quantification	184
7.2.3. BG4 quantification	185
7.2.4. BG4 validation using Enzyme-Linked Immunosorbent Assay (ELISA)	185

7.3. Chromatin Immunoprecipitation and Quantitative Polymerase Chain Reaction (ChIP-qPCR)	186
7.3.1. Chromatin fixation and sonication	187
7.3.2. Chromatin quality assessment	187
7.3.3. BG4 immunoprecipitation.....	188
7.3.4. Quantitative PCR and enrichment analysis.....	189
7.4. RNA-sequencing.....	191
7.4.1. RNA extraction	191
7.4.2. RNA quality assessment and sequencing.....	192
7.5. ATAC-sequencing	192
7.5.1. Profiling chromatin accessibility	192
7.5.2. Quality assessment and sequencing	195
7.6. Bioinformatics.....	195
7.6.1. ChIP-seq.....	195
7.6.2. ATAC-seq.....	196
7.6.3. RNA-seq.....	196
7.6.4. Data integration	197
7.7. Statistical tests	197

List of figures

Figure 1.1 Mechanisms of acquired drug resistance in HGSOC.	19
Figure 1.2 Epigenetic mechanisms used in transcriptional regulation in HGSOC.	24
Figure 1.3 ATAC-seq protocol.....	26
Figure 1.4 G4 structure and localisation.	29
Figure 1.5 G4-Seq methodology.....	32
Figure 1.6 Methods of mapping G4s within cells.....	36
Figure 1.7 Biological roles of G4s.	39
Figure 1.8 G4s transcriptional role in promoters and looping.....	43
Figure 1.9 G4s transcriptional role in R-loops and DNA methylation.....	45
Figure 1.10 G4 stabilisation and implications in cancer cells.....	50
Figure 2.1 PEO1 and PEO4 cells differential sensitivity to cisplatin.....	59
Figure 2.2 <i>BRCA2</i> mutational status of PEO1 and PEO4.....	60
Figure 2.3 BG4 expression, purification and concentration determination.....	63
Figure 2.4 BG4 selectivity validation.	65
Figure 2.5 Results of ChIP-qPCR of U2OS cells.....	68
Figure 2.6 Results of ChIP-qPCR of PEO1 cells.....	70
Figure 2.7 Unsuccessful ChIP-qPCR in PEO1 and PEO4 cells.....	72
Figure 2.8 Fragment size distribution analysis.....	74
Figure 2.9 Results of ChIP-qPCR for PEO1 and PEO4 cells.	76
Figure 2.10 Quality assessment of ChIP reads.	79

Figure 2.11 Correlation plot of ChIP aligned reads.....	83
Figure 2.12 IGV genomic tracks and G4 peaks in PEO1 and PEO4 cells.....	85
Figure 2.13 G4 peaks in PEO1 and PEO4 cells.....	87
Figure 2.14 Validation of G4 maps in PEO1 and PEO4 cells.	88
Figure 2.15 Genomic location distribution of PEO1 and PEO4 G4 peaks.	90
Figure 2.16 Pathway enrichment of G4-related genes in PEO4 cells.	91
Figure 3.1 RNA integrity validation.....	96
Figure 3.2 Bioinformatic analysis pipeline undertaken for NGS data.	97
Figure 3.3 RNA-seq replicate reproducibility and differential expression.	99
Figure 3.4 Differential gene expression pathway enrichment in PEO4.....	101
Figure 3.5 ATAC-seq sample validation.....	104
Figure 3.6 ATAC-seq data correlation, peaks and genomic distribution.	109
Figure 3.7 Differential gene expression of genes that gain or lose a G4.....	111
Figure 3.8 Differential expression of genes that gain a G4 at different locations.	113
Figure 3.9 Validation of G4 maps with alternative peak calling parameters.	116
Figure 3.10 Differential expression of genes that lose a G4 in PEO4.....	118
Figure 3.11 Differential expression of genes associated with a gain in G4, detected by ChIP and CUT&Tag, in two PEO models.....	121
Figure 3.12 Pathway enrichment of upregulated genes genes linked with a gain in G4s.	123
Figure 3.13 Pathway enrichment of downregulated genes linked with a loss in G4.....	124
Figure 4.1 Cytotoxicity of PDS and PhenDC-3 in PEO1 and PEO4.	132

Figure 4.2 PCA plot of PEO1 and PEO4 RNA-seq with G4 perturbation.....	133
Figure 4.3 Differential expression of genes linked with PEO1 G4s after G4 perturbation.	135
Figure 4.4 Differential expression of genes linked with PEO4 G4s after G4 perturbation.	136
Figure 4.5 Pathway enrichment of downregulated genes in PEO1 after G4 perturbation.	138
Figure 4.6 Partial re-sensitisation of PEO4 to cisplatin.....	144
Figure 5.1 Establishment of PEO1/4 and PEA1/2 cell line pairs.	148
Figure 5.2 Chromatin preparation ChIP-qPCR of PEA1 and PEA2 cells.	150
Figure 5.3 G4 peaks and validation in PEA1 and PEA2 cells.	154
Figure 5.4 Genomic location distribution of PEA1 and PEA2 G4 peaks.	155
Figure 5.5 Differential expression of genes associated with a G4 gain in PEA2 cells....	157
Figure 5.6 Differential expression of genes associated with a G4 loss in PEA2 cells.....	159
Figure 5.7 Pathway enrichment of genes linked with non-promoter PEA2 G4s.....	161
Figure 5.8 Differential expression of genes associated with G4 gain in drug resistance.	163
Figure 5.9 Correlation, peak overlap and genomic distribution of G4s.....	167

List of tables

Table 2.1 Table of three biological repeats of ChIP-qPCR PEO1 and PEO4 cells.....	77
Table 2.2 Alignment quality of ChIP samples.	81
Table 3.1 Top 10 most differentially expressed genes.	100
Table 3.2 GSEA enrichment of differentially expressed genes in PEO4.	102
Table 3.3 Alignment quality of ATAC samples.....	105
Table 3.4 GSEA enrichment of genes linked with a gained non-promoter G4.....	126
Table 3.5 GSEA enrichment of genes associated with promoter G4s.....	127
Table 4.1 GSEA enrichment of downregulated genes in PEO4 after G4 perturbation..	139
Table 4.2 GSEA enrichment of downregulated genes linked with promoter G4s after G4 perturbation.....	141
Table 4.3 GSEA enrichment of genes associated with non-promoter G4s after G4 perturbation.....	142
Table 5.1 Alignment quality of ChIP samples.	151
Table 5.2 GSEA enrichment of conserved G4s in PEO4 and PEA2.....	168
Table 7.1 Cytotoxicity drug and concentrations.....	183
Table 7.2 Primer sequences used in ChIP-qPCR.....	190
Table 7.3 Primer sequences for ATAC-seq.....	195

Statement of originality

I, Gem Flint, declare that the work presented in this thesis is original work, which I have performed and written up myself. In the instances where experiments were performed in collaboration with others, the experiment and the individual has been recognised directly within the text. All external sources have been referenced and acknowledged within the text and listed in the bibliography.

Copyright Declaration

The copyright of this thesis rests with the author. Unless otherwise indicated, its contents are licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International Licence (CC BY-NC).

Under this licence, you may copy and redistribute the material in any medium or format. You may also create and distribute modified versions of the work. This is on the condition that: you credit the author and do not use it, or any derivative works, for a commercial purpose.

When reusing or sharing this work, ensure you make the licence terms clear to others by naming the licence and linking to the licence text. Where a work has been adapted, you should indicate that the work has been changed and describe those changes.

Please seek permission from the copyright holder for uses of this work that are not included in this licence or permitted under UK Copyright Law.

List of publications

- J Robinson⁺, **G Flint**⁺, I Garner, S Galli, TE. Maher, MK Kuimova, R Vilar, IA McNeish, R Brown, H Keun, M Di Antonio* *G-quadruplex structures regulate long-range transcriptional reprogramming to promote drug resistance in ovarian cancer cells*. Genome Biology, 2025- DOI: 10.1186/s13059-025-03654-y
- S Galli*, **G Flint**, L Růžicková, and M Di Antonio* *Genome-wide mapping of G-quadruplex DNA: a step-by-step guide to select the most effective method*. RSC Chemical Biology, 2024- DOI: 10.1039/D4CB00023D
- L Monti*, G Firth, JRC Faria, JM Kelly, **G Flint**, S Galli, TE. Maher, and M Di Antonio* *Genome-wide mapping of DNA G-quadruplexes in Trypanosoma brucei chromatin reveals enrichment in coding regions*. In review, 2025, bioRxiv- DOI: 10.1101/2025.07.22.666098
- SP Nuccio, E Cadoni, R Nikoloudaki, S Galli, AJ Ler, C Sanchez-Cabanillas, TE Maher, E Fan, D Guneri, **G Flint**, M Zhu, LS Liu, CR Fullenkamp, Z Waller, L Magnani, JS Schneekloth, M Di Antonio* *Chemically modified CRISPR-Cas9 enables individual G-quadruplex and i-motif targeting, revealing ligand dependent transcriptional perturbation*. Accepted. 2025, Nature Communications

+ = both authors contributed equally to this work

Acknowledgements

The work presented here has been the result of four of the best and most rewarding years at Imperial College, which could not have been achieved without the support and guidance of some incredible scientists and people.

I would like to start by acknowledging my supervisors, Dr Marco Di Antonio and Prof Hector Keun, for their incredible mentorship throughout the project. I have truly learnt from the best, and without them I would not be the researcher I am today- forever challenging the dogma! It has been an honour to be part of the Di Antonio group, and I cannot wait to see all the incredible science that they will most certainly produce over the coming years.

Special thanks go to the members of the Di Antonio group who shaped my genomics journey: Jenna Robinson, Silvia Galli, Claudia Sanchez Cabanillas and Ludovica Monti. It was an uphill battle at times, but I am incredibly proud of what we have achieved through hard work and a lot of laughs. Of note, Jenna Robinson for training me in bioinformatics and dealing with all of my initial error messages with a smile on her face. Thank you also to the scientists who contributed directly to this work; Jenna Robinson for the collaboration on genomics protocols; Thomas Maher for the G4 ligands and just being the best friend I've gained from this PhD; Silvia Galli who helped antibody expression at times and Chelsie Li for helping me rush through ChIP samples as time was running out!

I must thank the EPSRC/Chemical Biology CDT and CRUK for funding this work, and to the amazing people this training programme introduced me too along the way. It has been an invaluable opportunity for which I am ever grateful.

Finally, a massive thank you goes to my wonderful friends and family who have been the greatest support throughout my PhD. They truly are my rocks, and they made the hard times so much easier. In particular, my Mum, who goes above and beyond for me every day and is my biggest support in life.

1. Introduction

1.1. High Grade Serous Ovarian Carcinoma

High grade serous ovarian carcinoma (HGSOC) is the most common and aggressive form of epithelial ovarian cancer (EOC), accounting for approximately 75% of all ovarian cancer cases.(1) It is classified at Grade 3, and arises from the serous membrane in the epithelial layer.(2) This means that the carcinoma is considered to be at the advanced stage at diagnosis, with poorly differentiated and rapidly proliferating cells.(3) The majority of EOC cases are diagnosed at this stage, due to the absence of symptoms presenting at the early stage.(2) When symptoms do occur, they include bloating, abdominal pain, and increased urinary frequency.(2) Additionally, the lack of early stage screening tools hinders detection at the early stages of disease.(4) Due to late detection, the disease has often progressed significantly, meaning the 10-year survival rate for patients diagnosed at an advanced stage is 15%, compared to 55% for those diagnosed at an early stage.(2)

EOC is very heterogenous, with multiple genetic and epigenetic alterations occurring in oncogenes and tumour suppressor genes.(5) This causes dysregulation of signalling pathways that contribute to disease.(5) However, HGSOC is characterised by a high frequency of *TP53* mutation in up to 96% of cases.(6) This is a gene with key tumour suppressor properties, such as cell cycle control, promoting deoxyribonucleic acid (DNA) damage repair (DDR), and initiation of cell death to prevent uncontrolled growth.(6) Due to the near ubiquitous nature of *TP53* mutation, it is believed this event may be one of the mutations associated with disease initiation and carcinogenesis.(7) This near universal *TP53* mutation is often accompanied by additional mutations, such as in *BRCA1/BRCA2* genes in 10-20% of HGSOC cases.(2) These mutations confer a high genetic risk factor for EOC, whether they are germline or somatic mutations. Both BRCA1 and BRCA2 proteins are involved in the repair of double strand DNA breaks, via the homologous recombination pathway.(8) Common mutations in the *BRCA* genes include insertions,

deletions or substitutions that cause a frameshift, or introduction of a nonsense mutation, that results in the truncation of the BRCA proteins and a loss in protein activity.(9) Hence, loss of BRCA1 or BRCA2 functionality contributes to genomic instability, by failing to repair a cells damaged double strand DNA breaks, and these mutated cells continue to divide, and accumulate more mutations, due to the *TP53* mutation and its role in cell cycle control. The overall result of combining the altered function of these two tumour suppressor genes, is an accumulation of genetic changes and genomic instability that is crucial for tumour development and uncontrolled cell division.(10)

1.1.1. HGSOC treatment

Typically, the gold standard treatment of HGSOC is cytoreductive surgery followed by adjuvant chemotherapy.(7) This is the preferable course of treatment if optimal tumour removal is possible, with little to no residual disease being left behind. The adjuvant chemotherapy is then administered to eliminate any remaining cancer cells, to help prevent disease relapse.(11) In some cases, neo adjuvant chemotherapy is administered prior to surgery, to shrink the tumour before removal, to make the surgery less invasive and improve patient outcome.(12) Regardless of administration order, platinum- and taxane-based chemotherapeutic agents are a standard component of HGSOC treatment. These include cisplatin, which acts by binding to DNA to form inter- and intra-strand crosslinks. These platinum adducts cause DNA damage, that in turn block transcription and cell division, and ultimately induces apoptosis in rapidly proliferating cancer cells.(13) Whilst an effective therapy, studies have highlighted the limitations of first-line chemotherapy due to the potentially harmful side effects and toxicity.(14) As such, alternative therapies that exploit tumour specific mutations, such as in *BRCA1/2* genes, have also been developed as a targeted therapy for HGSOC patients that harbour those mutations.

Of note, Poly(ADP-ribose) polymerase (PARP) inhibitors have been developed to exploit the DNA repair defect of *BRCA*-mutant cells. PARP ordinarily acts as part of the base excision repair process, to repair single strand DNA breaks.(15) In this case, PARP1 acts

as a sensor to detect single strand breaks, before undergoing an auto-post translational modification to create chains of PAR that recruit proteins and DDR enzymes to the site of DNA damage.(16) Inhibitors of PARP act by binding to the PARP enzymes active site block the synthesis of PAR chains, or through a PARP trapping mechanism, where the PARP enzyme is trapped to the site of DNA damage leading to stalled replication forks.(17)(18) PARP inhibitors operate on the principle of synthetic lethality, which is the concept that the targeting and disruption of two genes is sufficient to cause cell death, whilst the disruption of just of those genes is not.(19) In this instance, the *BRCA*-mutation that some HGSOC cells have is not sufficient to cause cell death, but rather advantageous for the cancer cell survival. However, when treating these cells with a PARP inhibitor, which prevents single strand DNA break repair, single strand DNA breaks accumulate and are converted to double strand breaks.(20) This causes irreparable DNA damage and overwhelming genomic instability of which the cancer cell cannot survive, and apoptosis is triggered, due to the inability of the cell to repair the accumulating double strand breaks through homologous recombination.(20) This strategy makes the treatment of cancer cells more selective, as it exploits the genetic vulnerabilities that are present in cancer cells and not in the healthy cells, such as *BRCA* mutation.

1.1.2. Drug resistance in HGSOC

Typically, patients initially respond well to treatment. However, around 75% of patients will acquire some form of drug resistance and develop recurrent disease.(21)(22) Whilst therapies such as PARP inhibitors and other targeted therapies have prolonged progression-free survival, resistance to both chemotherapy and other HGSOC drugs is a common occurrence and a major cause of fatality.(23) Coupled with the late detection due to lack of symptoms and poor screening, acquired drug resistance strongly contributes to the poor prognosis of HGSOC patients, and presents a considerable challenge in the treatment of HGSOC.

Resistance can be intrinsic or acquired, with the former arising from pre-existing genetic factors within HGSOC that make the treatment less effective from the start of the

treatment regime.(24) Resistance that develops after exposure to treatment, through an accumulation of genetic mutations or epigenetic modifications to reduce drug efficacy, is termed acquired resistance.(25) In the instance that a patient initially responds well to treatment, before becoming less responsive and developing resistance, this would be an acquired form. Acquired resistance can develop via many molecular mechanisms, which are multifactorial and can develop dependently and independently of each other.(21) These include increased drug efflux, decreased drug influx, increased DDR, increased cancer cell stemness, increased epithelial to mesenchymal transition (EMT), decreased apoptosis, altered autophagy, altered metabolism, and epigenetic modification (**Figure 1.1**). In essence, the distinct molecular mechanisms all aim to reduce programmed cell death, apoptosis, of the cancer cell by reducing the efficacy of the therapy.

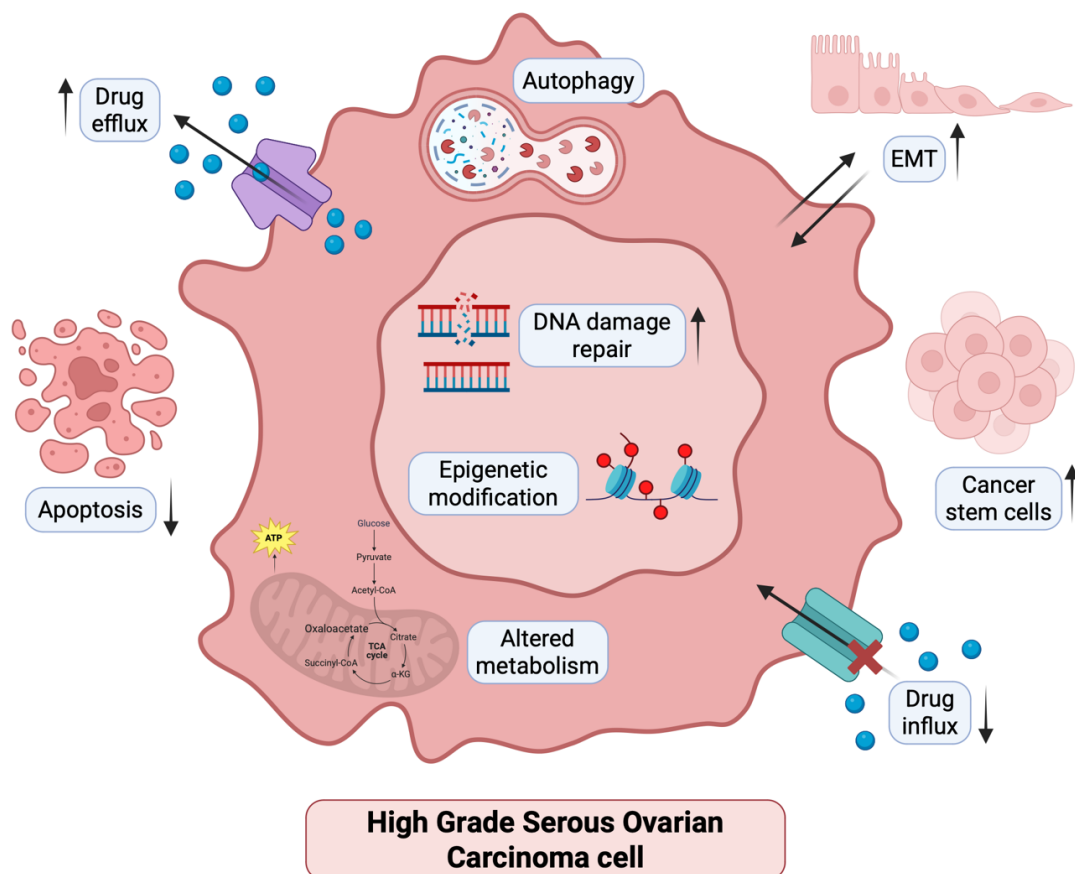


Figure 1.1 Mechanisms of acquired drug resistance in HGSOC.

These include increased drug efflux, decreased drug influx, increased cancer cell stemness and EMT, altered metabolism and autophagy, all contributing to decreased apoptosis to promote cancer cell survival.

1.1.2.1. Drug efflux and influx

Increased drug efflux and decreased influx act to reduce the intracellular concentration of the drug to an insufficient level to exhibit the desired therapeutic effect, and aid the development of resistance.(24) This abnormal transmembrane transport is mediated by overexpression of membrane transporter proteins that enable drug efflux, and reduced expression of transporters that aid drug influx. Most commonly, the adenosine triphosphate-binding cassette (ABC) superfamily that utilises adenosine triphosphate (ATP) to facilitate drug efflux has been shown to be overexpressed in HGSOC, specifically those carcinomas treated with cisplatin.(24) Alternatively, solute carrier (SLC) superfamily transporters have been shown to transport cisplatin, and its analogs, into the cell and has a low level of expression in platinum resistant ovarian cancer.(22)

1.1.2.2. DNA damage repair

Whilst aberrant DDR remains crucial for genomic instability and cancer cell survival, such as through *BRCA* mutation, an accumulation of DNA damage which is ineffectively repaired leads to apoptosis and cancer cell death.(26) Many HGSOC treatments, such as chemotherapeutic and PARP inhibitor drugs, act to induce DNA damage and promote this apoptotic effect. As a result, increased DDR helps to combat this therapeutic effect and maintain cancer cell viability. This increased DDR proficiency can occur via many mechanisms, including reinstating efficient homologous recombination, through reversion mutations in *BRCA*-mutant cancers.(27)(28)(29) This enables effective repair of double strand DNA breaks, limiting the accumulation of DNA damage and thus reducing drug efficacy. Furthermore, overexpression and elevated activation of cell cycle checkpoints also reduces DNA damage accumulation through promotion of DDR and also contributes to drug resistance in HGSOC.(30)

1.1.2.3. Cancer stem cells and epithelial to mesenchymal transition

Sub-populations of tumour cells that have stem cell like characteristics are termed cancer stem cells (CSC). These arise through EMT, which is a process where epithelial tumour cells lose their epithelial defining characteristics and gain mesenchymal traits that allow cell migration and invasion of other tissues through metastasis.(31) This process gives the cancer cells stem cell like properties, as they change their inherent traits, such as decreasing cell-cell adhesion and apical-based polarity, in order to progress and migrate.(31) EMT is thus associated with the acquirement of stem-like features, and both EMT and CSC have been associated with acquired drug resistance in HGSOC. CSCs have an inherent tendency to resist apoptosis, which is a key marker of drug resistance, and the CSC model for drug resistance proposes that it is the repopulation of drug sensitive HGSOC cells by resistant CSC cells that is a driver of therapy resistance.(32)(33) Cancer cells that are resistant to chemotherapy agents have also been shown to have acquired mesenchymal cell characteristics, which suggests EMT may be involved in the development of resistance.(34)(35) Additionally, CSC and mesenchymal cells tend to have a low cycling rate in comparison to rapidly proliferating cancer cells, which are the key target of platinum therapy.(36) EMT and CSC stemness is highly regulated by multiple signalling pathway to maintain a stem like state. These signalling pathways include Wnt/ β -catenin, TGF- β and TNF- α /NF- κ B signalling pathways, which promote cancer cell migration and invasion through regulation of EMT associated genes and CSC properties.(37)(38)(39) Aberrant activation and upregulation of these pathways has thus been associated with acquired drug resistance, and targeting of these pathways to reduce signalling has been shown to increase sensitivity of resistant cells to chemotherapy agents.(40)(41)

1.1.2.4. Autophagy and metabolism

Autophagy is the process by which cytoplasmic components are sequestered into vesicles called autophagosomes, and degraded upon fusion with lysosomal components.(42) This enables the cell to maintain the cellular biosynthesis and energy demand of the cell during metabolic stress.(43) Cisplatin treatment has been shown to activate an autophagic response in ovarian cancer cells, as a survival signal that aims to counteract

drug induced apoptosis.(42) This is termed cytoprotective autophagy, and inhibition of this form can aid cellular response to cisplatin.(44) Notably, autophagy has been reported for its close inter-play with the DNA damage response. It can impact DDR promoted by DNA damage during radiotherapy, and also regulates normal DDR pathways through degradation of deubiquitinating enzymes that impact the recruitment of BRCA1 and RAD51 to sites of double strand breaks.(45) Moreover, mammalian target of rapamycin (mTOR) is a serine/threonine protein kinase made of two complexes; mTORC1 and mTORC2.(46) The complex acts as a negative regulator of autophagy, and is tightly regulated through interaction with other signalling pathways, such as PI3K/AKT.(47) Upregulation of mTOR signalling results in inhibition of autophagy, whilst drug treatment generally promotes autophagy in EOC. mTOR is also a signalling pathway that is highly dysregulated in EOC with acquired drug resistance. Cisplatin resistant tumour cells often have constitutive activation of mTORC1 signalling, that also has roles in translation initiation and cellular senescence.(48)(46) Inhibition of mTOR signalling has been shown to promote cisplatin-induced apoptosis through suppression of this signalling pathway.(49) Moreover, activated mTOR signalling is associated with increased FANCD2 expression, which is a DNA interstrand crosslink repair protein.(50) Knockdown of FANCD2 in some platinum resistant HGSOC cell lines has resulted in increased sensitivity to platinum-therapy, and inhibiting the mTOR pathway has been shown to deplete FANCD2. This suggests that mTOR signalling could mediate platinum response and resistance via FANCD2 regulation and DDR, more so than its involvement with autophagy.(50)

These various molecular mechanisms all contribute to acquired drug resistance in HGSOC. What underpins them all, is an alteration in gene expression to up or downregulate the signalling pathways, that in turn influences drug transport, CSC, EMT, DDR, and autophagy. This signalling dysregulation allows HGSOC cells to evade apoptosis and adapt to survive the cytotoxic effects of drugs used to treat HGSOC.

1.2. Epigenetic regulation in HGSOC

One hallmark of cancer cells is the ability to disrupt these critical signalling pathways to maintain a resistant state, through epigenetic regulation.(51) Epigenetic regulation modulates how genes are expressed, without the alteration of the DNA sequence itself.(52) Epigenetic modifications act to regulate the genome, without relying on genetic changes, and can occur in response to environmental factors, such as drug treatment. In turn, epigenetic modifications acquired in response to therapy can influence how cancer cells respond to treatment.(53) Common epigenetic modifications include DNA methylation, histone modifications, and chromatin accessibility changes. Moreover, epigenetic modifications are dynamic and reversible, making them an attractive target for therapeutic intervention.(53)

1.2.1. DNA methylation and drug resistance

DNA methylation involves the transfer of a methyl group to the C5 position of cytosine to form 5-methyl-cytosine (5mC) and is catalysed by DNA methyltransferases (DNMT). This modification acts to regulate gene expression by recruiting proteins that either promote transcription or prevent the binding of transcription factors to DNA sites.(54) The majority of 5mC occurs at cytosines that forego a guanine base at CpG sites.(55) When there is an increased level of DNA methylation, primarily at CpG sites in gene promoter regions, this acts to repress or silence expression of that gene (**Figure 1.2 A**).(54) In the context of EOC, hypermethylation of the *BRCA1* promoter has been reported in 5-24% of cases, and is associated with increased sensitivity to cisplatin relative to unmethylated *BRCA1*.(56)(57) Conversely, hypermethylation of the *hMLH1* and *hMSH2* genes, that are involved in mismatch repair, can confer resistance to platinum therapy.(58)(59) However, less than 1% of hypermethylated genes are linked to therapy resistance induction.(21) This could mean that it is the hypermethylation of one or few key genes that is instrumental for the acquisition of drug resistance. Moreover, it has been proposed that DNA methylation could be used as an epigenetic prognostic biomarker of response to therapy and be profiled to this end.(33)

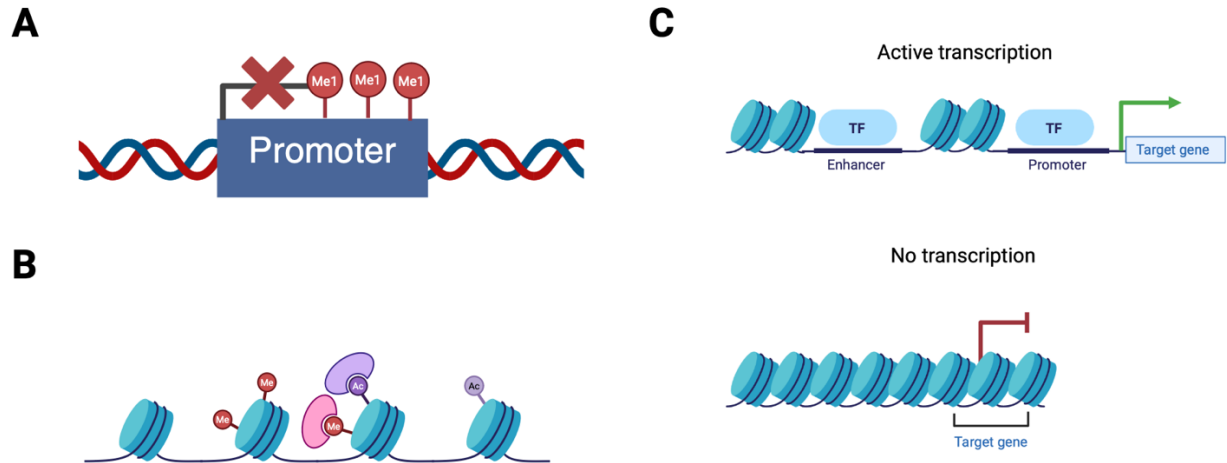


Figure 1.2 Epigenetic mechanisms used in transcriptional regulation in HGSOc.

A) Hypermethylation of promoter regions resulting in transcriptional repression, **B)** histone modifications such as methylation and acetylation with proteins binding to these modifications and **C)** changes in chromatin accessibility to repress or promote active gene expression due to accessibility of the underlying DNA to transcriptional machinery.

1.2.2. Histone modifications and chromatin accessibility

Within the nucleus, DNA coils around histone proteins to form nucleosomes. Histone modifications are the post-translational modification of the amino N-terminal histone tail that protrudes from the nucleosome.⁽⁶⁰⁾ These modifications regulate the chromatin structure through impacting nucleosomal interactions, but also through the recruitment of chromatin remodelling proteins that can re-position the nucleosomal arrangement.⁽⁶⁰⁾ These modifications can include acetylation, methylation, phosphorylation, and ubiquitination, and can be permissive or repressive depending on the mark properties and locations (**Figure 1.2 B**). Permissive marks are histone modifications that promote gene expression and maintain open, transcriptionally active chromatin, and include acetylation of lysine residues of a histone tail.⁽⁶¹⁾ This allows access of the transcriptional machinery to the underlying DNA, and in turn elevates gene expression. Repressive marks are associated with a condensed, closed chromatin state,

which is not transcriptionally active (**Figure 1.2 C**).⁽⁶¹⁾ Hence, there is a close interplay between histone modifications, chromatin accessibility, and transcriptional regulation. These modifications are dynamic and can switch from a permissive to a repressive state (and vice versa), in response to external stimuli such as drug treatment.

Chromatin accessibility can be mapped using Assay for Transposase Accessible Chromatin with high-throughput sequencing (ATAC-seq), developed by Buenrostro *et al.*⁽⁶²⁾ This methodology aims to identify open regions of chromatin, which may be associated with active gene expression. Changes in chromatin accessibility between disease states or cell lines may be indicative of epigenetic differences that are key for cell specific regulation.⁽⁶³⁾ This mapping approach does not probe for histone markers for a permissive chromatin state, as there are many different histone markers conferring this, and it would require an extensive profiling protocol. Instead, it utilises a hyperactive Tn5 transposase, that inserts sequencing adapters into accessible chromatin regions.⁽⁶²⁾ These fragments can then be sequenced through next generation sequencing (NGS), with read pile up indicating the presence of an accessible chromatin site (

Figure 1.3). ATAC-seq has become a core method for probing chromatin accessibility, and to this end epigenetic regulation, including in HGSOC.

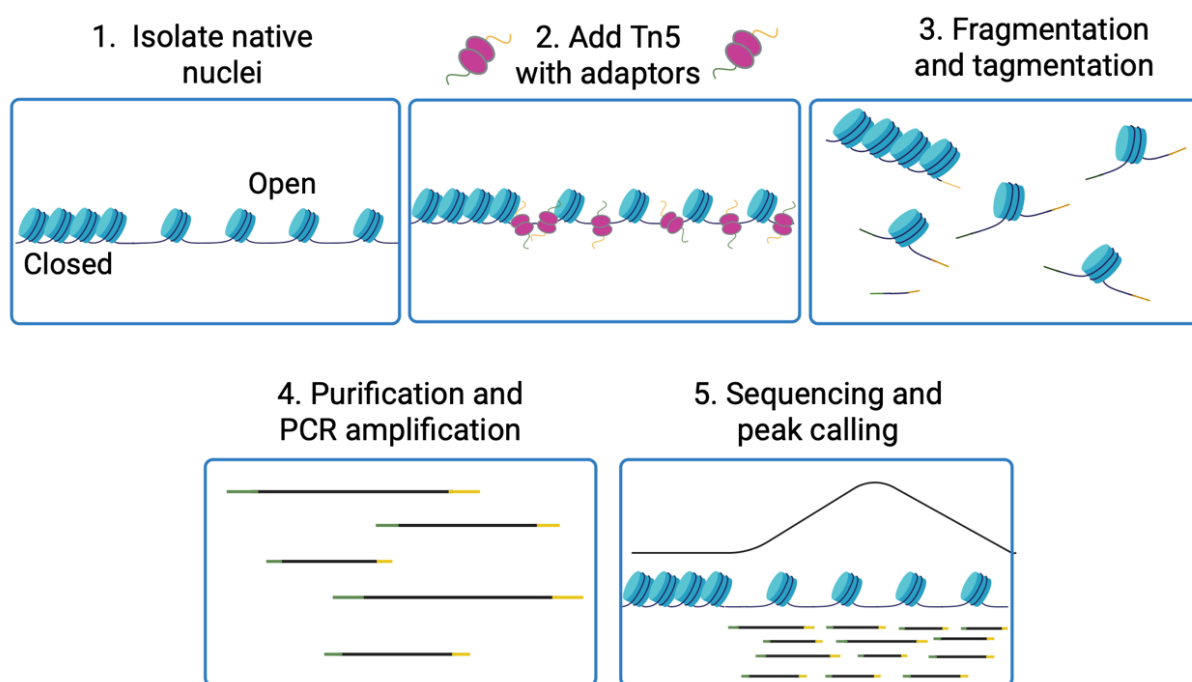


Figure 1.3 ATAC-seq protocol.

1) Native nuclei isolation with open and closed chromatin regions in their native state, **2)** addition of overactive Tn5 attached to sequencing adaptors that binds to open chromatin regions, **3)** fragmentation of chromatin and insertion of sequencing adaptors at open chromatin regions, **4)** purification of DNA fragments and PCR amplification prior to **5)** NGS and peak calling to identify read pile up in open chromatin sites, called accessible peaks.

1.2.2.1. Histone modifications and chromatin accessibility in HGSOC drug resistance

EOC drug resistance has previously been associated with an upregulation of histone deacetylase (HDAC) activity.(64)(65) This is the enzyme responsible for removing permissive histone marks from lysine residues, hence acting to compact the chromatin and downregulating gene expression.(66) Additionally, overexpression of histone lysine N-methyltransferases EHMT1 and EHMT2 has been associated with PARP inhibitor resistance.(67) These enzymes can help activate or repress gene expression through the transfer of a methyl group to lysine residues at position 4 (permissive) or 9 (repressive) of histone protein 3.(68)(69) HGSOC cells have also been shown to contain bivalent

promoters, which is when both permissive and repressive marks are present at the same gene promoter.(70) These promoters are in a bivalent state, and poised for either activation or repression depending on environmental factors, such as chemotherapy exposure. Indeed, H3K4me3 and H3K27me3 bivalent marks have been identified in pathways of high relevance to HGSOC, such as PI3K and TGF- β signalling.(71) These gene sets are significantly differentially expressed in acquired drug-resistant HGSOC cell lines, relative to chemotherapy sensitive cells, suggesting the importance of bivalent promoters at epigenetically regulating gene expression levels in response to drug exposure, to drive resistance.(71)

Furthermore, histone marks can inform of active enhancer sites, which are regulatory elements implicated in HGSOC drug resistance. Enhancers are non-coding regions of the genome that help activate target gene expression. They are mostly found at intronic and intergenic regions of the genome, and can be located as far as a million base pairs away from the target gene.(72)(73) They are comprised of many transcription factor binding sites, which bind transcription factors and chromatin modellers to aid chromatin looping of the enhancer complex, to make direct contact with promoter sites.(72)(74) Clusters of enhancers located close to together in the genome are termed super-enhancers, which have an even stronger enrichment for transcriptional coactivator binding.(75)(76) The acetylation of lysine residue 27 of histone 3 (H3K27ac) distinguishes active enhancer sites from inactive enhancers.(77) Moreover, this mark is associated with active promoters in mammalian cells, and the genomic locations vary significantly based on cell type and disease state.(77) These distal enhancer sites have been reported to be of importance to maintain transcriptional reprogramming associated with acquired drug resistance in ovarian cancer.(78)(79) Of note, super-enhancer clusters were significantly upregulated in HGSOC resistant cells, and targeting of these super-enhancers resulted in significantly decreased expression of the target genes. Ultimately, this resulted in an increase in sensitivity of resistant HGSOC cells to cisplatin.(78) Additionally, HDAC inhibitors have been shown to alter enhancer reprogramming that contributed to resistance of patient-derived ovarian cancer models to MEK inhibitors.(79) This highlights the importance of non-promoter regions, that can be identified by histone

modifications, in the epigenetic regulation of gene expression, leading to acquired drug resistance.

Chromatin accessibility changes have also been linked to acquired HGSOC drug resistance. Altered chromatin accessibility, primarily at intergenic regions, is associated with altered gene expression of genes enriched in key drug resistance pathways such as the DNA damage response.(80) Interestingly, drug-resistant HGSOC cells lines clustered together when considering their ATAC-seq profiles, suggesting common epigenetic chromatin changes associated with a drug-resistant state. Changes in accessibility particularly occurred at regions containing cis-regulatory elements. Hence, it is proposed that epigenetic changes at these regulatory sites, like through alteration in chromatin accessibility, can drive gene expression changes that contribute to acquired drug resistance.(80) Single-cell chromatin accessibility landscapes have also shown an enrichment of chromatin accessible sites of resistance-pathway-activating transcription factor binding motifs.(81) In particular, at binding sites of regulators involved with CSC and Wnt signalling. This is a trend also observed in other cancers, such as chemotherapy resistant breast cancer, which also associates increased chromatin accessibility with significantly upregulated genes involved with acquired drug resistance.(82) This corroborates the idea that changes in chromatin accessibility are associated with altered gene expression of pathways that contribute to drug resistance. In addition, non-promoter sites are again highlighted as being of importance in this context.

1.3. G-quadruplexes

Whilst chemical modification of DNA and histones are established epigenetic markers, increasing evidence suggests that structural changes of DNA can also elicit epigenetic regulation. In this context, DNA G-quadruplexes (G4s) are emerging as an epigenetic regulator, and have not yet been considered in the context of drug-resistant HGSOC.(83) These are non-canonical secondary structures that form in guanine-rich (G-rich) DNA regions.(84) Four guanine bases Hoogsteen hydrogen bond with each other to form a planar G-tetrad, with two or more tetrads stacking on top of each other through π - π

interactions to form the quadruplex structure (**Figure 1.4 A**).⁽⁸⁵⁾ G4s are stabilised by monovalent cations within the central cavity, and are sandwiched between the tetrads. Potassium ions preferentially stabilise the G4 structure, owing to the larger size in comparison to sodium and then lithium ions, which further decrease in stabilisation capacity.⁽⁸⁶⁾ G4s are predominantly located in open chromatin regions, and have more recently been associated with transcriptional regulation (**Figure 1.4 B**).⁽⁸⁷⁾ These factors advocate that G4s could be considered as epigenetic regulators, owing to the change in DNA structure and not the underlying DNA sequence, whilst regulating gene expression.

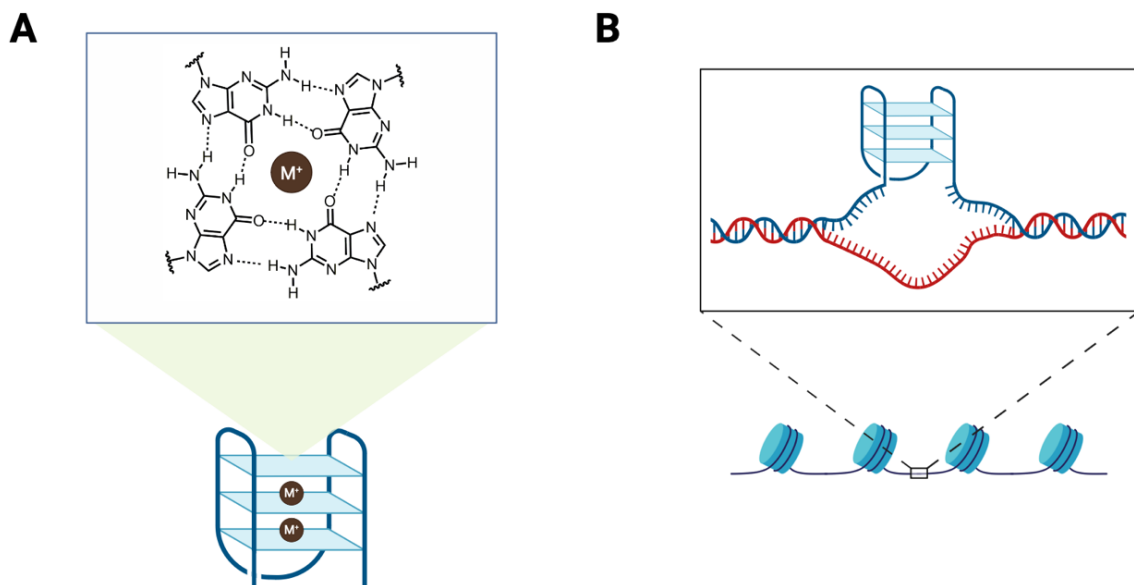


Figure 1.4 G4 structure and localisation.

A) The structure of G4s formed from four guanine bases Hoogsteen hydrogen bonding with each other into a G-tetrad that stack via π - π interactions, stabilised by a monovalent cation in the cavity of the tetrad. **B)** Schematic depicting the formation of G4s in open chromatin regions, where G4s are enriched. Figure adapted from (88).

1.3.1. Cellular detection of G4s

1.3.1.1. Computational prediction

The biological relevance of G4s was overlooked until the early 2000's, when the crystal structure of a human G4 sequence, extracted from telomeric DNA, was reported.(89) It was then thought that if such a structure could form *in vitro*, then maybe it could also form endogenously within cells. Initially, computational prediction methods were developed to identify putative G4-forming sequences within the human genome.(90) These methods included Quadparser, that uses a canonical G4-folding motif $d(G_3+N_{1-7}G_3+N_{1-7}G_3+)$, where 3 or more guanine bases (G) are interspaced with 1-7 alternative bases (N), which form the loops interspacing the G-tetrads.(91) Since then, several more sophisticated algorithms have been developed to reduce false negatives associated with Quadparser. This stemmed from the discovery of G4-forming sequences that were composed of only two G-tetrads, with long looping regions over 7 nucleotides long, and G4 structures containing bulge like regions.(92) This highlighted that G4 formation is not restricted to a canonical motif, and as such other methods were required to predict G4 distribution across the genome. These include G4Hunter, an algorithm to assess G-richness and G-skewness of a sequence, to give a likelihood of G4 formation.(93) Using this prediction, regions proposed to form stable G4s were identified and validated *in vitro*, using biophysical assays to confirm G4 formation under physiological conditions.(93) Indeed, these developed algorithms are able to predict non-canonical G4 structures that a simple strict motif search will enable, and hence identify more predicted G4s. However, this benefit comes with an increased false positive rate, which is an additional issue with the relying on the development of G4 predictive algorithms alone. Whilst these methods were successful in providing the first insight into the distribution of putative G4-forming sequences in the human genome, they failed to account for the cellular chromatin landscape which would likely influence G4 formation. As such, they were not accurate or comprehensive enough to probe the true biological relevance of G4 formation. Methods to experimentally map G4s in the human genome needed to be developed to this end.

1.3.1.2. *In vitro* mapping

The first experimental whole-genome map of G4s was first generated in 2015, using an *in vitro* method termed G4-Seq.(94) The technique exploited the DNA polymerase stalling ability of G4s, that in turn reduces the quality of NGS, which can then be leveraged to identify where in the genome G4s can form. In essence, isolated genomic DNA is sequenced in conditions that do not favour G4 formation, such as in lithium containing buffers. Then, the same genomic sequence is sequenced under G4 stabilising conditions, such as a potassium containing buffer or in the presence of a G4 stabilising ligand, such as pyridostatin (PDS), which acts to hold the G4 in formation (**Figure 1.5**). (94) At regions in which a G4 forms, there will be a significant drop in the sequencing quality in the G4-promoting conditions, relative to the conditions that do not favour G4 formation. These regions of differential sequencing quality were mapped back to the genome, to identify where in the human genome can form a G4 *in vitro*. Through this experimental G4 mapping, >700,000 sequences were identified to have the potential to fold into a G4, termed observed quadruplexes (OQS). (94) This value was much greater than that predicted using computational methods alone, which predicted around 400,000, further highlighting the need for experimental map generation as opposed to relying solely on computational prediction using primary sequences or G4 richness. (91)(93) Moreover, an enrichment of G4s was observed in promoter and 5' untranslated regions (5'UTR). This made it the first experimental method to detect G4 enrichment at key regulatory sites, such as promoters, in the human genome. An advantage of this method was that it could be adapted to generate OQS maps in many organisms, owing to the requirement of simply needing isolated genomic DNA from whichever organism was desired to be mapped. (95) Indeed, OQS maps were generated in 12 different species, which also showed G4 enrichment at promoter and 5'UTR sites, suggesting a regulatory role of G4s conserved across different species. (95)

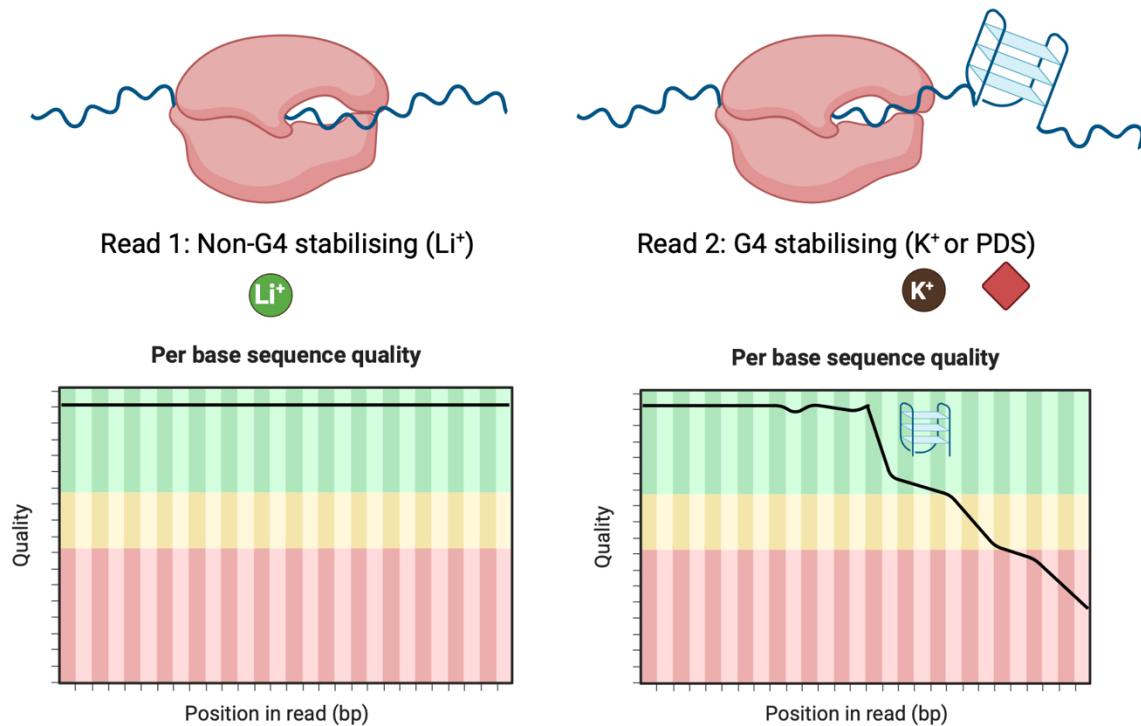


Figure 1.5 G4-Seq methodology.

Schematic of the two read methodology implemented by G4-Seq. Read 1 is sequenced in the presence of non-G4 stabilising conditions such as in lithium buffer, in which the G4 is unfolded, and the DNA polymerase moves freely along the DNA sequence, resulting in high quality NGS reads. Read 2 is the same genomic sequence, sequenced in the presence of G4 stabilising conditions such as a potassium buffer or with PDS. In these conditions, the G4 will be stabilised and result in DNA polymerase stalling which is registered as a decrease in sequencing quality. Where this quality decrease occurs signifies the site of G4 formation.

Whilst the development of G4-Seq was an important advancement in G4 mapping, relative to previous computational predictions, it had its limitations. Firstly, it could not be utilised to assess differences in G4 distribution across different cell types, or disease states within the same species. This is because it utilised isolated genomic DNA, which will be highly conserved between cell types of the same species, and as such the OQS profiles would be near identical. Moreover, as it is an *in vitro* mapping method, it still fails to account for the nature that DNA exists within cells; in chromatin form. The 3D structure

in which chromatin exists in the cellular context can vary significantly between different cell types.(74) As such, accessibility of chromatin differs greatly, and it was thought that this could alter G4 distribution within different types. To this end, endogenous mapping of G4s needed to be established to decipher how chromatin organisation impacted G4 formation and help uncover the true biological functions of G4s.

1.3.1.3. Endogenous G4 mapping

To generate the first endogenous G4 maps in the chromatin context, an antibody to selectively target G4 structures over other nucleic acid structures needed to be developed. This would allow direct targeting of G4 structures, which could then be extracted and sequenced to provide a map of where G4s form within human chromatin and cells. The most widely used antibody for this purpose, and for the visualisation of G4s in mammalian cells via immunofluorescence, is the single-chain variable fragment antibody; BG4.(96) BG4 was developed from phage display, using a library containing 2.3×10^{10} different clones. It was shown to have high selectivity for a broad range of G4 structures, binding with a low nanomolar affinity, as opposed to negligible binding to other nucleic acid structures, such as single stranded and double stranded DNA.(96) BG4 was the first antibody utilised to visualise DNA G4s in human cells, via immunofluorescence. Immunofluorescence studies demonstrated the dynamic nature of G4 formation, which is cell cycle dependent, further suggesting that G4s have a biological function within cells.(96)

Owing to the selective nature of BG4, it was then employed to generate the first endogenous G4 map via chromatin immunoprecipitation (ChIP).(97) This methodology was adapted from standard ChIP protocols, that are commonly used to map protein-DNA interactions or histone modifications.(98) Chromatin is first cross-linked by formaldehyde fixation to preserve the chromatin architecture, and the folded G4s, before sonication to produce DNA fragments in which G4s are then immunoprecipitated using BG4 (**Figure 1.6 A**).(99) The immunoprecipitated fragments are extracted using magnetic beads coated in an anti-FLAG antibody, targeting the FLAG tag present on

BG4.(96) These fragments are then subjected to NGS, to determine the sites of G4 formation. G4 ChIP-seq has been implemented to generate G4 maps in many cell lines, such as cancerous cell lines and human embryonic stem cells.(97)(99)(100) The number of G4s detected varies greatly depending on the cell line G4 ChIP-seq is performed on, ranging from 1,000 to 20,000 G4s detected.(87) This is much fewer than the number of OQS identified from G4-Seq, highlighting the importance of the cellular chromatin environment in regulating G4 folding within cells. However, of those G4 peaks detected using G4 ChIP-seq, >80% overlapped with OQS sites.(99) This suggests that there is a proportion of G4s that form within cells that cannot be identified using *in vitro* detection alone, further corroborating the idea that endogenous G4 mapping needs to be performed to gain a true understanding of G4 distribution. It is worth noting that the vast differences in G4 peak numbers detected between cells could be due to technical variation and differences in data quality between the individual studies reporting them. Therefore, it is difficult to distinguish the true differences in G4 peak numbers between cell lines unless a truly controlled study is performed to reduce inter-sample technical variation.

The generation of endogenous G4 maps via G4 ChIP-seq was instrumental in expanding the understanding of the biological role that G4s may play within the cells. It was noted that cancerous cells generally had a greater total number of G4s, relative to non-cancerous cells. Moreover, G4s were found to be enriched at oncogene promoters, which are typically highly expressed in cancerous cells.(97)(101) This suggested that G4s may have oncogenic potential, and be involved in transcriptional regulation at oncogenic sites.(102) It was after the generation of G4 ChIP-seq maps that G4s propensity to form in open chromatin regions was also discovered. Through intersection with ATAC-seq maps generated in the same cell lines, it was found that 98% of G4 peaks resided in accessible chromatin regions.(97) This is expected, owing to the bulkier structure that a G4 resides, and so a more open environment would be required in order to form. However, it is not expected that all accessible sites in the genome will contain a G4. This is because there are many more open chromatin regions as opposed to G4 sites, and chromatin opening is not reliant on a sequence specific motif to occur unlike G4s. Indeed, the mapping of accessible sites is also not confounded by G4 formation, as is exemplified by the high identified overlap of G4 and ATAC peaks. This is encouraging that the mapping

of accessibility is not impacted by non-canonical DNA structure and is robust enough to compare sites of accessibility and G4 formation. This further solidified the idea that G4 formation is tightly regulated by the chromatin architecture of a cell. Subsequently, it is proposed that G4 location between cell types will differ owing to the different chromatin landscapes that different cells adopt to maintain their distinct transcriptional profiles. Whilst a specific sequence is required for the formation of a G4, it is highly governed by the accessibility of the chromatin in order to form the alternative secondary structure. This is similar to other epigenetic modifications, by which they are not expected to form at the same location in every cells genome and instead also differ according to cell type.

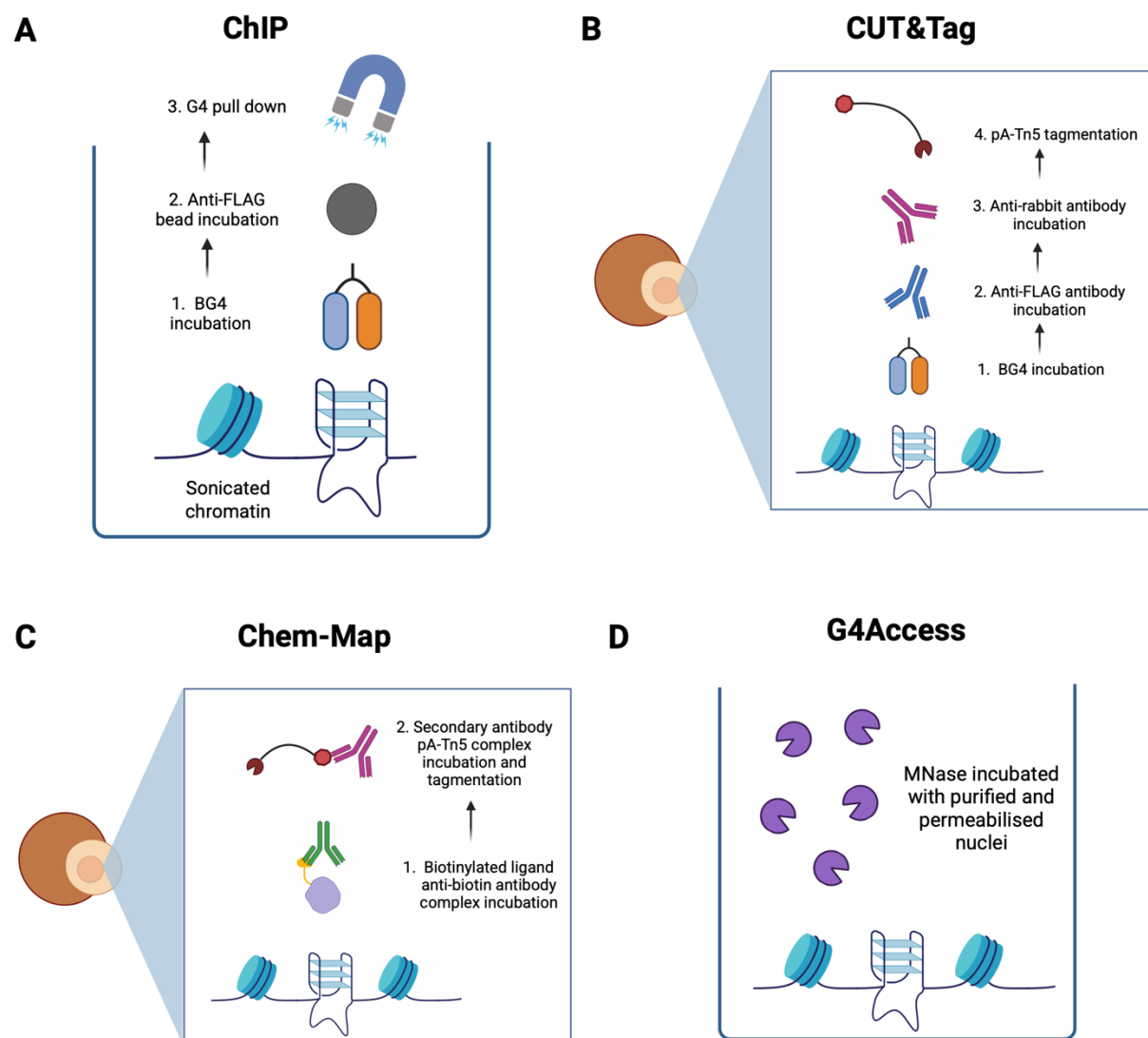


Figure 1.6 Methods of mapping G4s within cells.

A) ChIP-seq: chromatin is cross-linked prior to sonication to produce DNA fragments.

G4s are then immunoprecipitated using BG4 and extracted using anti-FLAG tag magnetic beads which bind the FLAG tag on BG4. **B) CUT&Tag:** BG4 binds G4 structures in fixed or native cells and nuclei, BG4 is then tagged with a transposase enzyme (Tn5), via various antibody binding steps, that cleaves the DNA under the BG4 binding and adds adapter sequences for NGS. **C) Chem-Map:** G4 targeting ligands are biotinylated and bind G4 structures in cells which are tagged with an anti-biotin antibody complexed with Tn5 that cleaves the DNA at the G4 site and inserts adapters for NGS. **D) G4Access:** A non-probe based approach that relies on MNase that typically cleaves DNA in open chromatin regions, yet is protected from cleavage when a G4 is present when low levels of MNase are present.

Since the development of G4 ChIP-seq, alternative G4 mapping methods have emerged to address some of the limitations associated with ChIP techniques. G4 ChIP data can have a low signal-to-noise ratio, and requires a relatively high sequencing depth of 30 million reads per sample, to achieve enough depth to confidently assign G4 peaks.(103)(99) To address this issue, BG4 has also been used to map the genomic locations of G4s using cleavage under target and tagmentation (CUT&Tag) (**Figure 1.6 B**).(104)(105) This method employs Tn5, which binds to BG4 through a series of antibody binding steps to enable signal amplification. In turn, the Tn5 cleaves the DNA at the site of antibody binding and G4 formation, inserting sequencing adaptors at the site of cleavage. These cleaved sites are then preferentially amplified during polymerase chain reaction (PCR). This generates a greater signal-to-noise ratio than in G4 ChIP-seq, as the untagmented fragments will not be amplified, as no adapter sequences will be inserted at these sites. This enables a lower sequencing depth of 10 million reads per sample to be used, as sufficient signal of G4s is seen over the background at a much lower depth.(106) Another advantage of CUT&Tag is the ability to perform G4 genomic mapping in native cells or nuclei, which is more similar to the physiological chromatin state of a cell as opposed to fixed and extracted chromatin.(106) It has, however, been proposed that untargeted CUT&Tag enriches at accessible chromatin sites thus decreasing the validity of maps generated with G4 CUT&Tag.(107) Nonetheless, negative controls using IgG antibodies

or using no targeting antibody show very minimal enrichment at open chromatin sites or known G4 sites, contradicting this claim. Moreover, given that G4s reside predominantly in open chromatin regions, and Tn5 cleaves in these regions, this would be expected that G4s identified by CUT&Tag enrich in these regions, and the notion that G4 CUT&Tag is not a viable approach to map G4s is hotly contested.

Additional G4 mapping methods have also been developed, that are not antibody-based approaches. These include Chem-Map, which follows an approach reminiscent of G4 CUT&Tag, but instead uses a biotinylated G4 targeting ligand in place of BG4 (**Figure 1.6 C**).⁽¹⁰⁸⁾ This approach generated robust G4 maps, which had a high overlap with detected G4s by CUT&Tag using BG4, indicating that the detected G4s are independent of probe-based method used.⁽¹⁰⁸⁾ A probe-free G4 mapping approach has also been developed, to remove the need for highly selective G4 probes with a strong binding affinity like required in ChIP, CUT&Tag and Chem-Map.⁽¹⁰⁹⁾ G4Access utilises a low concentration of the micrococcal nuclease, MNase, which cleaves DNA at open accessible sites (**Figure 1.6 D**). When G4s form in open chromatin regions, G4Access proposes that the DNA is then protected from MNase cleavage.⁽¹⁰⁹⁾ Whilst beneficial to have a probe-free method to map G4s that removes probe bias, this method requires *in vitro* validation of sequences. As such around 80% of the genomic G4 sequences identified through G4Access had the propensity for fold into a G4 *in vitro*.⁽¹⁰⁹⁾ This mapping method has its own quite drastic limitations. The mapping is purely based on the enzymatic activity of MNase, which preferentially cleaves at A- T-rich sites upstream of G-stretches. As such, cleavage will occur at these sites purely based on genomic sequence, and not due to presence of a G4 structure, thus requiring *in vitro* validation of sequences, and which may not actually form within a cell.

Multiple mapping methods have been developed to map endogenous G4s. They span probe based to non-probe-based approaches, and have all be used to develop G4 maps in a variety of cell states, types, and organisms, with various advantages and disadvantages to each mapping method. The generation of G4 mapping techniques was instrumental in

developing the understanding of the biological roles of G4s in cellular biology and disease, and G4 distribution, and association with transcription, could now be probed.

When conducting comparisons of the epigenome between different cell lines, it is important to ensure sufficient normalisation of the datasets is implemented to ensure a fair comparison. This includes comparisons of datasets such as G4 maps, accessibility peaks, and RNA. This can be done in the peak calling stage of data processing, such as when using MACS2, which normalises the control input file to the sample file generated to ensure an equal sequencing depth between the two is compared. This means that if the sample file has vastly more reads than the control that it won't be called as a false positive peak, whilst if the control file has many more reads that it won't be a false negative peak. This method of peak normalisation helps ensure fair comparison between the epigenomes of different samples and cell lines. Alternatively, spike-ins are another routine method for normalisation between samples. This is the process of adding a known amount of external DNA to the sample, which can then be used in the bioinformatics processing stage to normalise and scale samples to each other, knowing they all have received the same amount of spike-in control. Whilst an effective method for sample normalisation, instances where background and total signal for the G4 or ATAC samples remains constant, normalisation at the peak calling stage is sufficient.(110)

1.3.2. Biological relevance of G4s

G4-mapping was instrumental in the exploration of the biological processes in which G4s are associated. G4 formation has now been widely investigated and implicated in multiple biological processes (**Figure 1.7**). These roles include telomere maintenance, DNA replication, phase separation events, regulation of chromatin architecture, and transcriptional reprogramming.(111)(112)(113)(114) How and why G4s form, and are involved with such processes, remains a key question. Further understanding of the mechanisms by which G4s have an impact on cellular processes could unveil new therapeutic targets to perturb a wide variety of cellular functions in diseased cells.

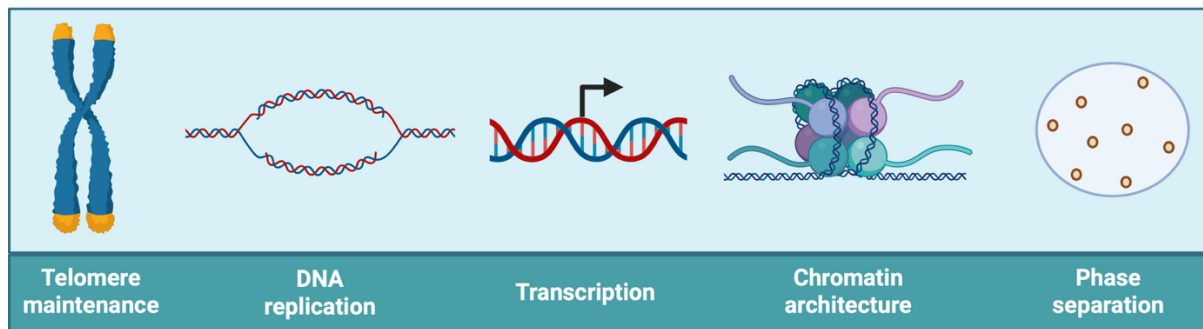


Figure 1.7 Biological roles of G4s.

Proposed biological roles of DNA G4s in telomere maintenance, DNA replication, transcriprional regulation, chromatin architecture and phase seperation.

1.3.2.1. Telomere maintenance and DNA replication

The idea that G4s may form at telomeric regions originated prior to generation of genomic G4 maps, instead from the crystal structure of G4 formation forming in a human telomeric DNA sequence.(89) Telomeres are the DNA/protein complexes which cap the ends of chromosomes, and have a protective function against damage to the chromosomal ends.(115) In mammalian cells, the telomeres terminate in a single-strand, which is highly G-rich, and has the propensity to form multiple G4 structures.(116) Indeed, G4s have been shown to form at telomeric regions within cells via BG4 immunofluorescence, with 20-25% of the G4 foci localising at telomeric regions.(96) However, the resolution of immunofluorescence is not sufficient enough to attribute G4s to the single strand overhang region of the telomere, but rather locates them at the telomere in general.(111) The exact role of G4s in telomere maintenance is still disputed. It has been suggested that G4s may play a protective capping role at telomeres, and have a positive effect on telomere maintenance, through colocalization with telomerase within cells.(117)(118) However, it has also been proposed that telomeric G4s can also act as an impediment to telomere replication, resulting in genomic instability. It is worth noting that studies that concluding this utilise a G4 stabilising ligand that holds the G4 structures in place, and prevents the resolution of G4s, resulting in increased DNA damage signals at telomere sites and impaired telomeric replication.(119)(120) Therefore, it could be postulated that it is the excessive stabilisation of telomeric G4 structures that results in

genomic instability, rather than the presence of the endogenous structures themselves. This is due to replication stalling as the telomerase cannot process a stabilised G4 structure, hence the stabilised G4 structure essentially inhibits telomerase and subsequently prevents extension of the 3' overhang and prevent loss of genetic material.(121) Therefore, stabilisation of telomeric G4s leads to progressive shortening of the telomeres which can be a highly mutagenic process and results in instable chromosomes lacking the protective cap and chromosomal imbalances.(122)

G4 prevalence is thought to vary throughout the cell cycle, with an increase in G4 foci being observed in S-phase, which is the DNA replication phase of mitosis, and the prevalence of G4 foci is reduced upon DNA polymerase inhibition.(96) This provided evidence that G4 formation increased during DNA replication. The mechanism driving this could be due to the active unwinding that helical DNA undergoes to replicate DNA. This creates a more favourable environment for G4 formation, as DNA becomes transiently single stranded, as it is uncoiled ready for replication.(123) Indeed, the formation of G4s during replication can cause replication fork stalling if they are not actively resolved, as DNA polymerase is unable to progress through a G4 structure.(124) G4 formation within cells is regulated by helicase enzymes, such as BLM, WRN and FANJ.(125)(126)(127) These enzymes can actively resolve G4s, to allow DNA replication and maintain genomic stability. Mutations in these resolvases increase the incidences of G4s in the genome and can contribute to replication fork stalling and genomic instability at sites of DNA replication.(128) However, G4 formation is not always deemed to be inhibitory to DNA replication. G4 forming sequences are enriched at DNA replication origins and may have functional importance in DNA replication initiation.(129)

1.3.2.2. Phase separation

G4s have the ability to trigger liquid-liquid phase separation events.(113) This is the ability of a homogenous solution to spontaneously divide into two distinct liquid phases.(130) This process is mediated by weak interactions occurring between nucleic acids and nucleic acid binding proteins.(113) A biological example of this phenomena is

membrane-less organelles in mammalian cells, such as the nucleolus, where the self-assembly of proteins and RNA occurs to form a liquid-like condensate within the nucleus.(131) Super-enhancers have also been shown to exert their transcriptional effect through the formation of condensates, via liquid-liquid phase separation, to localise transcriptional regulatory proteins close to their target genes.(132) G4s have recently been implicated in the triggering of phase separation events in the absence of any proteins.(114) Therefore, they may indeed be involved in phase separation events occurring in cells, although where, how, and why this occurs is still to be uncovered. It has been proposed that G4s may be linked to phase separation events around promoter sites, to aid transcriptional activation at the CCND1 gene.(133) As such, the capacity of G4s in phase separation may be linked to transcriptional regulation, although this concept is still largely unexplored.

1.3.2.3. Transcriptional regulation and chromatin architecture

The regulatory role that G4s play in chromatin architecture and transcription is key in classifying them as potential epigenetic regulators. The idea that G4s are associated with transcriptional regulation is multifaceted, occurring through multiple mechanisms which will be discussed below.

One of the first ideas that G4 could be involved in transcriptional regulation came from the publication of a crystal structure of a G4 formed from the sequence of the human *c-MYC* promoter, a key regulatory site.(134) This concept that G4s may have regulatory role was further solidified after computational G4 predictions found that they were enriched at promoter sites throughout the genome.(101) Indeed, >40% of human promoter regions contain at least one G4-forming motif.(101) The generation of endogenous G4 maps corroborated this, and enabled association of G4s and gene expression for the first time.(97) Here it was reported that promoters containing a G4 had significantly higher levels of gene expression, relative to promoters that did not contain a G4 (**Figure 1.8 A**).(97) It was speculated that this association with elevated transcription may be a consequence of G4s preferable formation in accessible chromatin sites, which are linked

to active transcription. Hence, G4s forming at promoters will also be accessible promoter sites, which are usually actively transcribed. However, it has been reported that it is the formation of a G4 structure in some promoter sites that is responsible for increased expression of the target gene, and not just the accessibility of the promoter.(135) A promoter site was investigated in two cell lines that had the same promoter sequences and accessibility, yet a G4 formed in one cell line and not the other. Interestingly, the cell line with the promoter containing a folded G4 consistently showed increased gene expression.(135) This is mediated by transcription factor binding to promoter G4s, with higher levels of binding being reported at promoter sites containing G4s.(87)(135) This suggests that G4s may act as transcription factor binding hubs in regulatory promoter sites, to elevate transcription of the proximal gene. In fact, this idea has been confirmed for the G4 present in the *c-MYC* promoter, in which transcription factors have been shown to interact with the *c-MYC* promoter through the preferential interaction with the G4 structure, as opposed to the duplex sequence.(136) This evidences G4s involvement in transcriptional regulation through transcription factor binding at key regulatory promoter sites.

Transcription factors do not only bind promoters, but also to more distal sites relative to the target gene, such as enhancers. These enhancer sites physically interact with the target promoter to elevate gene expression, even though they can be thousands of base pairs apart. This interaction is mediated by transcription factors which bind to motifs present in the promoter and enhancer sites, which are then bound together via a mediator complex. This interaction initiates chromatin remodelling and DNA looping to bring the regulatory elements into close proximity.(137)(138) Cohesin is the major protein involved in extruding the DNA to form these promoter-enhancer loops, bringing the two in close proximity.(139) The CCCTC-binding factor (CTCF) then acts to anchor and stabilise the DNA loop through binding to DNA sites, stalling cohesin.(140) G4s are also enriched at chromatin loop boundaries, and adjacent boundaries containing G4s readily interact with each other owing to the high enrichment of chromatin remodelling protein binding sites.(141) Moreover, the presence of G4s at loop boundaries was found to reinforce the anchoring ability of CTCF, suggesting that they may play a role in

establishing loop boundaries that could bring promoter and enhancer elements together (Figure 1.8 B).(141)

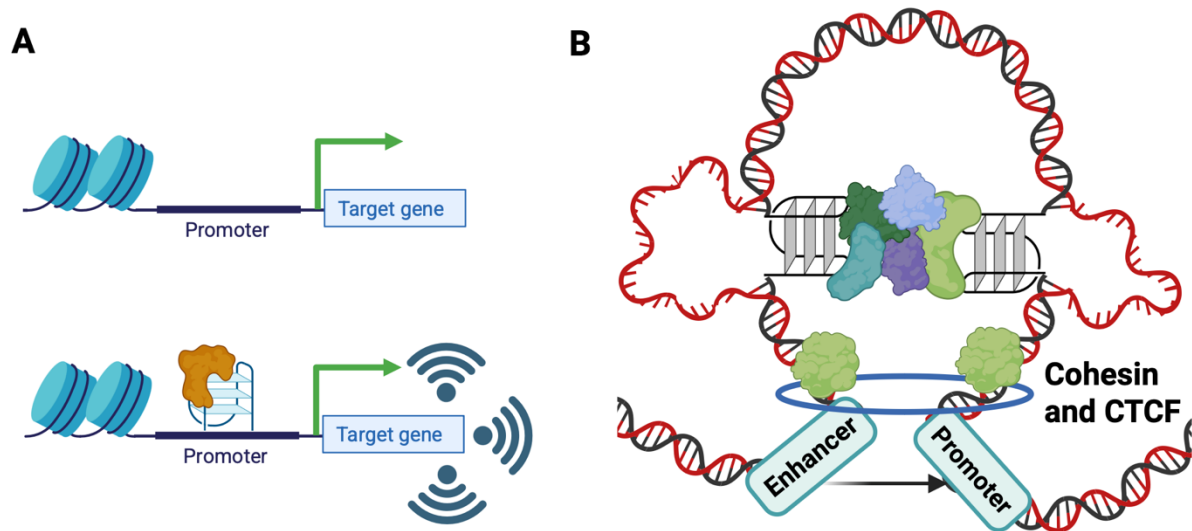


Figure 1.8 G4s transcriptional role in promoters and looping.

A) Promoter regions containing G4s have significantly higher levels of gene expression compared to promoters which do not have a G4. This is because transcription factors (yellow) bind to G4s and G4 containing promoters at a higher level to promote transcription. **B)** G4s are enriched at chromatin loop boundaries and helped to reinforce the anchoring ability of CTCF and cohesin to maintain chromatin looping, which is the mechanism by which promoters and enhancer elements are brought into proximity. Their ability to bind transcription factors aids this by reinforcing the anchoring of distal elements.

This idea has been consolidated by a study investigating the co-localisation of G4s and R-loops at CTCF sites, reporting that attenuation of R-loops, or deletion of specific G4-forming sequences, led to a reduction in CTCF binding and gene expression.(142) R-loops are a three-stranded nucleic acid structure comprised of a DNA-RNA hybrid and a displaced single-strand of DNA.(143) They form during replication and transcription, and R-loop formation is favoured during DNA negative supercoiling, at G-rich regions, or in

the presence of DNA nicks.(144) R-loops are intertwined with chromatin modulation. It is suggested that R-loops may be a product of chromatin modulation for transcriptional regulation, but that they may also contribute to chromatin closing prior to transcriptional repression.(145) Due to the greater thermodynamic stability of GC base pairing between RNA and DNA, R-loops preferentially form at GC-rich regions, which also have the propensity to form G4s in the non-template strand (**Figure 1.9 A**).(142) RNA:DNA hybrids form most stably with C-rich regions of DNA, leaving a G-rich single-stranded DNA non-template displaced which can then fold into a G4.(113) Formation of a G4 on the non-template strand has been shown to stabilise these R-loops, and promotes transcription via successive R-loop initiation, and RNA polymerase initiation.(146) Hence, the association of G4s and R-loops in the context of transcriptional regulation is complex. On one hand, co-occurrence of these structures at chromatin loop boundaries and CTCF binding sites can modulate transcription via the promotion of regulatory sites coming into proximity.(140) On the other hand, induction of G4 formation through inception of an R-loop at a promoter or transcriptional start site can also promote transcription proximally, through successive R-loop formation.(146) Indeed, both structures are associated with genomic instability upon over stabilisation or enrichment, which can negatively impact transcription. Down-regulating G4 resolvases, or targeting G4s with ligands to stabilise them, increases the number of G4s within cells and in turn the number of R-loops.(147)(148) This can stimulate DNA damage via double strand DNA breaks, exemplified through the enrichment of 53BP1 and γ H2AX double strand break markers after G4 ligand incubation. In cells that are deficient in DDR, such as *BRCA*-mutant cells, this accumulation of DNA damage can trigger apoptosis.(148)(149) This highlights how G4s can be exploited to promote genomic instability within cells.

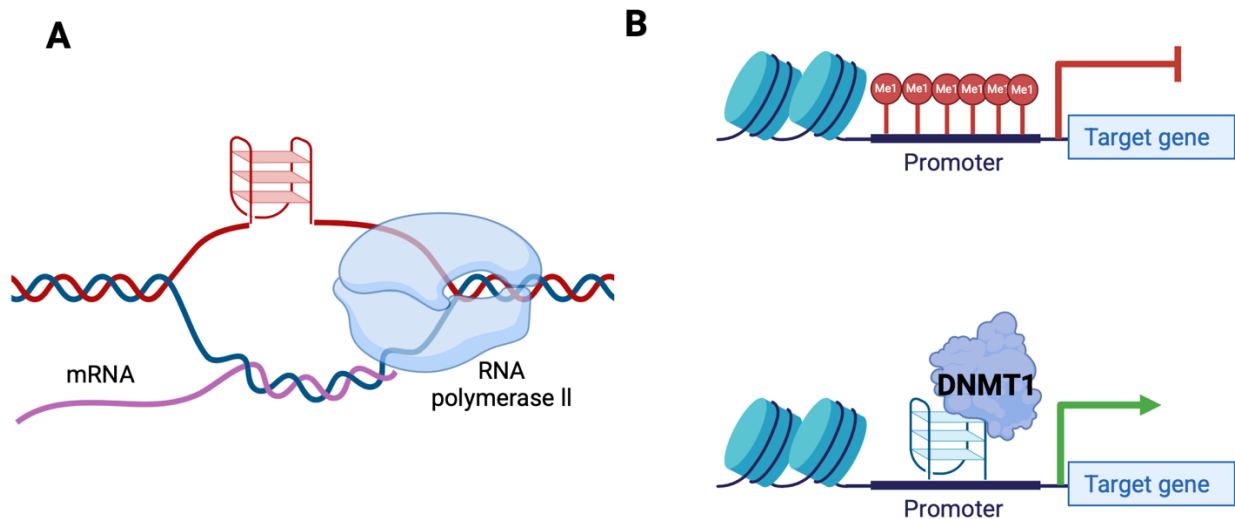


Figure 1.9 G4s transcriptional role in R-loops and DNA methylation.

A) G4s can form on the non-template strand of R-loops and stabilise the loop formation which promotes transcription. **B)** G4s can sequester DNMT1, which protects nearby CpG islands from methylation. This can facilitate transcription as it prevents hypermethylation of CpG promoter sites which inhibits transcription.

G4s have also been shown to associate with a chromatin loop-promoting transcription factor; YY1.(150) Here, G4 formation was disrupted at YY1 binding sites via CRISPR genome editing of the underlying G4-forming sequence. This resulted in a significant reduction of chromatin looping events, and the corresponding decreased expression of genes regulated by the DNA loops.(150) Interestingly, the altered gene expression was not limited to genes proximal to G4 sites, but via distal enhancer sites that interact with promoters via chromatin looping, as shown by 3D chromatin mapping; Hi-C.(150) G4s involvement in long-distance chromatin interactions has further been supported through reports that artificially inserting G4-forming sequences into the genome induces significantly more 3D chromatin looping interactions, relative to when a non-G4-forming sequence was inserted.(151) Furthermore, recruitment of enhancer histone marks and transcriptional co-activators was increased upon the addition of G4 structures, but not in the control.(151) These studies demonstrate that G4s may be involved with regulation of chromatin architecture, which in turn regulates gene expression. Moreover, it highlights

that G4s may work cooperatively with other epigenetic regulators. This suggests that G4s regulatory impact on transcription can be distal as well as proximal, like in the enrichment at promoter sites, and further consolidates G4s involvement in the regulation of chromatin architecture to facilitate transcriptional regulation.

G4s have been shown to actively bind chromatin remodelling proteins, further implicating them in the maintenance of chromatin architecture. These include CSB, SMARCA4 and ATRX, which could act to maintain open and active chromatin states to promote transcription.(152)(153)(154) Chromatin modulators can bind the genomic sites that G4s have been detected, and that this binding is enriched at gene promoters, a highly regulatory site for transcription.(153) Furthermore, G4s can change local chromatin architecture via nucleosome exclusion, to prevent the formation of nucleosomes where the G4 is present and facilitate transcription.(141) It has recently been reported that this is particularly prominent at sites of maximal promoter activity, and suggested that G4s play an active role in controlling nucleosomal exclusion at these sites.(155) Additionally, this study eludes to the idea that transcription associated with promoter G4s does not precede G4 formation, but instead stabilises the structure. This suggests that G4 formation at certain locations isn't a random byproduct of elevated transcription, in which formation is then beneficial to the cell and hence stabilised, but rather a mechanism leveraged by cells to initiate transcription at certain regions.(155) This idea is in its infancy, but provides an interesting perspective that G4s could drive transcriptional regulation through nucleosomal exclusion at promoter sites.

Owing to the close interplay that G4 formation has with the chromatin landscape, other epigenetic drugs, such as histone deacetylase (HDAC) inhibitors, can act to regulate the G4 landscape within a cell. Functioning HDAC removed acetyl marks on histones, which then promotes chromatin condensation and downregulates transcription of those regions. Inhibiting HDAC prevent deacetylation of histones by HDAC proteins, thus leaving acetyl marks on histones and maintaining an open chromatin state.(66) As G4s preferentially form in regions of open chromatin, HDAC inhibitors promote G4 formation through maintaining an open chromatin landscape.(156) This highlights the close

relationship that G4s have with other epigenetic marks, and how targeting other epigenetic modifications can have a secondary impact of the G4 landscape within a cell. Moreover, it exemplifies how G4 stabilisers can have opposite effects on gene expression compared to the targeting of other epigenetic marks. For instance, inhibiting HDAC and thus promoting a relaxed chromatin state would increase gene expression in those relaxed regions. Conversely, the stabilising of G4s in those regions would downregulate gene expression.

G4s have also been shown to regulate transcription via potential regulation of the methylome. G4 presence has been tightly correlated with hypomethylation of CpG sites, and 5mC has been shown to inhibit formation of G4s in both open and closed chromatin states.(100)(157) *In vitro*, G4s have been shown to bind DNMT1, and in turn impair its catalytic activity, in a sequence independent capacity.(100) It is thus proposed that G4s sequester DNMT1 to protect nearby CpG sites from methylation (**Figure 1.9 B**). Furthermore, the sites further up and down stream G4 structures have been shown to have a higher-than-average level of methylation, highlighting that this is a G4 site specific depletion of methylation.(100) This can prevent hypermethylation of CpG sites, particularly in promoters, which normally suppresses transcription. However, it can also be proposed that 5mC methylation directly inhibits G4 formation to repress transcription, highlighting a close interplay between the two mechanisms of transcriptional regulation.(157) Indeed, knocking out DNMT1 increases the number of G4s present within cells, and also promoted global transcriptional activity, as would be expected.(157) The association between DNA methylation and G4 formation is also dependent on G4 stability.(158) The more stable a G4 is, the more methylation efficiency decreases, indicating that modulating G4 stability could be a viable tool for regulating methylation, and in turn gene expression, within cells.(158)

The use of G4 stabilisers to counteract these endogenous epigenetic regulatory effects is an intriguing concept which is contested in the literature. For instance, the stabilising of a G4 structure to prevent transcription factor binding and the 'binding hub' role of a promoter G4 would act to downregulate transcription and therefore reverse the

endogenous impact of native G4 formation. Conversely, the presence of a G4 may prevent nucleosome formation at these sites and thus would maintain an open promoter state and subsequently allow for active transcription. The roles of G4 stabilisation could work in harmony together at different regions, promoting different things. However, it could be argued that the ability of a G4 to bind transcription factors to enable chromatin looping events and active transcription at promoter sites is more crucial to maintaining elevated levels of transcription. This is because even if a G4 is stabilised and prevents chromatin closing at a promoter site, the binding of the stabiliser and subsequent prevention of transcription factor binding negates the transcription promoting state of the chromatin and as such stabilisation hinders the role of endogenous G4 formation anyway.

Overall, G4s can play a multifaceted role in transcriptional regulation that is tightly linked with maintenance of chromatin architecture to achieve this biological role. This can range from acting as transcription factor binding hubs in promoter regions to facilitate elevated binding and transcription, to modulating chromatin looping events to bring together distal regulatory elements, such as promoters and enhancers. There is an increasing understanding that G4s could be considered to be epigenetic regulators, as they have the capacity to modulate transcription whilst leaving the underlying genetic sequence in which they form untouched, the prerequisite of an epigenetic regulator. Instead, a change in DNA structure is enough to facilitate transcriptional changes, which has evidenced at multiple regulatory sites.

1.3.3. G4s in cancer

Cancer was one of the first diseases that G4s were thought to be biologically implicated. G4s form more so in cancer cells, relative to healthy cells, and this is a phenomenon observed in multiple different cancer types.⁽⁹⁷⁾⁽¹⁵⁹⁾ This has also been validated using both G4 genomic mapping techniques, and BG4 immunofluorescence, which shows a significantly higher level of G4 formation in human stomach and liver cancers, breast cancer cells, lung, liver, and head and neck cancers.⁽⁹⁷⁾⁽⁹⁶⁾⁽¹⁶⁰⁾⁽¹⁶¹⁾ Besides differences in the global numbers of G4s, the enrichment of G4s and G4-forming motifs in certain genomic

locations suggests a biological function of G4s in cancer. G4s and G4-forming motifs are typically enriched in oncogene promoters, whilst there is a lack of G4s or G4-forming motifs in promoters of tumour-suppressor genes.(162)(97) This highlights that the formation of a G4 in an oncogene promoter may be leveraged by the cell to sustain elevated gene expression of the oncogene, contributing to carcinogenesis. It has been reported that destabilising alterations to G4 structure in some oncogenic promoters, such as MYC, KRAS and KIT, correlates to decreased gene expression (**Figure 1.10**).(134)(163) Conversely, decreased expression of tumour suppressor genes also contributes to cancer progression, and lack of G4-forming motifs in these regions mean that a G4 cannot form to help maintain increased levels of expression at these sites, to combat carcinogenesis.

1.3.3.1. Telomeric G4s and cancer

Due to the implication of G4s in biological processes, such as transcriptional regulation and telomere maintenance, as well as the enrichment of G4 structures in cancer cells, it was thought that perturbing these structures through chemical stabilisation could be used as an effective anti-tumour strategy.(159) Stabilisation of G4s often acts to reverse the biological impacts that G4 formation has endogenously. Hence, G4 stabilisation could impact telomerase activity and replication, promote or inhibit transcription, stall DNA replication, and induce genomic instability, consequently triggering apoptosis. Telomerase activity is upregulated in up to 90% of cancers, which allows cancer cells to replicate without telomere shortening, giving cancer cells unlimited replicative potential.(164) As G4s are enriched at telomeric regions, it is hypothesised that the stabilisation of these structures in these locations can block the activity of telomerase to prevent uncontrolled cell division. This would be a targeted therapy towards cancer cells, owing to the upregulation of telomerase in these cells, and not in somatic cells. Indeed, the G4 stabilisation at telomeric sites has been reported to reduce proliferation of tumour cells via telomerase inhibition or dysfunction, depending on the ligand used (**Figure 1.10**).(111)(165)(166) Targeting the telomere via uncapping of the protective ends is a treatment reminiscent of that used to target CSC, highlighting that targeting G4s can be a viable strategy for many different cancer cell states.(166)

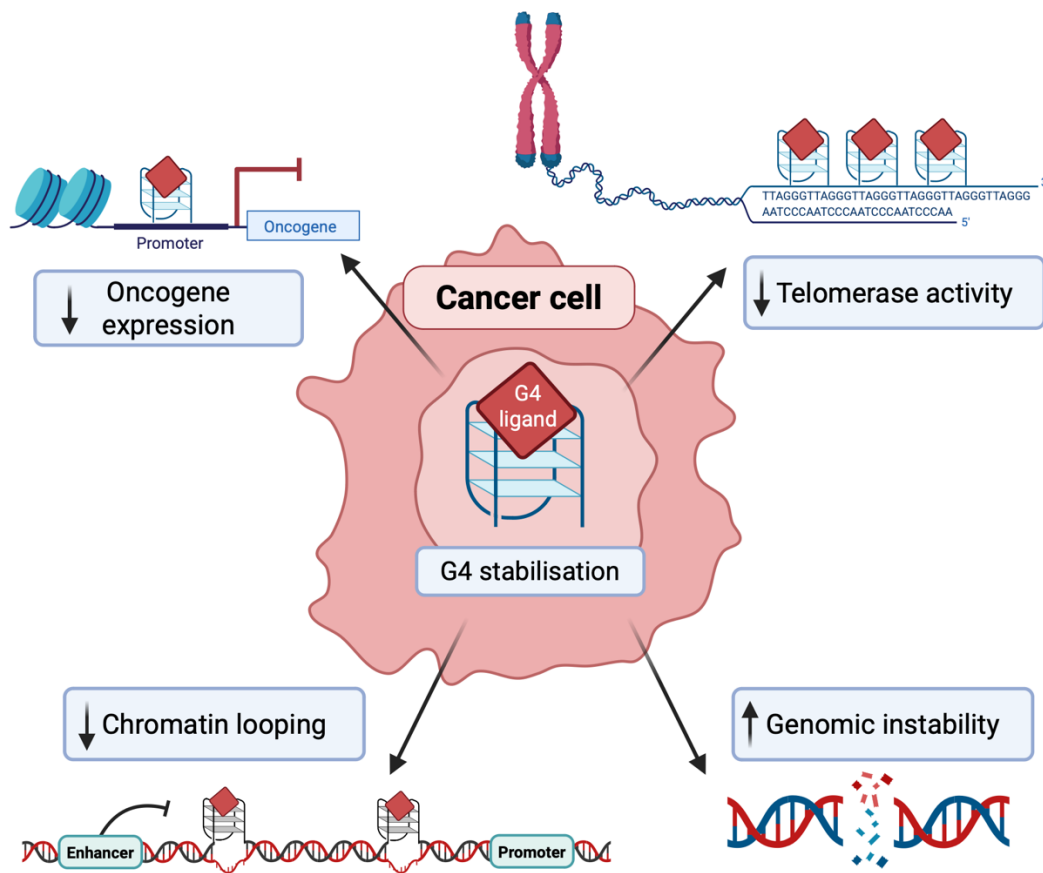


Figure 1.10 G4 stabilisation and implications in cancer cells.

These include reduced expression of oncogenes, decreased chromatin looping events, reduced telomerase activity and increased genomic stability via the induction of double strand DNA breaks.

1.3.3.2. G4s and genomic instability in cancer

Targeting G4s to induce apoptosis via unsustainable genomic instability is one of the primary mechanisms by which G4s can be exploited in cancer therapy. This is exemplified by the combination of G4 stabilisation with an inherent deficiency in homologous recombination, that leads to the accumulation of DNA double strand breaks and checkpoint activation (**Figure 1.10**).⁽¹⁶⁷⁾⁽¹⁶⁸⁾ In the absence of functional homologous recombination, via *BRCA*-mutation, the DNA damage induced by G4 stabilisation by PDS

is unable to be effectively repaired. This results in genomic instability, which drives checkpoint activation, cell cycle arrest, and apoptosis.(167) It is important to note that the stabilisation of G4s often has the opposite biological impact to that of an endogenous G4. This is because G4 ligands prevent G4-protein interactions and stabilise the structure to prevent resolution. Thus, stabilisation of G4s promotes genomic instability, as it cannot be actively resolved to allow for regular biological functions of the cell to occur, which often results in DNA damage occurring. This finding has been recapitulated in patient-derived xenograft tumours that have developed resistance to PARP inhibitors.(169) PDS was shown to hold the G4 structures in place, disrupting replication and leading to fork stalling which induced DSBs within the tumours. The toxicity of PDS was further potentiated *in vivo* by the cancer chemotherapy drug paclitaxel, highlighting that G4 stabilising drugs can be used in combination with traditional cancer treatments in *BRCA*-mutant cancers.(169) G4 stabilisation is not just limited to PDS in these instances. CX-5461 is a G4 stabilising molecule that has been reported to be selective lethal in *BRCA1/2*-deficient cancer cells, and patient-derived xenograft models.(149) This includes tumours resistant to PARP inhibitors, once again strengthening the concept that G4s are a viable therapeutic target in cancers with impaired DDR. Moreover, CX-5461 is currently in Phase I clinical trials for breast and ovarian cancers with DDR deficiencies.(170) However, CX-5461 is not exclusively a G4 stabiliser, and was initially noted for its ability to inhibit RNA polymerase I-mediated transcription, and induction of senescence and autophagy.(170) Therefore, it's capacity to exert its therapeutic potential may stem from several mechanisms of action, of which G4 stabilisation is included. Regardless, CX-5461 has been reported to be a promising therapy for the treatment of HGSOC, in combination with homologous recombination-deficiency.(171) The combination of CX-5461 with PARP inhibitors has been shown to exacerbate replicative stress and DDR activation, which cannot be efficiently repaired in a DDR-deficient cell, triggering apoptosis. Moreover, this study demonstrates that CX-5461 has efficacy against cisplatin-resistant HGSOC patient-derived xenografts.(171) However this work does not recognise or comment on the role of G4 stabilisation in this context, instead focusing on the inhibition of RNA polymerase I-mediated transcription as the mechanism of therapeutic efficacy.(170) Since the promising therapeutic potential of CX-5461, other G4 ligands (MTR-106) have been developed which inhibit *BRCA*-deficient cancers *in vitro* and *in vivo*.(172) Once again, this was demonstrated in PARP inhibitor resistant cells, and when

using in combination with PARP inhibitors it resulted in a synergistic effect.(172) This further strengthens the idea that G4s can act as a therapeutic target to produce a synthetically lethal effect in *BRCA*-deficient cancers, and in a PARP inhibitor resistant cell state. This highlights the potential of G4 targeting in the treatment of resistant cancer types.

G4s have been shown to have a close interplay with PARP in the context of cancer. Not only do G4 stabilising ligands show therapeutic potential in tumours resistant to PARP inhibitors, but they also have been shown to interact with the PARP protein itself. PARP-1 is enriched at actively transcribed promoters, similarly to G4s.(173)(174) Furthermore, PARP-1 has been shown to bind certain G4s, and this in turn stimulates enzymatic activity.(175) This is because G4s are recognised by PARP as DNA damage, or they can recognise the DNA damage that G4s can cause if the formation of G4s is sustained and not actively resolved. This interaction has been therapeutically exploited through the generation of a dual-action G4 stabilising and PARP inhibiting compound.(176) This molecule acts by stabilising G4 formation *in vitro* and in cells, which induces DNA damage that PARP is recruited to and subsequently inhibited by the PARP inhibitor moiety of the compound.(176) In *BRCA*-mutant cells, this induces an antiproliferative effect, and further supports the idea that G4s can be targeted in anti-cancer treatment. Moreover, *PARP1* has been shown to contain a G4-forming motif in its promoter.(177) Selective targeting of this G4 could provide an alternative mechanism of overcoming resistance to PARP inhibitors, as the stabilisation of this G4 may actively reduce expression of *PARP1*. This is because some mechanisms of PARP inhibitor resistance include upregulation of *PARP* expression and the alteration of the PARP1 enzyme structure to prevent PARP inhibitors from binding yet still maintaining the enzymatic activity of PARP1.(178) Hence, a mechanism to actively downregulate expression of *PAPR1*, such as through G4 targeting, could provide an attractive solution. Indeed, some PARP inhibitors have been reported to exhibit mild binding to the G4 present in the *PARP1* promoter, which may provide a starting point for the design of molecules to target this specific G4 structure.(179)

G4s have also been associated with genomic instability in cancer through the positive correlation between the local densities of G4s and single nucleotide variants (SNVs), across a cohort of 13,480 cancer patients.(180) SNVs are point mutations in the DNA sequence and are the most common type of genetic variation in the genome. In a spatial association study, it was reported that there was a strong significant correlation between sites of SNVs and OQS from G4-Seq.(94)(180) An increase in the number of local OQS sites was associated with increased genomic instability, indicated by an increase in SNVs. Moreover, the majority of SNVs were in the flanking regions of the OQS site, rather than at the site of potential G4 formation. This provides alternative evidence that G4s are implicated in genomic instability, and that they could be prognostic markers in cancer.(180)

1.3.3.3. G4s and acquired drug resistance in cancer

In the context of drug resistance in cancer, G4 formation has been reported to be altered in drug-resistant glioblastoma cells, relative to the drug sensitive counterparts.(181) This is the first indication that G4 formation may be altered within cancer cells, in relation to acquired drug resistance.(181) This is unsurprising, however, given that G4s are dynamic structures implicated in transcriptional regulation, and that altered expression of signalling pathways is a key driver of acquired drug resistance. However, the impact that the altered G4 distribution has on the transcriptome remains unexplored. This provides rationale that G4 distribution may be altered in the development of acquired drug resistance, whether that is passively, or having an active role in maintaining a drug-resistant state. Furthermore, G4 targeting has been used a mechanism to treat *BRCA*-deficient cancers, that are resistant to cisplatin or PARP inhibitors.(169)(149)(171) This corroborates the idea that G4s can be exploited therapeutically to overcome resistance to other cancer treatments, either via G4s potential epigenetic regulatory capacity in the acquisition of resistance, or exacerbating the DNA damage potential that the stabilisation of G4s elicits.

1.3.3.4. G4s and epigenetic regulation in cancer

Whilst the majority of research into G4s in the context of cancer has focused on exploiting the DNA damage inducing capacity of G4 ligands in combination with DDR-deficiency, less work has investigated the role G4s may play in epigenetic regulation in cancer.(83)(182) The enrichment of G4s in cancer cells, at oncogene promoters, and at chromatin loop boundaries to aid distal regulatory element interaction, implies that they could have play an epigenetic regulatory role in cancer cells, that contributes to cancer cell development and evolution.(97)(150)(135) In liposarcoma cells, promoter G4s have been associated with high levels of gene expression in open chromatin regions, highlighting the capacity of G4s in transcriptional regulation in cancer cells.(135) Moreover, the decreased gene expression associated with G4 perturbation in oncogene promoters further supports the idea that G4s have an epigenetic role in cancer.(163)(167) One study in human pancreatic ductal adenocarcinoma also highlighted how G4 targeting results in a down-regulation of a large number of genes which are implicated in survival, metastasis and resistance.(183) This highlights the transcriptomic effects that G4 targeting can have, in addition to induction of DNA damage. Furthermore, one study found that different breast cancer patient-derived tumour xenograft models had differences in differentially enriched G4 forming regions.(184) These regions are present at promoters of highly amplified genes with high levels of expression. As response to treatment and development of resistance can vary widely between patients, it was hypothesised that this difference in landscape, and subsequent elevation of expression, could be a biomarker for treatment response, as an epigenetic adaptation mechanism. Indeed, xenografts with more enriched differentially enriched G4 regions were more sensitive to G4 stabilisation, due to the higher abundance of G4 targets.(184) It could be proposed that perturbing G4s through stabilisation could reduce chromatin looping events that bring promoter and enhancer elements together (**Figure 1.10**). However, to date, it has not been probed in depth how G4 formation may change as cancer cells develop. Changes in the G4 landscape as cancer cell evolve may be associated with gene expression changes within the cell, from an epigenetic regulatory perspective.

1.4. Project rationale and aims

There is ever emerging evidence that G4s can be considered as epigenetic regulators within cells, owing to G4s transcriptional regulation capacity.(83) Moreover, G4s are enriched in cancer cells, and a molecule with G4 stabilising capacity is already in clinical trials for the treatment of breast and ovarian cancer, highlighting G4s as attractive targets in cancer therapy.(97)(170) However, how the G4 landscape changes as cancer cells evolve is a relatively unexplored concept. It could be hypothesised that as G4s are dynamic structures with similarities to epigenetic regulators, that the distribution of G4s across the genome may change as cancer cells evolve.(83) This could be reminiscent of how other epigenetic regulators change in distribution, such as changes in chromatin accessibility, to upregulate and downregulate gene expression of genes crucial to maintain the cellular state.(80) One instance when an epigenetic profile of a cancer cell may alter is in response to therapy. This is based on survival, and cancer cells readily adapt and evolve quickly to develop resistance to therapy, in order to survive the toxic environment. Indeed, acquired drug resistance is governed by a plethora of epigenetic adaptations to maintain altered expression of signalling pathways contributing to the molecular mechanisms of drug resistance.(59)(40)(70) Epigenetic adaptation is a relatively fast process, in comparison to the acquisition of genetic mutation, likewise, is the formation and resolution of G4s within cells. Furthermore, G4 landscapes in glioblastoma cells have been reported to be altered between the drug sensitive and drug-resistant states.(181) However, how these changes in G4 distribution may contribute to altered transcription of pathways implicated in drug resistance remains unexplored.

HGSOC is the most aggressive form of EOC, and 75% of patients will acquire some form of drug resistance over the course of their treatment.(21) Disease relapse, because of drug resistance, is a major cause of fatality, and resistance can develop via many molecular mechanisms. As resistance can develop over a short period of time, it has been speculated that epigenetic changes may be highly implicated in the development of drug resistance in HGSOC.(70) In fact, altered DNA methylation states, histone modifications, and chromatin accessibility changes have been implicated in the development of drug resistance in HGSOC.(185)(64)(80) Acquired drug resistance in HGSOC is also an example

where distal regulatory sites are prominent in epigenetic adaptation, to facilitate transcription through chromatin looping events, bringing together enhancer and promoter sites.(78)(186) As such, acquired drug resistance in HGSOC was deemed to be an ideal model to investigate if the G4 landscape changes as a result of cancer evolution, leading to therapy resistance, as it is reported to be highly regulated by epigenetic adaptation. Changes in G4 landscape can then be associated with changes in gene expression, to identify if the gain or loss of G4s is correlated to increased or decreased gene expression. Ultimately, it can be investigated if the changes in the G4 landscape are associated with the altered expression of signalling pathways implicated in drug resistance, indicating if G4s have an epigenetic regulatory capacity, in the context of drug-resistant HGSOC.

Therefore, it is hypothesised that the G4 landscape is altered as ovarian cancer cells evolve in response to treatment, and that it is the gain in G4s contributing to transcriptional upregulation, which leads to resistance to therapy. To this end, this research project aims to investigate if G4 formation may be used by HGSOC cells to epigenetically contribute to drug resistance. This will be achieved via the completion of specific aims, which are outlined below:

- 1) Generation of genomic G4 maps in drug-sensitive and drug-resistant paired, patient-derived HGSOC cell lines using G4 ChIP-seq, to uncover if there are differences in the G4 landscape (Chapter 2).
- 2) Intersection of G4 maps with transcriptomic (RNA-seq) data to investigate how the gain or loss of G4 in drug-resistant HGSOC cells is associated with gene expression changes in different genomic locations, and how this compares to changes in chromatin accessibility (ATAC-seq) (Chapter 3).
- 3) Perturbation of G4s via stabilisation to assess how G4 targeting impacts the transcriptome of HGSOC cells, and if this can alter response to cisplatin (Chapter 4).
- 4) Comparison of G4 maps and transcriptomic effects in an alternative HGSOC drug-resistant cell line pair, to assess whether any impacts seen are cell specific or more broadly implicated in HGSOC drug resistance (Chapter 5).

Overall, this research project provides the first in depth investigation into how the formation and distribution of G4s changes in the development of acquired drug resistance in HGSOC cells. This will provide a greater insight into the epigenetic regulatory role that G4s can play in the context of cancer evolution and assess the viability of G4 targeting as an approach to combat acquired drug resistance.

2. The optimisation of genomic G4 mapping in HGSOC cells

2.1. Chapter overview

To investigate the epigenetic impact that G4s might have on drug resistance in ovarian cancer cells, the genomic locations of G4s must first be determined in drug-sensitive and drug-resistant ovarian cancer cell lines. To ensure the differences between the two states of therapy responsiveness were not due to the inherent differences between the cell lines and are instead more likely to be mediated by acquired epigenetic adaptations, cell lines derived from a singular patient were used. The generation of this data can unveil whether G4s are enriched in drug resistance-relevant pathways, as opposed to random formation, that may be implicated in acquired drug resistance in ovarian cancer cells.

2.2. PEO1 and PEO4 paired ovarian cancer cell lines

To investigate how G4s may be implicated in the development of drug resistance in ovarian cancer, the established paired cell lines PEO1 and PEO4 were selected. These cell lines were derived from a single patient with HGSOC, before (PEO1), and after (PEO4), the development of chemotherapy resistance (**Figure 2.1 A**). Therefore, the genetic backgrounds of the cell lines are comparable, with the differential response to chemotherapy being the main distinction between them. To confirm a difference in sensitivity to cisplatin was experimentally appreciable between the two cell lines, a 72-hour cell viability assay was conducted using increasing concentrations of cisplatin. The GI_{50} value for PEO4 cells was 7-fold higher than that of PEO1 (**Figure 2.1 B**). This is in line with the original 3-fold difference initially reported for these cell lines (187), and comparable to previous studies that report a 3-11-fold difference in

sensitivity.(188)(189)(190)(191) These results confirm that PEO4 cells exhibit a lower drug sensitivity to cisplatin than PEO1, substantiating their selection as the cell line model to probe G4 involvement in ovarian cancer cell drug resistance.

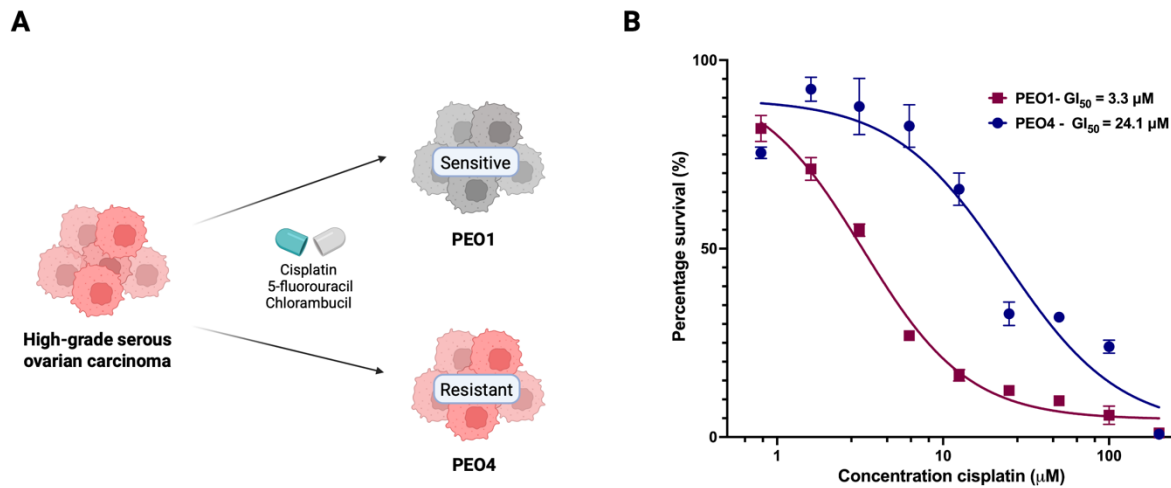


Figure 2.1 PEO1 and PEO4 cells differential sensitivity to cisplatin.

A) Schematic of PEO1 and PEO4 derivation from a single patient before and after resistance to chemotherapy. **B)** MTS 72-hour cell viability assay measuring PEO1 and PEO4 response to increasing concentrations of cisplatin. Error bars show standard deviation from experiment conducted in technical triplicate. GI_{50} for PEO1 is $3.3 \pm 0.8 \mu\text{M}$, and $24.1 \pm 2.1 \mu\text{M}$ for PEO4.

The PEO4 cell line was acquired 10 months after PEO1, which is a relatively short time for treatment resistance to occur. This supports the idea that epigenetic adaptation may be involved in the development of drug resistance, as opposed to genetic mutations.(80) However, there are key genetic differences between PEO1 and PEO4 that could strongly contribute to acquired drug resistance. Namely, the polarity in BRCA2 status. PEO1 cells have a mutation in exon 11 of the *BRCA2* gene (Y1655X) that causes a stop codon, resulting in truncated, non-functional BRCA2 protein.(192) Consequently, homologous recombination is impaired in the PEO1 cell line, due to the crucial role in this pathway of DDR that BRCA2 plays.(193) However, PEO4 cells have developed another mutation in

the same amino acid (Y1655Y), that restores BRCA2 functionality, by preventing the stop codon from occurring. To confirm that these mutations were present in the obtained PEO1 and PEO4 cell lines, exon 11 of the *BRCA2* gene from RNA-seq data was visualised on IGV (**Figure 2.2**).⁽¹⁹⁴⁾ Indeed, the PEO1 and PEO4 cell lines had the expected mutations, confirming the difference in *BRCA2* status, and consequent homologous recombination capacity between the two cell lines. Other genetic differences between the two cell lines stem from copy number alterations. PEO1 cells have no normal copies of chromosomes 1, 9, 11 or 17, whilst PEO4 cells had only one normal copy of each.⁽¹⁹⁵⁾ The alteration in normal copies of these chromosomes could be the root of transcriptional differences between the two cell lines, owing to the fact that they have differing underlying gene doses.

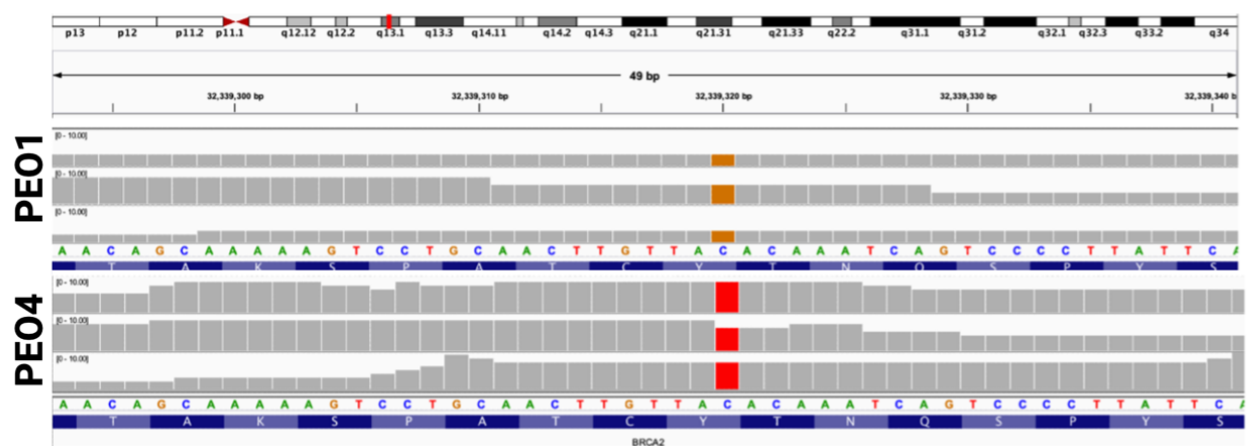


Figure 2.2 *BRCA2* mutational status of PEO1 and PEO4.

Genomic tracks from RNA-seq data in PEO1 and PEO4 cells visualised in IGV. In comparison to the reference genome and amino acid sequence, PEO1 cells have a C to G mutation (with the nucleotide marked brown corresponding to a mutation to G from the displayed reference sequence) causing a Y1655X mutation at exon 11. In PEO4 cells, there is a C to T mutation causing a Y1655Y mutation (with a nucleotide marked red corresponding to the mutation to T from the displayed reference sequence).

It has thus been speculated that this difference in DDR ability is the main reason behind the acquired drug resistance in PEO4 cells. Cross-linking drugs such as cisplatin cause DNA damage, which can result in apoptosis if repair of the lesions is unable to occur.(196) Hence, *BRCA*-mutant cancer cells are more sensitive to DNA damaging drugs than those with functional *BRCA* protein.(197) However, it has been reported that an additional mutation can occur in PEO1 cells that functionally reinstates *BRCA2*.(192) Despite the restored functionality, PEO1 still exhibits a 6-fold greater sensitivity to cisplatin in comparison to PEO4.(192) This indicates that the *BRCA2* status alone is not the sole contributor to drug resistance in the PEO1/4 cell line model, and epigenetic mechanisms may still be implicated. Based on these observations, the PEO1 and PEO4 cell lines were deemed suitable to study the potential epigenetic role of G4s in acquired drug resistance, despite the underlying genetic differences between the cells.

2.3. BG4 generation and validation

To obtain accurate genomic maps of G4s in the PEO1 and PEO4 cell lines, a G4-specific antibody must first be generated to ensure selective enrichment of G4 structures. The selectivity of the antibody is essential to ensure minimal off-target binding, leading to false positives. The G4 targeting antibody selected for this aim was BG4.(96) Expression was achieved using an autoinduction method optimised by Dr Denise Liano, a previous PhD student in the Di Antonio group.

2.3.1. BG4 expression and purification

In short, *E. coli* cells were transformed with a plasmid encoding for BG4 via heat-shock. To select for successfully transformed cells, kanamycin was added to the suspension as the plasmid also conferred antibiotic resistance. Expression of the BG4 antibody was auto induced through the activation of the T7 RNA polymerase.(198) The induced cells were grown to allow expression of the BG4 antibody in the periplasm. Following this, a series of cell lysis and treatment steps were undertaken to provide an endogenous protein mix, containing BG4. The target antibody was then purified using a column, and cobalt resin

which is a his-tag interactor, utilising the his-tag present at the C-terminal end of BG4. The antibody elution was concentrated by the buffer exchange method, using a spin column, ready for validation and use in G4 genomic mapping.

To assess the purity and concentration of BG4, the concentrated elution was subjected to Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis (SDS-PAGE). This technique involves the denaturation of proteins, using the detergent SDS, and then separating them based on the proteins molecular weight following the application of an electric field. BG4 has a molecular weight of around 30 kDa, and so the appearance of a band at this weight should indicate the successful expression and purification of the target antibody. The presence of this band was indeed seen in both the technical replicate lanes containing the BG4 concentrated elution, indicating expression had been achieved (**Figure 2.3 A**). To determine the concentration of the BG4, a serial dilution of a standard solution of bovine serum albumin (BSA) was run alongside the BG4 samples on the SDS-PAGE. The intensity of the BSA bands (determined using ImageJ (199)) were plotted against known concentration, using a simple linear regression determined by GraphPad Prism. The concentration of BG4 could then be ascertained by interpolation of the graph, using the average band intensity from the technical repeats of BG4 obtained from the SDS-PAGE (**Figure 2.3 B**). It was determined that the concentration of BG4 was 0.51 mg/ml, which is equivalent to 1.7 μ M. It is important to note that the quantification may not be totally accurate, as there is no error bars present on the standard curve as $n=1$. However, in the context of BG4 quantification, and approximation is sufficient for the downstream genomic applications.

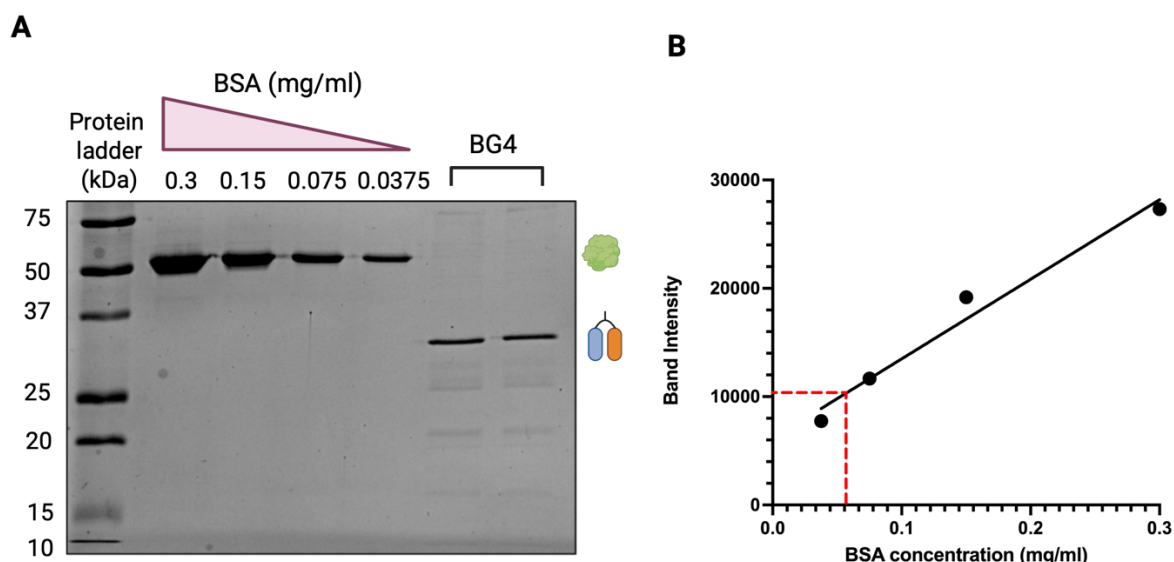


Figure 2.3 BG4 expression, purification and concentration determination.

- A)** SDS-PAGE gel showing a protein ladder next to a serial dilution of BSA standards used for concentration quantification. The lanes containing the concentrated BG4 elution show a band at around 30 kDa equating to the BG4 antibody. **B)** The simple linear regression plot of BSA band intensity and concentration used to determine BG4 concentration. The red dotted line corresponds to the band intensity of BG4 determined from the SDS-PAGE gel of n=1, and the interpolated concentration.

2.3.2. BG4 selectivity validation

Once sufficient amounts of pure BG4 had been obtained, its selectivity for G4 DNA needed to be determined via enzyme-linked immunosorbent assay (ELISA). In essence, an ELISA can test the selective binding of BG4 to G4 DNA over non-G4 DNA, by using oligonucleotides that either have the capacity to form a G4 structure or not. To test BG4 binding to G4 DNA, an oligonucleotide was used that was composed of a sequence found in the *c-MYC* gene promoter, that has been validated to form a G4 using circular dichroism by Dr Jenna Robinson, a previous PhD student in the Di Antonio group. The second oligonucleotide was a mutated *c-MYC* sequence that no longer had the ability to form a G4, to test BG4 binding to a similar DNA sequence yet not a G4 structure. An additional

control of wells containing no oligonucleotides was used to assess background binding of the antibody to the plate.

Biotinylated oligonucleotides were attached to a streptavidin coated 96-well plate, and BG4 was added to the wells using a 1:2 serial dilution. After washing the plate to remove any weak interactions, a secondary antibody was added to the wells to bind to the BG4, with its anti-FLAG moiety. This anti-FLAG antibody was also attached to a horse radish peroxidase (HRP) enzyme, which catalyses the TMB (3, 3', 5, 5'-tetramethylbenzidine) chromogenic substrate to cause a colour change from clear to blue. The reaction was then stopped using sulphuric acid, resulting in a colour change from blue to yellow, of which the absorbance can be measured at 450 nm. The intensity of the yellow colour is proportional to the amount of BG4 binding within the wells, and therefore selectivity of BG4 to the different oligonucleotide environments can be determined.

Using this methodology, it was found that BG4 bound strongly to the G4 forming *c-MYC* DNA, with a $K_d = 10.4 \pm 2.8$ nM (**Figure 2.4**). Whilst this K_d value is slightly higher than that reported in the initial BG4 generation paper, it remains in line with other reported values in the literature, and so was deemed to have good binding to G4 DNA.(96)(200) This contrasts with the mutant *c-MYC* and no oligonucleotide control, in which a K_d value was unable to be accurately calculated. Moreover, BG4 binding to either the mutant oligonucleotide, or to the 96-well plate, was near identical highlighting that BG4 has no preference for other DNA forms over general, non-specific binding to its environment. These results validated that a selective G4 antibody had been successfully expressed and purified for use in G4 mapping experiments. This ensures that selective enrichment of G4 structures over non-G4 structures could be confidently achieved.

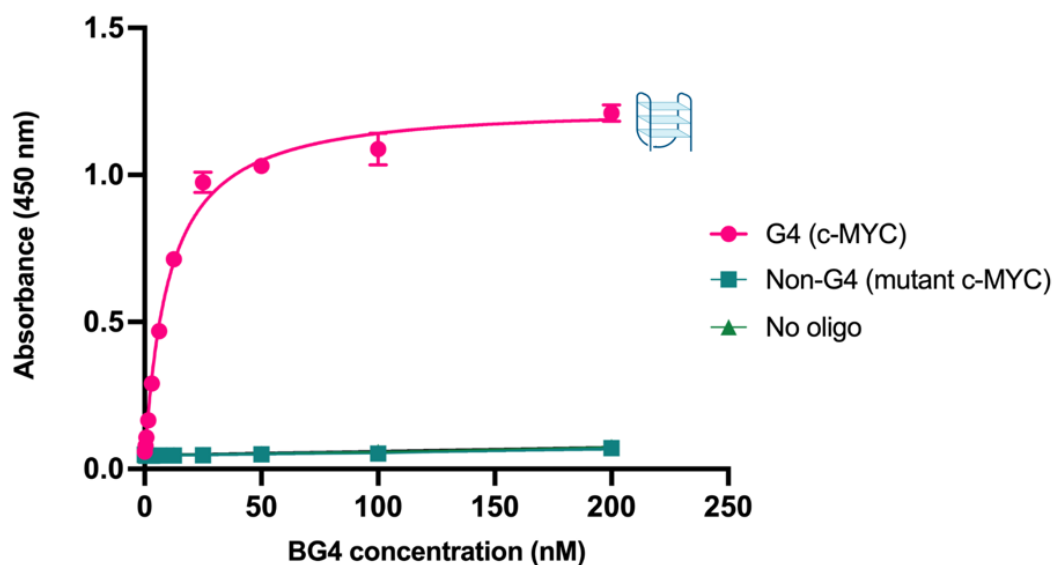


Figure 2.4 BG4 selectivity validation.

The BG4 binding ELISA curves for the G4 forming sequence (c-MYC), non-G4 forming (mutant c-MYC) and no oligo control well. The binding affinity of BG4 to c-MYC is $K_d = 10.4 \pm 2.8$ nM. Error bars are the standard deviation from two technical replicates.

2.4. Chromatin immunoprecipitation and quantitative polymerase chain reaction (ChIP-qPCR)

One of the most established methods to detect endogenous DNA G4s in human chromatin is with BG4 ChIP-seq.(97) This method cross-links cells to preserve the chromatin architecture, and folded G4s, before shearing via sonication to produce fragments between 100-600 bp. The sonicated chromatin then undergoes immunoprecipitation, using the validated BG4 antibody, to selectively pull-down fragments containing G4 structures. The G4 fragments are precipitated using magnetic beads coated with an anti-FLAG antibody, to selectively bind to BG4s FLAG tag. Prior to sequencing, ChIP efficiency can be assessed using qPCR, with primers targeting regions of the human genome previously validated to contain G4s, and others that target non-G4 regions.(99) A successful ChIP is considered to be when G4 regions are 5-fold enriched over non-G4 regions, and the average percentage input of G4 regions is >5%.(99) These parameters ensure that a good signal to noise ratio has been achieved through selective enrichment

of G4 fragments over non-G4 fragments, to allow for accurate peak calling of true G4 sites later on in the ChIP-seq pipeline.

This method to map G4s in PEO1 and PEO4 cells was selected due to the qPCR validation aspect of the protocol. This step allows validation of enrichment prior to next generation sequencing, meaning optimisation of the procedure in new cell lines could be conducted and assessed in a time and cost-effective manner. Alternative G4 mapping methods were published throughout the duration of this project, however in the context of this work ChIP-qPCR and sequencing was optimised and performed.(104)(108) Whilst G4 ChIP generates noisier data than other G4 mapping methods, such as CUT&Tag and Chem-Map, it was the most established and validated method of G4 mapping at the start of the project.

2.4.1. ChIP-qPCR in U2OS cells

To confirm the ChIP-qPCR could be performed successfully, it was first conducted on a human osteosarcoma cell line (U2OS), in which BG4 ChIP-qPCR has previously been published and validated.(99) This would ensure the ChIP protocol could be successfully reproduced on validated cell lines, before establishing it in PEO1 and PEO4 cells. To this end, chromatin from two batches of 4 million U2OS cells were cross-linked and extracted, before sonication with either 10 or 15 cycles of pulsed ultrasonic waves. 4 million cells were initially selected as it was the number present in a confluent T-75 cm³ flask, and because the original G4 ChIP protocol suggested between 1 million and 10 million cells dependent on the cell line.(99) To assess the fragment size distribution, the chromatin samples were run on a 2% agarose gel against a reference base pair ladder.

First, the sheared chromatin samples underwent reverse cross-linking and protein denaturation to release the DNA fragments and allow accurate quantification of fragment size. The DNA fragments were then run on an agarose gel to separate them based on their size, after applying a voltage to mobilise the fragments. The size of the fragments was

then qualitatively assessed by comparing them to a reference base pair ladder. For both samples that were sonicated for either 10 or 15 cycles, the bulk of the DNA fragments were sheared between the optimal 100-600 bp range (**Figure 2.5 A**). This indicates that the chromatin extracted from 4 million U2OS cells is the ideal amount to ensure successful sonication with these numbers of sonication cycles. Moreover, the concentrations of the two samples were 243 ng/ μ L (10 cycles) and 232 ng/ μ L (15 cycles), which fall within the desired chromatin concentration range (100-300 ng/ μ L), and show minimal variation between the two chromatin preparations. Subsequently, both preparations were deemed to be suitable for BG4 immunoprecipitation.

After ChIP had been performed, qPCR analysis was conducted to assess the enrichment of G4 regions over non-G4 regions. Throughout the duration of this project, primers targeting G4 regions were used to this end (MAZ, RPA3, KIF14 and SPRED2), which have been previously validated in multiple cell lines to target regions containing G4 structures.(99) As a non-G4 control, additional primers were used (ESR, TMCC1, IL36G), which target regions that do not form G4s in multiple cell lines.(99) The results of the qPCR show successful enrichment of G4-regions over non-G4 regions, and both the fold enrichment and G4 percentage input parameters exceed the desired thresholds. For the U2OS chromatin sonicated for 10 cycles, an average of 6.2% of the G4 regions were retained relative to 0.91% of the non-G4 regions, resulting in a 6.8-fold enrichment (**Figure 2.5 B**). A 7.1-fold enrichment was achieved for the 15-cycle sonicated U2OS chromatin, through an average of 9.2% of the G4 forming regions being retained from the input sample, in comparison to 1.3% of the non-G4 regions (**Figure 2.5 C**). These results indicated that chromatin could be extracted, sonicated and immunoprecipitated successfully for G4 DNA, reproducing published findings. This provided confidence to replicate this protocol in PEO1 and PEO4 cells, to generate G4 maps in these cell lines for the first time.

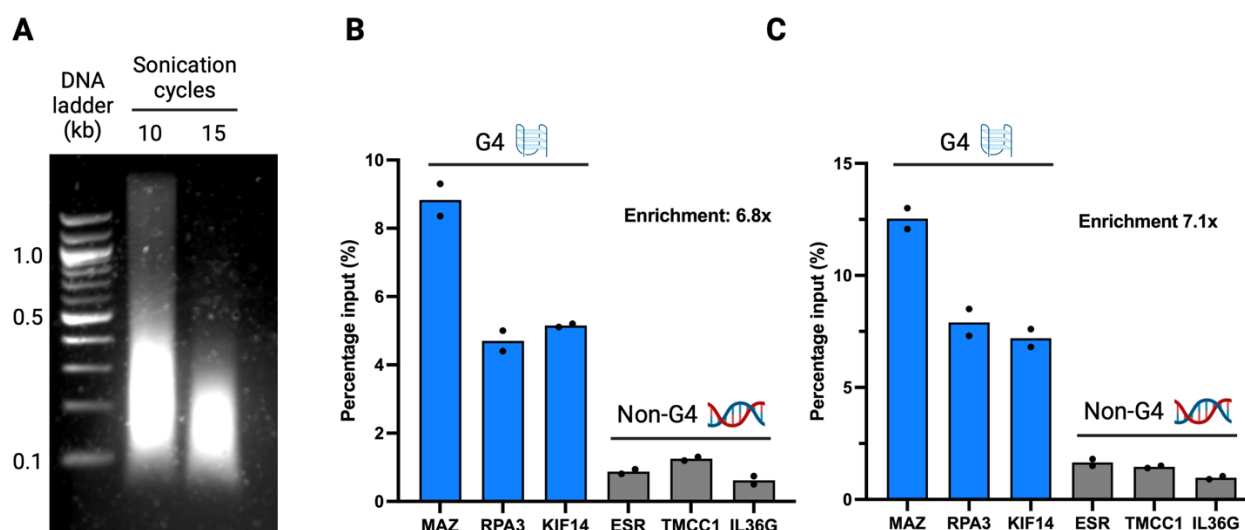


Figure 2.5 Results of ChIP-qPCR of U2OS cells.

A) Agarose gel showing DNA fragment size distribution of 4 million U2OS cells sonicated for 10 or 15 cycles. **B)** BG4 ChIP-qPCR of U2OS chromatin sonicated for 10 cycles showing a 6.8-fold enrichment of G4 regions (blue) and non-G4 regions (grey). **C)** BG4 ChIP-qPCR of U2OS chromatin sonicated for 15 cycles showing a 7.1-fold enrichment of G4 regions (blue) and non-G4 regions (grey). Percentage input values plotted as an average of n=2 technical replicates, with individual values for each genomic region represented as a dot.

2.4.2. Optimisation of ChIP-qPCR in PEO1 and PEO4 cells

Chromatin was cross-linked and extracted from 8 million PEO1 cells following the same protocol as for U2OS. 8 million cells were initially chosen due to the smaller morphology PEO1 relative to U2OS, with a confluent T-75 cm³ flask containing double the amount of PEO1 cells than U2OS cells. Subsequently, double the number of PEO1 cells were chosen for chromatin preparation considering the success of the U2OS G4 ChIP-qPCR, from samples prepared from a confluent T-75 cm³ flask. Two batches of 8 million PEO1 cells were sonicated for 10 or 15 cycles, and the fragments sizes assessed via agarose gel. Both chromatin samples had most of their DNA fragments between the optimal 100-600 bp range (**Figure 2.6 A**). However, inconsistencies arose with the handling of the samples,

as the concentrations differed; the 10-cycle sonicated chromatin was below the concentration threshold at 61 ng/ μ L whilst the 15-cycle sonicated chromatin was more than twice as concentrated at 141 ng/ μ L. This could be due to errors in the initial cell counting, or through cell loss throughout the cross-linking and sonication procedures. The variance between the samples highlights the importance of careful pipetting during wash steps throughout cross-linking and sonication, to ensure reproducible chromatin preparations can be obtained.

The 15-cycle sonicated PEO1 chromatin was taken forward for ChIP-qPCR, as the concentration was above the desired 100 ng/ μ L threshold, and the fragment distribution appeared to be within the 100-600 bp range. Encouragingly, a 4.9-fold enrichment of G4 regions (7.16%) over non-G4 (1.46%) was obtained, which is just shy of the desired 5-fold enrichment (**Figure 2.6 B**). Whilst lower than the fold enrichment achieved from U2OS chromatin, these results were suggestive that BG4 ChIP could be successfully used in PEO1 cells for genomic G4 mapping. Interestingly, the MAZ gene had a much lower percentage input relative to the other G4 control regions. It may be that the formation of a G4 in this region is more variable in PEO1 cells than in the other cell types that this region had been previously validated. However, there seemed to be large variation between the two technical repeats, in comparison to the other G4 control regions. Hence, this may be a technical error, as opposed to a true biological interpretation. Subsequently, additional biological replicates should be performed to assess this.

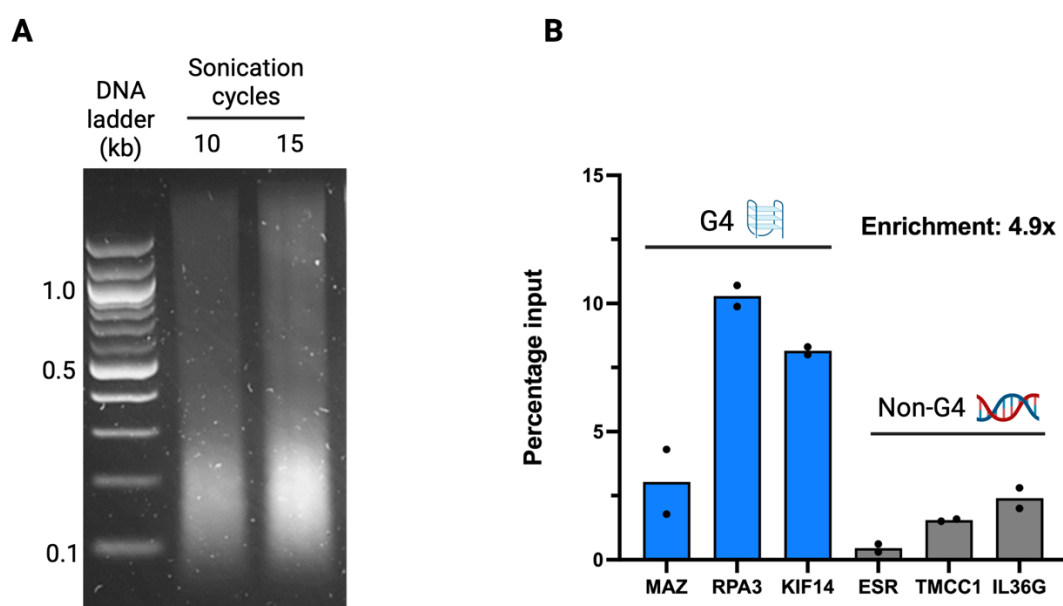


Figure 2.6 Results of ChIP-qPCR of PEO1 cells.

Agarose gel showing DNA fragment size distribution of 8 million PEO1 cells sonicated for 10 or 15 cycles. **B)** BG4 ChIP-qPCR of PEO1 chromatin sonicated for 15 cycles showing a 6.8-fold enrichment of G4 regions (blue) and non-G4 regions (grey). Percentage input values plotted as an average of $n=2$ technical replicates, with individual values for each genomic region represented as a dot.

It was hypothesised that the reason for the decreased fold enrichment of G4 regions over non-G4 regions in PEO1 cells relative to U2OS cells, could be due to the lower concentration of the chromatin sample (132 ng/ μ L and 243 or 232 ng/ μ L, respectively). A lower chromatin concentration means a reduced number of G4 targets present in the sample, which in turn increases the relative contribution of non-specific background noise from non-specific BG4 binding. Alternatively, PEO1 cells could have a lower abundance of G4s relative to U2OS cells, and so the lower enrichment may be attributed to this. To probe whether increasing chromatin concentration could improve the fold-enrichment, chromatin samples were prepared from 8 and 10 million PEO1 cells and sonicated for 15 cycles. The increase in cell number whilst keeping sonication buffer volume constant, would increase the sample concentration, and allow ChIP to be

performed on a more concentrated sample. However, increasing the cell number to 10 million resulted in inefficient sonication, as seen by the concentrated portion of high base pair fragments towards the top of the gel (>1 kb) (**Figure 2.7 A**). These fragments were much larger than the optimal 100-600 bp range, and resulted in minimal enrichment of G4-regions over non-G4 regions (**Figure 2.7 C**). Long DNA fragments are more likely to contain non-specific binding sites, which increases the background signal, and makes quantification of the target G4 regions harder to distinguish. Hence, it is unexpected that poorly sonicated chromatin, such as that generated from 10 million PEO1 cells, yields unsatisfactory enrichment.

ChIP-qPCR was also conducted on the 8 million PEO1 cell repeat that was sonicated for 15 cycles, to assess if successful ChIP-qPCR could be repeated for another biological repeat. Despite the apparent successful sonication indicated by the agarose gel (**Figure 2.7 A**), no enrichment of G4-regions over non-G4 regions was detected (**Figure 2.7 B**). This was surprising, since the chromatin had been prepared in an identical manner to the initial successful PEO1 sample. The concentration of the new PEO1 sample was higher than the initial one (180 ng/ μ L), however it would be perceived that this increase would be beneficial for ChIP enrichment and not abolish it completely. Similar results were obtained for PEO4 ChIP-qPCR. Whilst 12 million cells resulted in ineffective sonication and subsequent lack of enrichment, chromatin generated from 10 million cells (that looked optimal on an agarose gel) also yielded negligible enrichment (**Figure 2.7 D&F**).

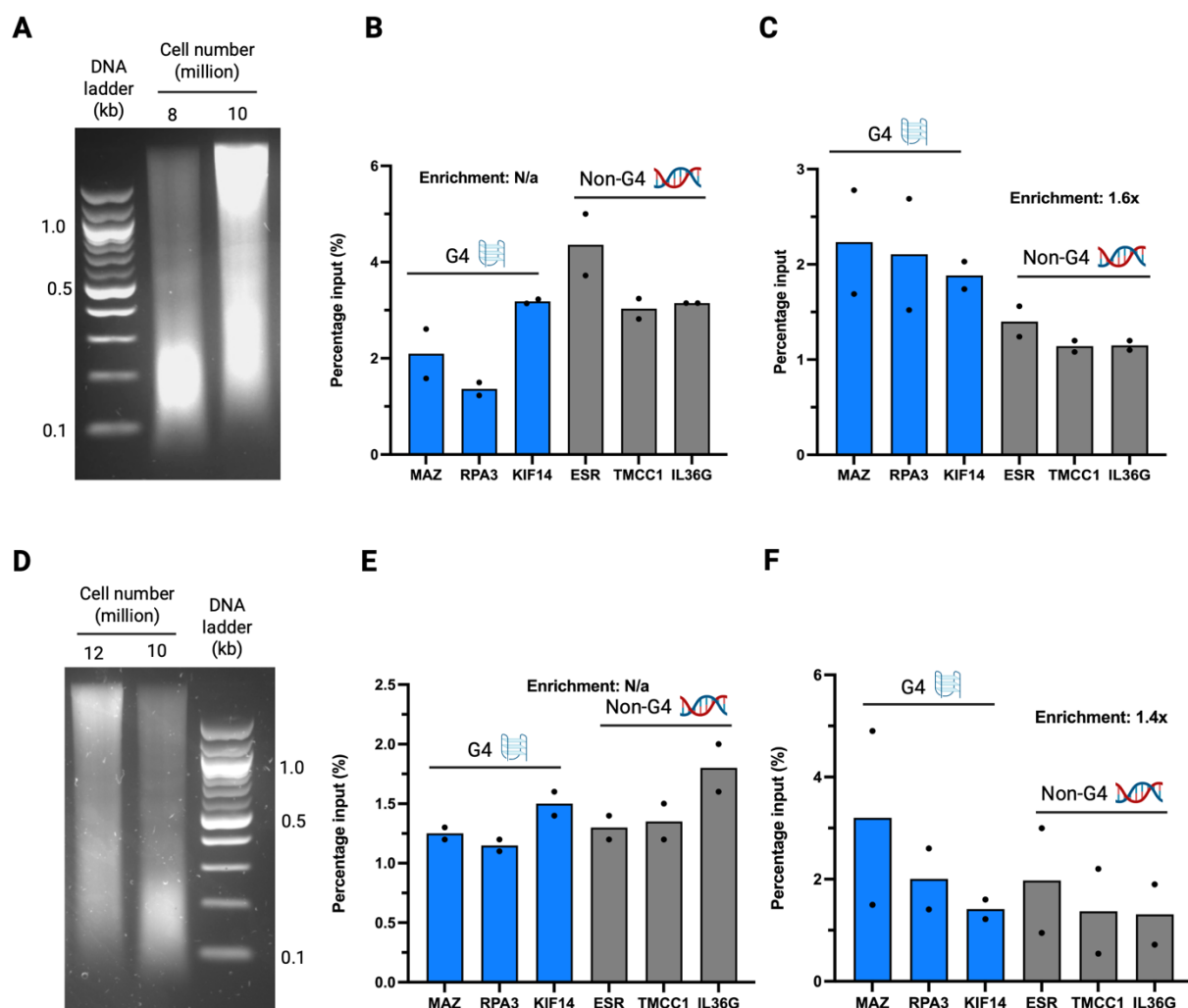


Figure 2.7 Unsuccessful ChIP-qPCR in PEO1 and PEO4 cells.

A) Agarose gel showing DNA fragment size distribution of 8 and 10 million PEO1 cells sonicated for 15 cycles. **B)** BG4 ChIP-qPCR of PEO1 chromatin from 8 million cells showing no enrichment of G4 regions (blue) and non-G4 regions (grey). **C)** BG4 ChIP-qPCR of PEO1 chromatin from 10 million cells showing 1.6-fold enrichment of G4 regions (blue) and non-G4 regions (grey). **D)** Agarose gel showing DNA fragment size distribution of 10 and 12 million PEO4 cells sonicated for 15 cycles. **E)** BG4 ChIP-qPCR of PEO4 chromatin from 12 million cells showing no enrichment of G4 regions (blue) and non-G4 regions (grey). **F)** BG4 ChIP-qPCR of PEO4 chromatin from 10 million cells showing 1.4-fold enrichment of G4 regions (blue) and non-G4 regions (grey). Error bars

represent standard deviation of two technical replicates. Percentage input values for qPCR plotted as an average of n=2 technical replicates, with individual values for each genomic region represented as a dot.

With experimental factors such as BG4, anti-FLAG beads, and buffers remaining constant between the successful and unsuccessful attempts, it was determined that the cause of the ineffectiveness of the ChIP must arise from the quality of the chromatin samples themselves. Upon further scrutiny, it was noticed that the qPCR profiles did not differ too significantly between the chromatin samples, regardless of whether the profiles looked optimal or substandard on the agarose gel. Thus, it was hypothesised that the chromatin sonication of these samples, which looked ideal on the gel, was in fact not as successful as it was initially thought to be. Indeed, there did appear to be a faint smear on the gel at fragment lengths above 600 bp. Initially, this was not thought to be an issue, as smears were also seen within this range for successful U2OS and PEO1 ChIP-qPCR samples. However, it was decided that use of an alternative method to visualise DNA fragment size distribution would help uncover what these smears on the gel meant for the distribution profile. During the project, an automated electrophoresis machine (Agilent TapeStation) was purchased for this purpose. This machine was used for sensitive quantification of DNA fragment distribution up to 5000 base pair. Re-analysing the chromatin fragments this way revealed some high base pair fragment tails were present, as depicted by the smear on the gel (**Figure 2.8 A-D**). Whilst these look negligible on the agarose gel, they become more apparent when quantified on the TapeStation. Subsequently, these high base pair tail regions have the propensity to hinder immunoprecipitation efficiency. These results revealed that ChIP efficiency is highly dependent on the chromatin sonication quality, and the presence of any high base pair fragments reduces success significantly. Thereafter, the TapeStation was used to assess chromatin sonication quality, owing to its enhanced sensitivity relative to an agarose gel.

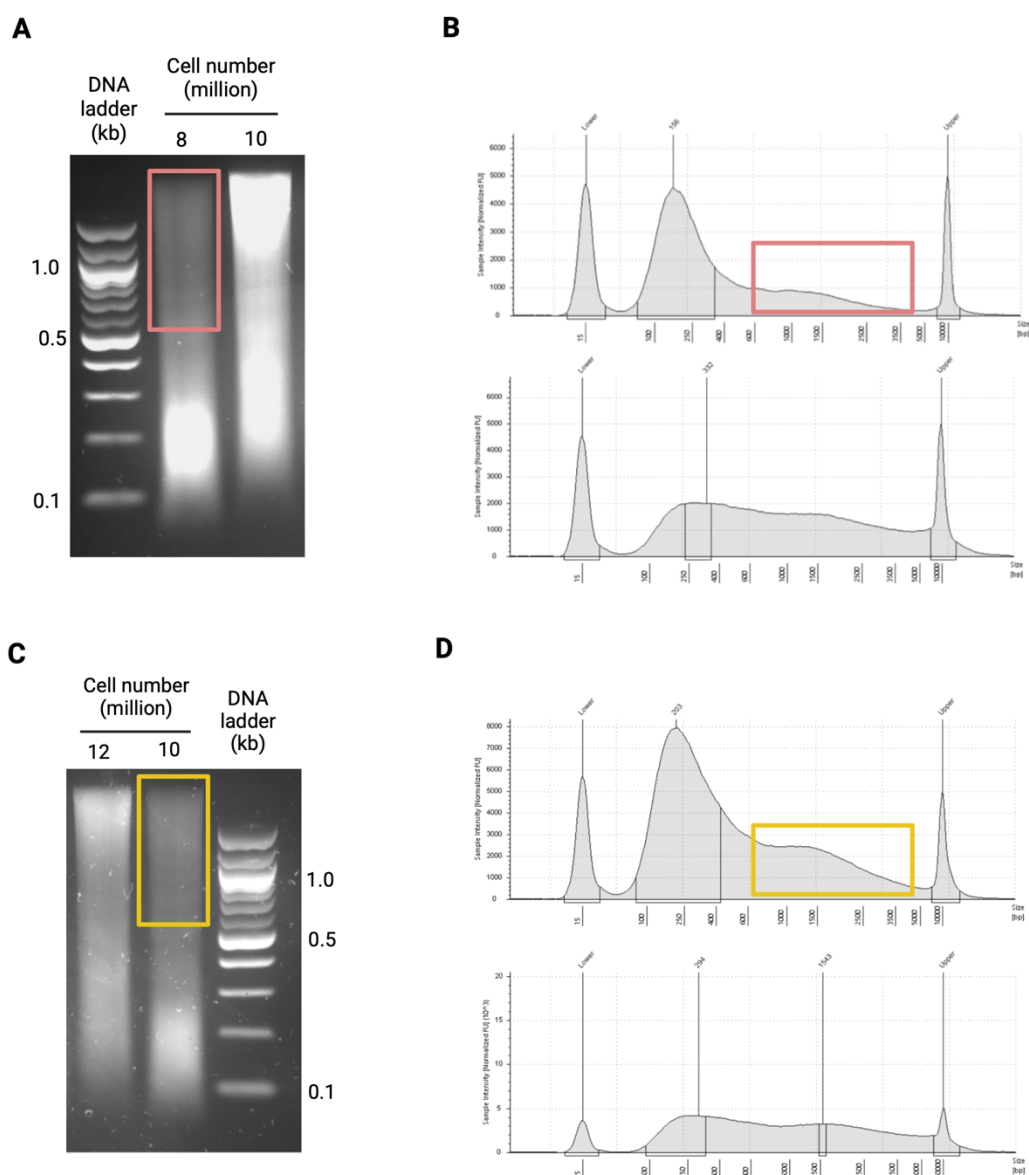


Figure 2.8 Fragment size distribution analysis.

A) Agarose gel showing DNA fragment size distribution of 8 and 10 million PEO1 cells sonicated for 15 cycles, with the smear highlighted by the box (salmon), and **B)** the corresponding TapeStation D5000 profiles for 8 million (top) and 10 million (bottom) PEO1 cells, with the section corresponding to the smear highlighted by the box (salmon). **C)** Agarose gel showing DNA fragment size distribution of 10 and 12 million PEO4 cells sonicated for 15 cycles, with the smear highlighted by the box (yellow), and **D)** the corresponding TapeStation D5000 profiles for 10 million (top) and 12 million (bottom) PEO4 cells, with the section corresponding to the smear highlighted by the box (yellow).

2.4.3. Successful generation of PEO1 and PEO4 ChIP-qPCR samples

Due to the higher base pair tails associated with 8 million PEO1 and 10 million PEO4 cells, the cells numbers for each cell line were reduced to adequately decrease the proportion of higher base pair fragments, by increasing sonication efficiency. Chromatin from 6 million PEO1 cells was cross-linked and sonicated for 15 cycles, at a concentration of 170 ng/ μ L. Quantification on the TapeStation revealed minimal fragments >600 bp, and this translated into a successful 8.6-fold enrichment of G4 regions over non-G4 regions in the ChIP-qPCR (**Figure 2.9 A&B**). This demonstrated that the higher base pair fragments were indeed the cause of the ChIP inefficiency, and that a reduction in cell number to facilitate more thorough sonication to within a 100-600 bp range could overcome this issue. Furthermore, the cell number reduction to 8 million PEO4 cells yielded the same success. The high base pair tail region on the TapeStation profile was reduced, and a 6.8-fold enrichment of G4-regions over non-G4 regions was obtained from ChIP-qPCR (**Figure 2.9 C&D**). The chromatin concentration was also a respectable 238 ng/ μ L. These parameters thus represent the ideal chromatin preparation conditions for PEO1 and PEO4 cells, to obtain sufficiently sonicated chromatin at an optimal concentration.

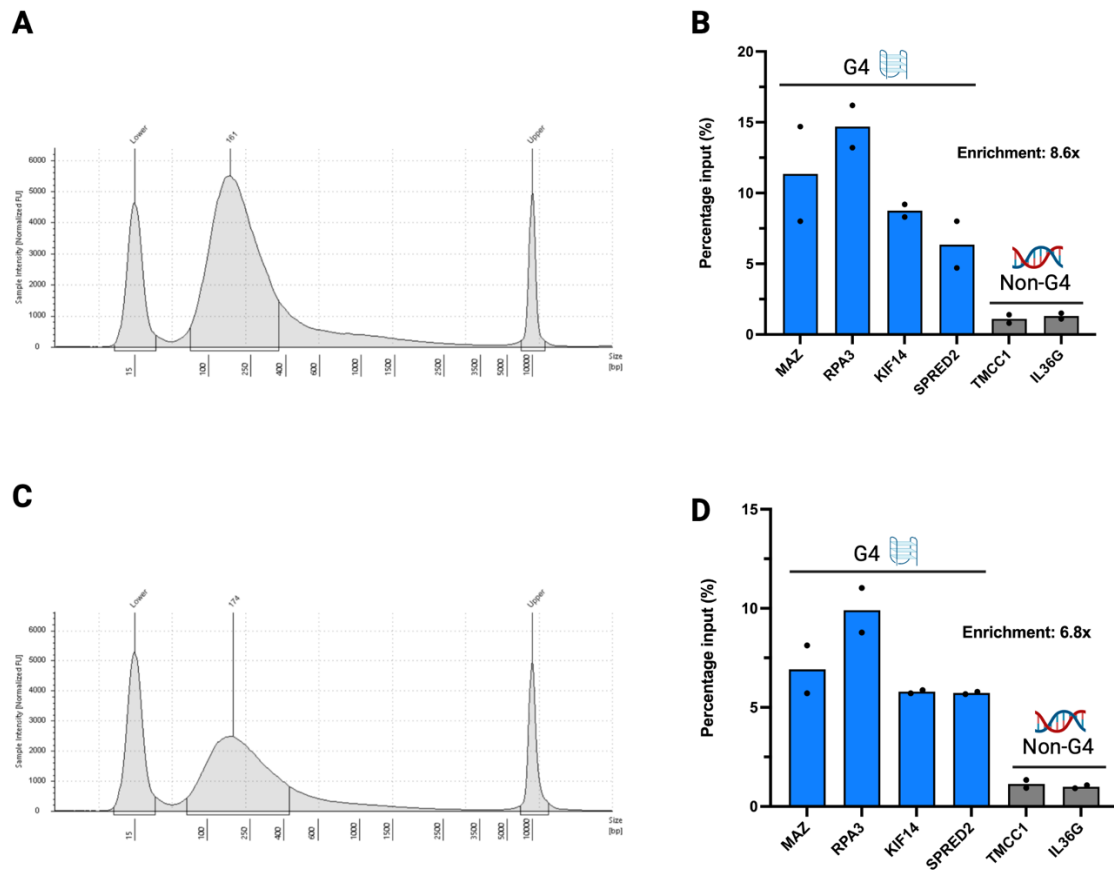


Figure 2.9 Results of ChIP-qPCR for PEO1 and PEO4 cells.

A) TapeStation D5000 profile of 6 million PEO1 cells after 15 cycles of sonication, and **B)** BG4 ChIP-qPCR of the sonicated PEO1 chromatin 8.6-fold enrichment of G4 regions (blue) and non-G4 regions (grey). **C)** TapeStation D5000 profile of 8 million PEO4 cells after 15 cycles of sonication, and **D)** BG4 ChIP-qPCR of the sonicated PEO4 chromatin 6.8-fold enrichment of G4 regions (blue) and non-G4 regions (grey). Percentage input values for the qPCR plotted as an average of n=2 technical replicates, with individual values for each genomic region represented as a dot.

Once these conditions had been established, ChIP-qPCR was repeated in PEO1 and PEO4 to obtain three biological replicates per cell line. The ChIP samples selected for next generation sequencing are detailed in

Table 2.1. All samples that were produced passed the threshold of >5% of G4 regions retained after immunoprecipitation. Additionally, all samples bar one passed the 5-fold enrichment criteria, with the exception being the original PEO1 sample generated at 4.9-fold enrichment. Overall, these samples were deemed suitable for sequencing and sent to The Francis Crick Institute for library preparation and Illumina paired end sequencing.

Cell line	Average % input G4 regions	Average % input non-G4 regions	Enrichment
PEO1	10.3	1.2	8.6x
PEO1	7.2	1.5	4.9x
PEO1	11	1.4	7.9x
PEO4	25.9	4.4	5.9x
PEO4	7.1	1.1	6.5x
PEO4	13	2.2	5.9x

Table 2.1 Table of three biological repeats of ChIP-qPCR PEO1 and PEO4 cells. Average % input is an average % input of all values calculated for primers for either G4 or non-G4 regions and from two technical replicates, all samples were sent for NGS.

2.5. Bioinformatic analysis of G4 ChIP-seq data

As previously discussed, successfully enriched G4 ChIP samples of PEO1 and PEO4 cells underwent Illumina NGS at The Francis Crick Institute. DNA fragments were attached to an Illumina flow cell via artificial adapter sequences, that were incorporated into the DNA strands via PCR during library preparation. Once NGS was performed, the output file in *fastq* format contained the detected nucleotide sequences, and the reads associated quality scores. However, in this format, the files could not be used to assess where in the genome the G4s were derived from. They first had to undergo a series of bioinformatical processes to match a sequence to a reference genome..

In summary, raw data (*fastq* files from NGS) was first assessed for its quality, and artificial sequencing adaptors removed. Reads were then aligned to a reference human genome, to identify where they are derived from. A further quality control step was undertaken post alignment, to remove sequencing artefacts and sort the aligned reads. Peaks were then called, which are regions of high read build up. These peaks were deemed to be sites of G4 immunoprecipitation, and peaks present in 2/3 biological repeats were taken forward for downstream analysis. The genomic locations of these peaks were determined via genomic annotation. The aim was to identify regions of high read build up in the human genome, that are classed as sites of G4 formation. From here, peaks identified in PEO4 and not in PEO1 were classed as a peak gain, and the converse a peak loss. It was hypothesised that the locations of these differences in G4s could be linked to genes and pathways associated with acquired ovarian cancer drug resistance, and this concept was probed using pathway enrichment analysis.

2.5.1. Assessment of raw data quality

When a base is called in NGS, a Phred score (known as a quality score (Q)), is associated with the detected base that indicates the probability of an error in the call. The Q score ranges from 0-40, with a Q10 meaning there is a 1 in 10 chance of an incorrect base being called, or a 90% chance that the base called is correct, whilst a Q30 score means there is a 1 in 1,000 chance of error, or 99.9% chance of a correct base being called.(201) To assess base call confidence of ChIP data, the genomics tool *fastqc* was used.(202) This tool gives an overview into the quality metrics of raw NGS data prior to any processing, and can give a visual breakdown of any issues that may have arisen in the NGS pipeline. The Q scores associated with the ChIP data were >35 on average, indicating high quality data sets (**Figure 2.10 A**). Additionally, the quality of the base calling did not deteriorate significantly the further along the read the base calling progressed, further corroborating the idea of highly reliable sequences being called (**Figure 2.10 B**). The *fastqc* tool also identified the presence of artificial adaptor sequences within the reads. In all samples, 5-15% of the reads contained some portion of the Illumina Universal Adaptor (**Figure 2.10 C**). Due to the artificial nature of these sequences, it is standard practice to remove these portions of the reads, so they don't hinder read alignment later in the pipeline. To this

end, the genomics tool *fastp* was used to trim the adaptor sequences from the genomic reads.(203) Moreover, *fastp* was adapted to simultaneously remove any reads <20 bp in length, as these have a higher probability of being sequencing artefacts, and are less likely to map uniquely to one genomic location.(203) After implementing *fastp*, *fastqc* was re-run on the trimmed datasets to confirm successful adaptor contamination removal (**Figure 2.10 D**).

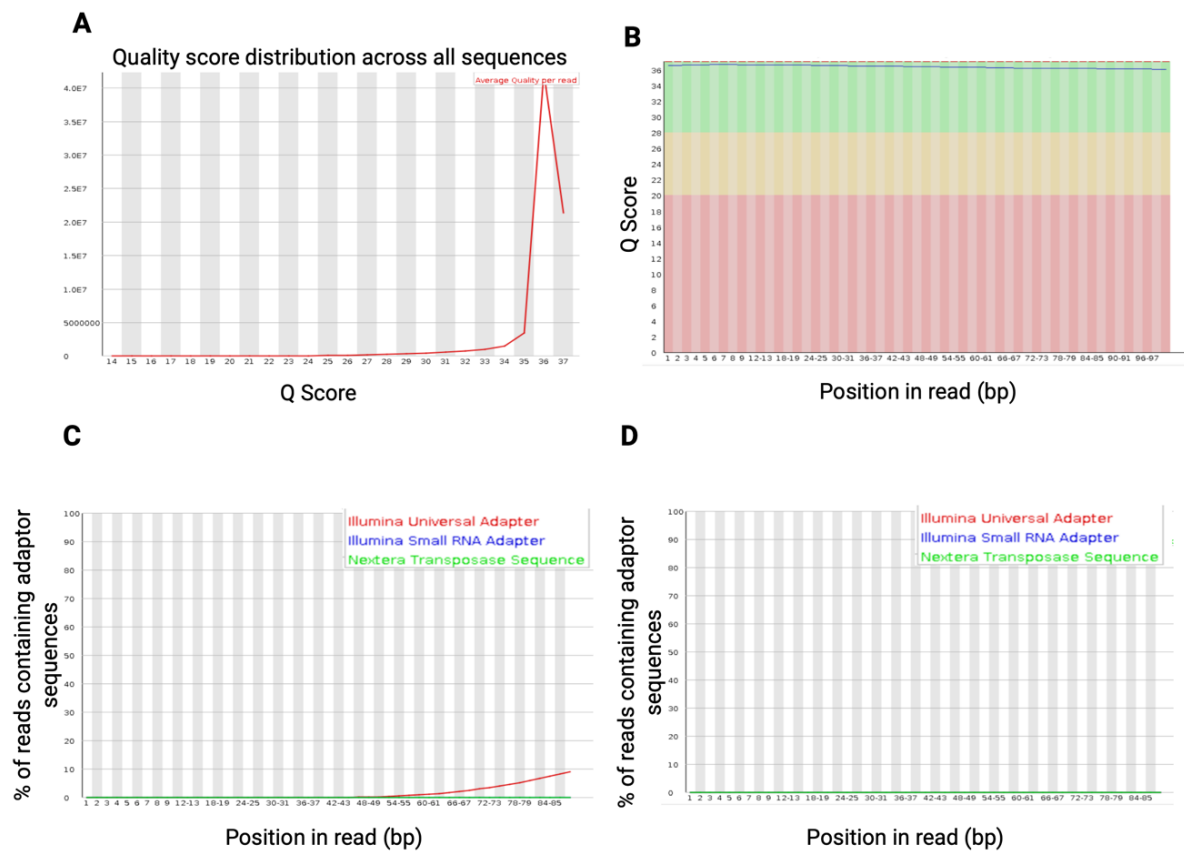


Figure 2.10 Quality assessment of ChIP reads.

Example *fastqc* outputs from PE04 G4 ChIP Biological replicate 1, Read 1. **A)** Distribution of mean Q Score across all reads, showing most reads have an average of Q36. **B)** The mean quality score against the position in the read the base was called. **C)** The percentage of reads containing contaminating Illumina universal adapter sequences, with 10% of reads in this sample containing adapter sequences. **D)** Percentage of reads containing contaminating Illumina Universal adapter sequences after trimming using *fastp*, showing complete removal of adapter sequences.

2.5.2. Read alignment

After the pre-processing and quality assessment was completed, the read's genomic locations were determined via alignment to a reference human genome. The hg38 reference genome was used to this end, as it is the most up to date genome with the most genomic coverage, compared to its predecessors.(204) *Bowtie2* was then employed to first index the hg38 genome, before aligning the reads.(205) As the reads were sequenced in paired-end format, both the forward and reverse reads were aligned. There are multiple metrics to consider when assessing the quality of the read alignment using *bowtie2*. Firstly, each mapped read is assigned a mapping quality score, ranging from 0-40. The higher the score, the more similar the read is to the genomic sequence it has been mapped to. Additionally, mapping 'uniqueness' is also an important parameter. Theoretically, a read should map uniquely once in the human genome, as it is derived from a singular location. However, this is not always the case. Portions of the human genome are highly repetitive and reads may align more than once. In this instance, the alignment with the highest mapping score, i.e. with the highest sequence similarity, is determined to be the primary alignment, whilst all others are secondary alignments. Finally, as paired end reads have a standard orientation and distance for how the paired reads should align from each other, this metric is also assessed. This is termed properly paired or concordant alignment. Reads should align concordantly once, meaning the paired reads align in the correct orientation, at the expected distance apart, and only once throughout the human genome.

Upon aligning the reads using *bowtie2*, a summary text file was generated informing of overall alignment, multi-mapping, and concordant alignment. These summary files were used as a data quality control, to assess how successful the mapping was, and if there were any issues with multi-mapping that could arise from sequencing artefacts. For all ChIP samples, the overall alignment of the reads was high at >97% (**Table 2.2**). Additionally, the unique and concordant reads that aligned once was >74% for all samples. This is indicative of good sequencing libraries (>70%), containing minimal artefacts, resulting in successful alignment.(206) Moreover, when assessing the percentage of reads that were duplicated, they account for most reads mapped more than

once meaning that they have most likely arisen from amplification in the library preparation and sequencing stages.

Condition	Rep	Total no. of mapped reads	Overall read alignment	No. of uniquely mapped reads	Uniquely and concordantly aligned	Read duplication
PEO1- Input ChIP	1	43,213,890	99.15%	33,745,726	78.09%	16.23%
	2	40,911,389	99.01%	32,283,177	78.91%	11.96%
	3	41,092,857	98.95%	33,227,684	80.86%	12.37%
PEO1- G4 ChIP (tech reps merged)	1	74,917,590	98.02%	55,169,313	73.64%	20.80%
	2	78,323,547	98.58%	61,233,349	78.18%	28.43%
	3	72,986,658	98.61%	57,479,355	78.75%	23.71%
PEO4- Input ChIP	1	37,027,380	99.15%	28,825,815	77.85%	16.30%
	2	40,310,168	99.07%	31,514,489	78.18%	17.80%
	3	39,120,683	99.03%	30,537,605	78.06%	16.27%
PEO4- G4 ChIP (tech reps merged)	1	72,045,440	97.98%	53,731,489	74.58%	19.38%
	2	65,423,348	98.02%	48,419,819	74.55%	19.19%
	3	67,332,871	97.45%	49,833,057	74.01%	18.16%

Table 2.2 Alignment quality of ChIP samples.

Table of total number of reads, overall alignment rate, uniquely and concordantly aligned read number and percentage of total mapped reads and read duplication percentages for PEO1 and PEO4 input and G4 ChIP samples.

2.5.3. Post alignment read sorting and filtering

A second round of quality control checks and filtering were conducted post alignment. The purpose of this was to remove reads which may not be a true biological

representation, and lead to false positive peaks in the G4 ChIP datasets. For instance, reads were removed that had a low mapping quality score (<20), and secondary aligned reads. Additionally, reads mapped to “blacklisted” regions were removed. These regions are defined as unstructured, and of high signal, that have artificial read build-up.(207) As such, removal of these regions is standard practise for ChIP-seq analysis, to remove signal-artefact, and to reduce likelihood of false positives in the peak calling step. Finally, duplicate reads were removed from the data. As previously mentioned, these arise from PCR over-amplification during library preparation and sequencing. Consequently, these reads are artefacts, and not of biological relevance. Overall, post filtering, the aligned read files contained reads mapping uniquely once to non-blacklisted regions, of a mapping quality >20 , with no read duplicates. This gave confidence that the reads remaining were a true reflection of the biological landscape, and not artefacts generated from library preparation or sequencing.

2.5.4. Sample clustering

Similarities between samples were assessed using *deepTools plotCorrelation*, using the Spearman method. This was conducted to assess reproducibility between biological replicates and highlight for the first time if any differences occur between different cell lines prior to peak calling. The Spearman correlation coefficient uses read coverage of the aligned and filtered datasets to assess how similar they are to each other. The similarity of the samples is assigned a score from 0-1, with the closer to 1 meaning the more strongly correlated the samples are. Using this methodology, correlation between PEO1 and PEO4 G4 ChIP and input control samples was assessed. The correlation plot divides into four distinct quadrants, with the input and G4 ChIP samples being the least correlated to each other, with scores ranging from 0.5-0.7 (**Figure 2.11**). This is unsurprising, given the input is a genomic background control, whilst the G4 ChIP samples have undergone antibody immunoprecipitation. Hence, read coverage will differ due to read pile up in peaks in the G4 ChIP samples, but not in the input. The most strongly correlated quadrant was that associated with PEO1 and PEO4 G4 ChIP samples. A very strong correlation was identified between biological replicates of the same cell line (>0.97 PEO4, >0.96 PEO1), indicating high reproducibility of the immunoprecipitation,

and robust data sets. PEO1 and PEO4 G4 ChIP samples also correlated well to each other (0.88-0.91), underscoring the conserved similarities between the immunoprecipitated samples. However, the correlation score between PEO1 and PEO4 G4 ChIP samples was lower compared to that between biological repeats of the same cell line. This indicates that there are still differences between the immunoprecipitated samples of the different cell lines, due to cell-type specific factors. Overall, the correlation plot of ChIP samples shows high reproducibility between biological repeats of the same cell line, whilst indicating that there are some differences between the PEO1 and PEO4 cell lines. This result provides the first evidence that the G4 landscape could differ between these cell lines.

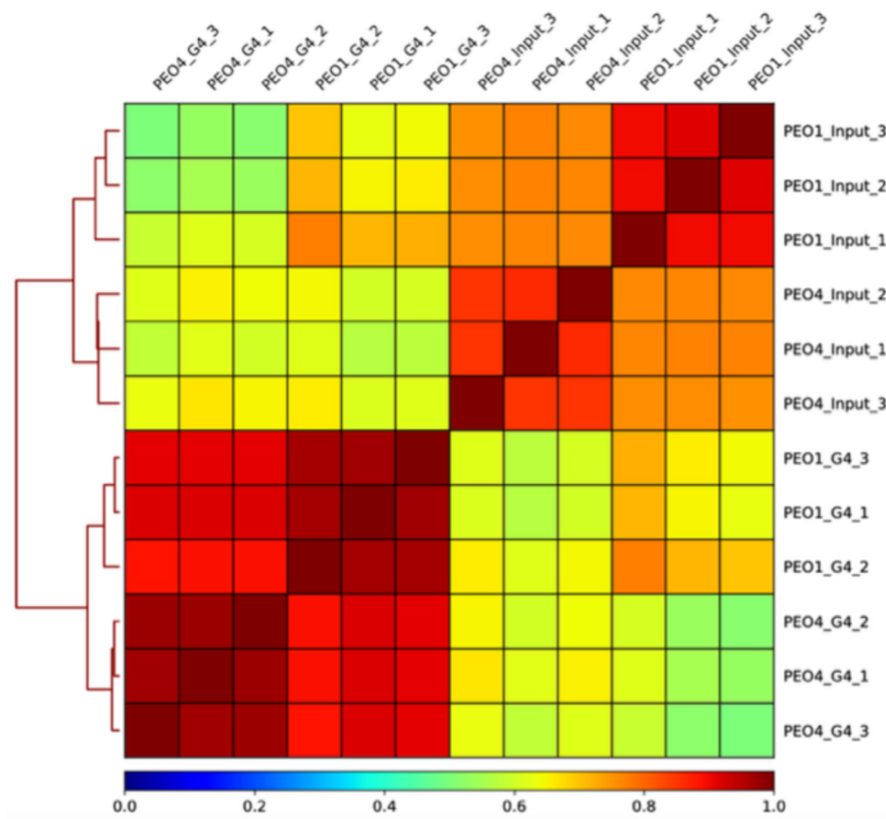


Figure 2.11 Correlation plot of ChIP aligned reads.

Heatmap Spearman correlation plots of PEO1 and PEO4 G4 ChIP and input control samples. Created using *deepTools plotCorrelation*.

2.5.5. Peak calling

To ascertain regions of high read build up, that likely correspond to G4s presence, Model-based Analysis of ChIP-seq 2 (macs2) was used to define these regions as 'peaks'.(208) Macs2 operates through scanning the genome and identifying read-enriched regions, that arise through immunoprecipitation. This can either be conducted via normalisation to the input control, that accounts for background genomic signal in the instance of ChIP-seq, or in the absence of this control. In short, 1,000 regions of high amplitude are located, and assigned as a peak at the mid-point between the forward and reverse reads. From these high-quality peaks, a peak calling model is generated specifically for the dataset and applied to the rest of the reads to detect peaks. When an input file is present, the regions of high read density are identified relative to the read level in the input file at that particular region.

Peaks were called in this way for G4 ChIP samples for PEO1 and PEO4 and visualised using IGV.(194) As a peak calling quality control, RPA3 and IL36G regions were visualised, due to their previous use as G4 forming and non-G4 forming regions in the initial qPCR validation of ChIP success prior to NGS. As expected, in the RPA3 promoter site of all biological repeats there was a distinct region of read enrichment that was determined by macs2 to be a G4 peak (**Figure 2.12 A**). This corresponds to the area targeted by the primers in ChIP-qPCR, which consistently showed enrichment over non-G4 forming regions. Conversely, when assessing the IL36G region, there was no discernible read build up, relative to the input samples, leading to no peak detection in this location (**Figure 2.12 B**). These results were encouraging that the peak calling method selected was successful at identifying true G4 peaks, as they replicated what had been observed in the ChIP-qPCR validation step, and showed formation in previously validated G4 sites.(99)

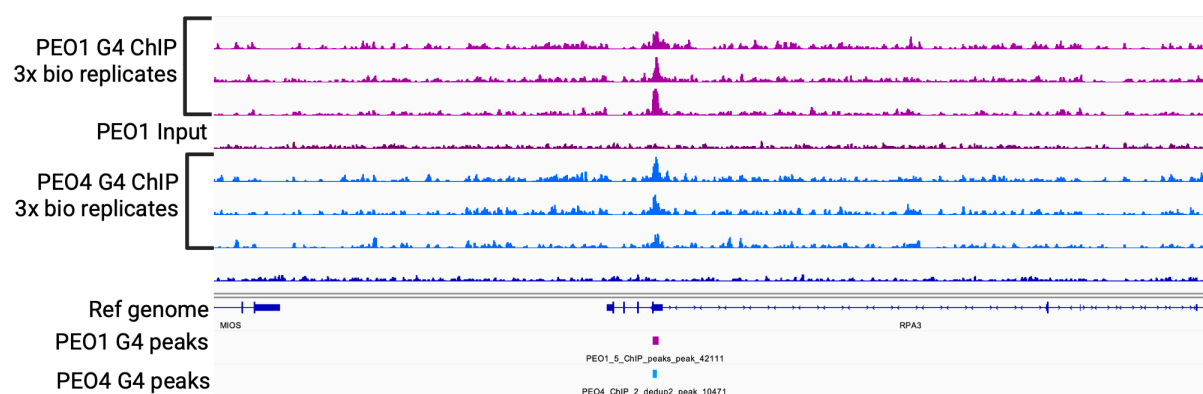
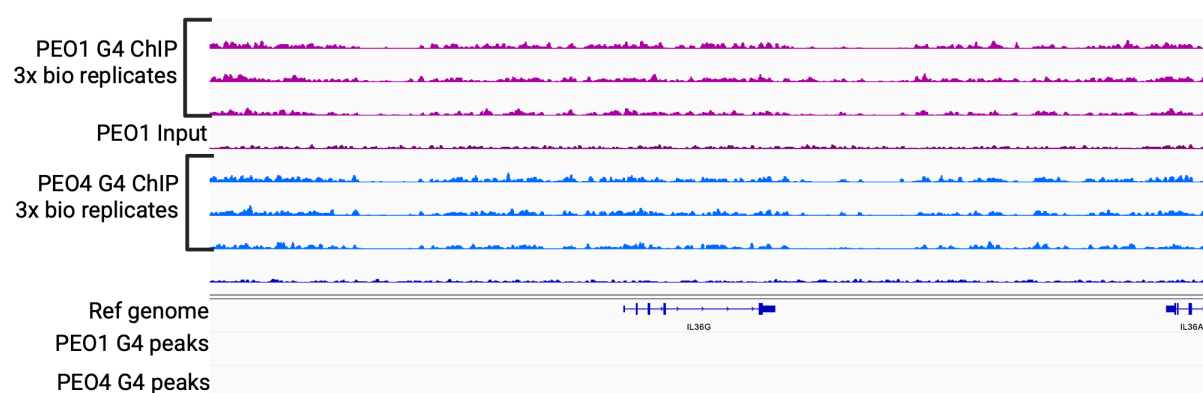
A**B**

Figure 2.12 IGV genomic tracks and G4 peaks in PEO1 and PEO4 cells.

Regions shown are of **A)** the RPA3 G4 forming region for all 3 ChIP-seq sample replicates, with a G4 peak in both cell lines, **B)** the IL36G non-G4 forming region for all 3 ChIP-seq sample replicates, with no G4 peak in either cell line.

To obtain finalised peak maps in the PEO1 and PEO4 cell lines, peaks were identified that were present in 2/3 of the biological replicates. This was done to increase the reliability of the peak calls based on a simple majority principle that >50% of the samples identify a peak. It enables a reduction in false positive peaks, that may occur through technical variation in a singular sample and allows for more reproducible G4 maps. 2/3 was also chosen over 1/3 or 3/3 as it attempts to balance the ratio of false positive and false negative peaks that could occur. For instance, if only peaks present in all three replicates are considered, this may result in missing true G4 peaks that could not be called in one

replicate yet called in the others. The reason for this could be due to technical error in the ChIP, or the enrichment not quite reaching the threshold to be called. Conversely, if peaks are considered that are just in 1/3, then this may result in false positive G4 peaks being called as some peaks may be technical artefact or be false positive peaks anyway as the false discovery rate is at 5%. Therefore, by considering those that are in 2/3 replicates, it reduces the chance of technical artefact peaks being considered yet does not miss the peaks which may be lacking in one sample due to technical variation. The intersection is done by bedtools intersect, which counts a peak as shared if even 1 bp overlaps between the two.

For the PEO1 cells, 12,984 G4 peaks were identified in this manner, with a relatively high variation in peak numbers detected between individual biological replicates (**Figure 2.13A**). This could be due to the variation in enrichments achieved for each biological repeat, ranging from 4.9-fold to 8.6-fold. Lower enrichment generally means a reduction in the number of peaks identified. This is due to an increase in false negative peaks, as read build up in true G4 regions may not surpass the enrichment threshold, relative to input, to qualify as a peak. In a more enriched sample, these sites are more likely to be detected, leading to a higher peak count. Regardless, the finalised G4 peak map was generated from 2/3 biological repeats, and so the variation between the samples is negated to an extent. There was less variation between PEO4 G4 peaks detected across the three biological repeats, with 6,392 peaks identified in 2/3 replicates (**Figure 2.13A**). This could be due to the limited variation between the enrichment of the three samples (5.9-6.5-fold), relative to within the PEO1 samples. Interestingly, there was over half the number of G4s detected in PEO4 cells than in PEO1. This could be due to the technicality of lower enrichment levels achieved in the PEO4 ChIP samples, or could be a true biological difference. To probe this idea, additional PEO4 G4 ChIP samples could be obtained at a higher fold-enrichment to assess whether this impacts the ratio of peaks between PEO1 and PEO4 cells. As this study aimed to assess epigenetic differences between PEO1 and PEO4 cells that could lead to acquired drug resistance, unique G4 peaks to each cell line were identified. This was achieved using bedtools intersect to compare the peak files and assess overlap. Here, the peaks are defined as shared if they overlap if even 1 bp overlaps between the two peaks. The peaks are classed as unique to

one cell line or the other if 0 bp overlap between the two peaks, however that does not mean that two peaks may not still be in close proximity. PEO4 cells had 1,934 unique peaks, termed gained G4s (**Figure 2.13 B**). Conversely, PEO1 cells had 8,526 unique G4 peaks, which were in turn lost G4s in the PEO4 cells. There were 4,458 G4 peaks conserved between the two cell lines. Indeed, this provided evidence that the G4 landscape does vary between PEO1 and PEO4 cells and could potentially be associated with the development of drug resistance, given that is the major variable and adaptation between the two cell lines.

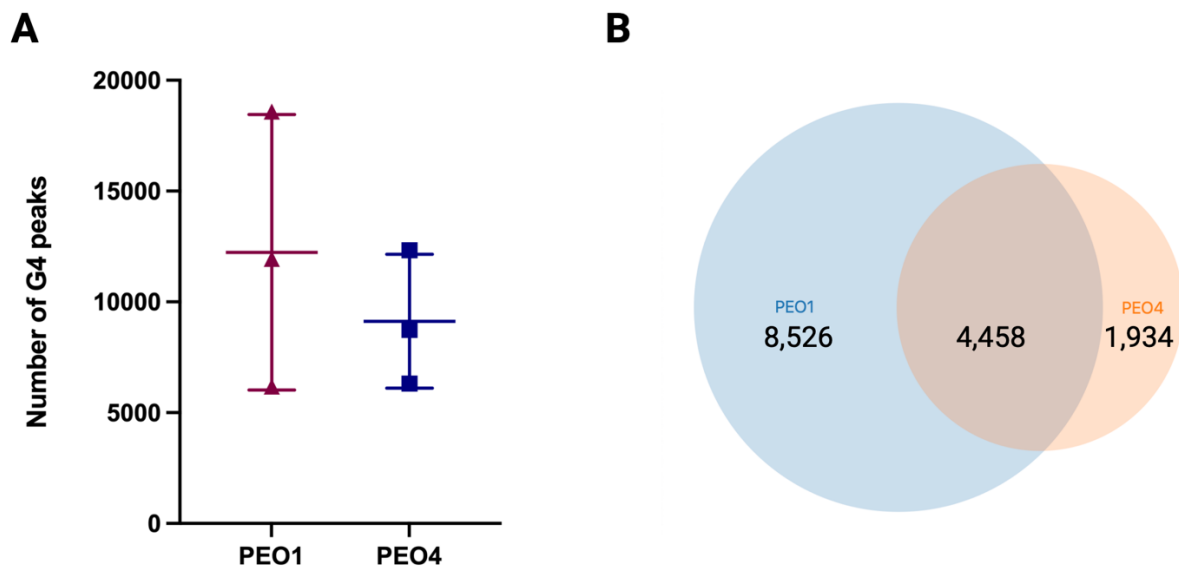


Figure 2.13 G4 peaks in PEO1 and PEO4 cells.

A) G4 peaks called in all 3 biological replicates, **B)** G4 peaks unique to PEO1, PEO4, or conserved between the two cell lines. Made with BioInfoRx.

To further validate the legitimacy of the generated G4 maps, two additional quality control checks were performed. Firstly, G4 peaks were intersected with an *in vitro* generated G4 map, by G4-Seq.(94)(95) These maps contain observed quadruplexes (OQS), that were formed from isolated DNA and give a good indication into sites that have the propensity to form G4s. In all cases, 70-80% of G4 peaks detected through G4 ChIP-

seq overlapped with OQS sites, in line with previous studies (**Figure 2.14 A**).⁽⁹⁹⁾ This result gave confidence that G4s had been successfully immunoprecipitated, and bioinformatically detected, due to the high overlap with *in vitro* detected G4 sites.

In addition, the most enriched sequence motif detected from the PEO1 and PEO4 peaks were identified using MEME suite.⁽²⁰⁹⁾ For both cell lines, the most enriched motif was a highly G-rich sequence, found in >98% of peaks (**Figure 2.14 B**). This was encouraging, given that G4s form from G-rich sequences, and further validated the successful detection of G4 structures in PEO1 and PEO4 cells.

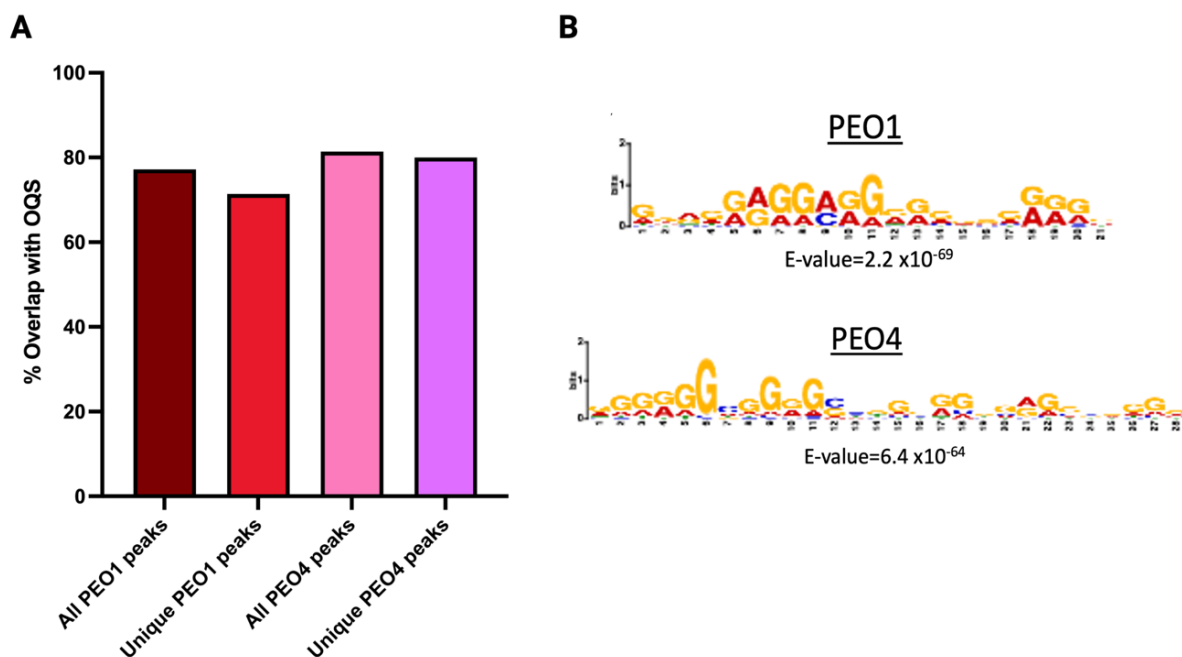


Figure 2.14 Validation of G4 maps in PEO1 and PEO4 cells.

A) Percentage overlap of G4 peaks with OQS sites for all PEO1 and PEO4 G4 peaks and those unique to either cell line and, **B)** the most enriched motif in the PEO1 (top) and PEO4 (bottom) G4 peaks. Generated by *MEME suite*.

2.5.6. Genomic distribution of G4 peaks

Once the G4 peak maps had been established, the genomic locations of G4s were determined using *homer annotatePeaks*. This function associates the peak with its nearest gene and annotates the genomic location it is found. Of note, these locations include promoters (-1 kb to +100bp of transcription start site), introns (non-coding section of the gene which is initially transcribed before being removed via splicing prior to translation), and intergenic regions (non-coding regions that are located between genes). The most enriched site of G4 formation in PEO1 and PEO4 cells was at gene promoters- at 34% and 59%, respectively (**Figure 2.15**). This is in agreement with previous studies that report there is an enrichment of G4s at promoter regions, which are key regulatory sites within cells.(97)(184)(135)(102) Nonetheless, when considering unique G4 peaks to either PEO1 or PEO4 cells, this promoter localisation dropped by 10%. In correspondence, there was an increase in localisation at intron and intergenic loci. This is of interest, as non-promoter regions have been highlighted as being of importance to development of drug resistance in ovarian cancer.(80)(78) This finding indicates that changes in G4 abundance also occurs in these regions. It is worth highlighting that the genomic distribution of G4s is disproportional to the abundance of these genomic loci. Intergenic regions make up around 50% of the human genome, whilst it is thought promoters account for <1% of the genome.(210)

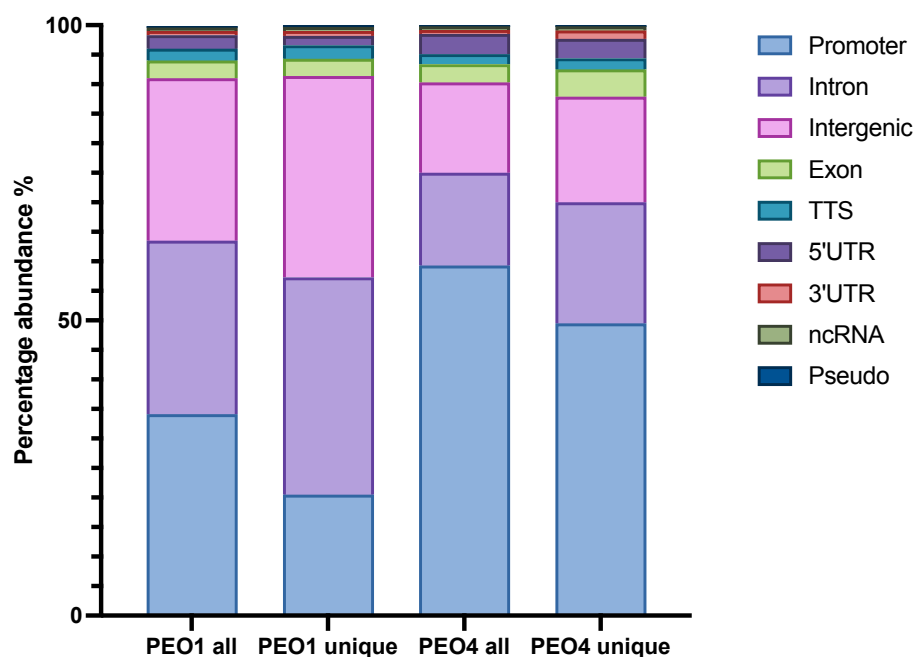


Figure 2.15 Genomic location distribution of PEO1 and PEO4 G4 peaks.

The most enriched site of G4 formation is in promoter regions for both PEO1 and PEO4, yet this localisation decreases by 10% when considering peaks unique to either PEO1 or PEO4.

2.5.7. Pathway enrichment of G4 distribution

To assess whether G4 gain or loss occurred randomly across of the genome, or if it is enriched in genes associated with specific biological pathways, pathway enrichment analysis was conducted using *ShinyGO*.⁽²¹¹⁾ This tool evaluates a list of genes against a background gene set and identifies whether the gene list is overrepresented in any KEGG pathways. This analysis was conducted to provide the first insight into which biological pathways gained G4s may be contributing to in PEO4 cells. Whilst it does not consider the transcriptome and gene expression changes, it offered preliminary insight into which biological pathways G4 formation could be associated with. For both G4 gain and loss in PEO4 cells, KEGG pathways are enriched for these gene sets (**Figure 2.16 A&B**). Notably, pathways such as mTOR and Hippo signalling are enriched amongst gained G4s, and play roles in tumorigenesis which contributes to cancer progression and drug resistance in ovarian cancer.⁽²¹²⁾⁽⁴⁶⁾⁽²¹³⁾ Conversely, MAPK, Rap1 and Ras signalling pathways are enriched amongst the genes linked with a loss of G4 in PEO4 cells, which have also linked with ovarian cancer progression and therapy resistance.⁽²¹⁾⁽²¹⁴⁾⁽²¹⁵⁾ Therefore, it was deemed that to make meaningful observations into the impact of gain or loss of G4s in PEO4 cells, association with gene expression changes would first need to be conducted. This is because the gain or loss of G4 may be passive and not correlated to an increase or decrease in gene expression levels that contribute to acquired drug resistance. Nonetheless, the enrichment of KEGG pathways was encouraging that G4 formation was not necessarily a randomly occurring event in PEO4 cells and instead localised in specific pathways that have associations to drug resistance.

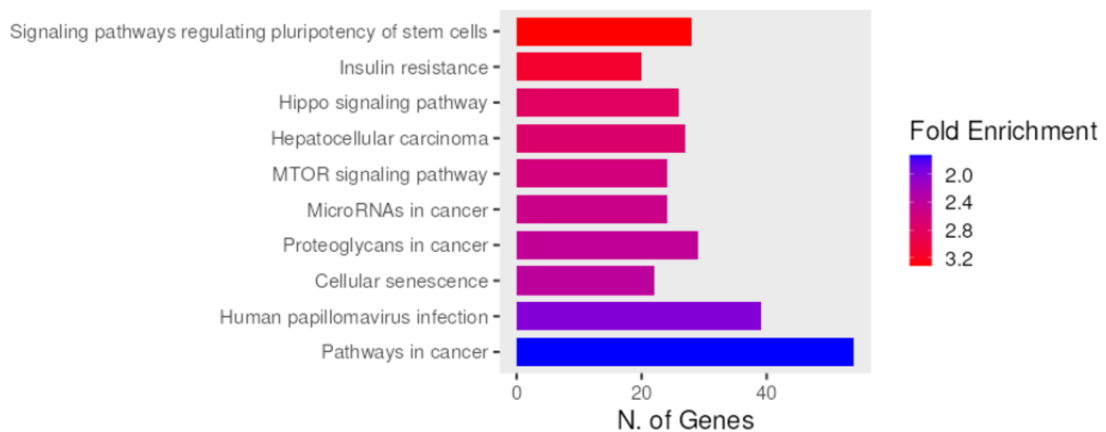
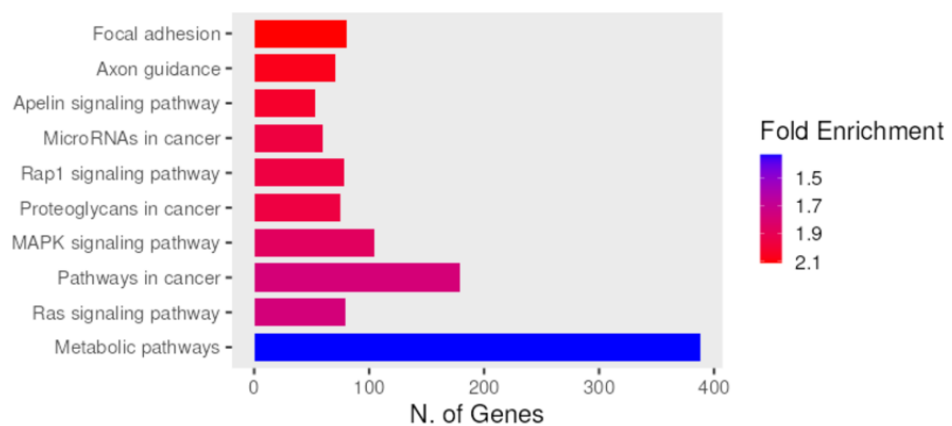
A**B**

Figure 2.16 Pathway enrichment of G4-related genes in PEO4 cells.

KEGG pathway enrichment of genes associated with **A)** a gain in G4 or, **B)** a loss of G4 in PEO4 cells. The number of genes enriched in each pathway and the fold enrichment of the pathway are indicated. The fold enrichment is a measure of how the percentage of genes within gene list present in the pathway divided by the percentage in the background. Calculated and made with *ShinyGO*.(211)

2.6. Chapter summary and conclusions

To summarise, G4 maps were successfully generated in PEO1 and PEO4 cells for the first time. This was achieved through expression, purification and validation of a G4 selective antibody for use in the immunoprecipitation of G4 structures. Upon successful immunoprecipitation, and validation by qPCR, samples were sequenced and bioinformatically processed to identify regions of G4 formation. Differences in the G4 landscape were detected between PEO1 and PEO4 cells, showing for the first time that variation in G4 formation can occur between a drug-sensitive and drug-resistant cell lines, as the cells evolve. Notably, promoter loci were the most prominent location for G4 formation, accounting for 50% of new G4 peaks. However, 40% of new PEO4 peaks were attributed to intron and intergenic regions, which are of known importance in acquired drug resistance in ovarian cancer cells. Assessment into the pathways that G4s form in PEO4 cells unveiled enriched pathways previously reported for their involvement in acquired drug resistance. This provided the first insight into the potential association between G4 formation and drug resistance. However, this was a correlative preliminary finding and not directly linked to gene expression changes that could be associated to acquired drug resistance. Hence, transcriptomic data needed to be generated to investigate how changes in G4 landscape might be linked to gene expression changes that could contribute to acquired drug resistance in ovarian cancer cells.

It is important to acknowledge the limitations associated with the generation of endogenous G4 maps using ChIP-seq. Firstly, it is difficult to gain a true value of G4 presence within the cells, owing to the enrichment required in order to detect peaks. As stated previously, samples with a lower enrichment of G4 regions over non-G4 regions identified via qPCR will tend to have less G4 peaks detected than those with a higher enrichment. Of course, the enrichment of different regions will also differ to the general enrichment ascertained by qPCR of specific regions, and so the general enrichment value may not be a solid metric of how successfully immunoprecipitated that sample is. To circumvent this, samples produced should ideally have similar enrichment to negate inter-sample technical variation, with as high an enrichment as possible. Furthermore, ChIP is heavily reliant on optimal sonication to ensure successful target immunoprecipitation. This is a confounding factor that requires careful consideration

when preparing samples and is most certainly a limitation of the method. Alternatively, validating the G4 maps using another G4 mapping method that does not rely on enrichment or sonication could be a useful progression to add confidence that the generated G4 maps are true representations of the G4 landscape within cells. This could be achieved via G4 CUT&Tag, which tags the DNA preferentially at G4 sites after BG4 binding and allows for specific enrichment via qPCR of G4 regions. Moreover, there may be bias introduced by only using BG4 to target the G4 structures. Whilst a highly selective antibody, using another probe-based method such as Chem-Map would reduce any probe-based bias for G4 detection. Another confounder in G4 map generation is the potential for artefact enrichment in the library preparation PCR stage. For instance, G-rich regions tend to be preferentially amplified during PCR. Hence, ensuring a suitable control (such as an input sample in ChIP, or a non-targeting control for CUT&Tag) is key to ensure the G4 peaks detected are as a result of specific targeting, and not biased PCR amplification. In the instance of this project, this was not an issue due to stable input files that had no preferential amplification at G-rich G4 sites that could result in false positives, but it is worth being aware of for future studies.

3. Intersection with transcriptomic and accessibility data

3.1. Chapter overview

Once differences in the G4 landscape had been identified between PEO1 and PEO4 cells, the transcriptional differences between the two also needed to be ascertained. This would allow investigation into whether G4 distribution changes were associated with gene expression changes reported to contribute to acquired drug resistance in ovarian cancer, or whether they were merely passively forming with no transcriptional impact. To be able to determine the relative transcriptional impact of G4 formation in PEO4 cells, chromatin accessible sites were also profiled, as changes in chromatin accessibility are of high importance in the epigenetic regulation of drug-resistant ovarian cancer. Hence, the transcriptional effect of G4 formation or loss was compared to a gain or loss in chromatin accessible site.⁽⁸⁰⁾ The transcriptional impact of altered G4 presence was analysed globally, and in three genomic locations: promoters, intronic and intergenic regions. Whilst the primary focus of previous studies has been on the transcriptional impact of G4 formation in promoter regions ⁽¹⁰¹⁾⁽⁹⁷⁾⁽¹³⁵⁾⁽⁸⁷⁾⁽²¹⁶⁾, non-promoter regions were also considered in this project due to the relevance of these sites in ovarian cancer drug resistance.⁽⁸⁰⁾⁽¹⁸⁶⁾⁽⁷⁸⁾ In addition, intergenic regions are of interest as that is where enhancer elements generally reside. Enhancers are regulatory elements with the ability to regulate transcription of a gene from thousands of base pairs away.⁽²¹⁷⁾ Moreover, it has recently been discovered in the Di Antonio group that G4's have a high overlap with the locations of transcriptional enhancer sites, which causes more potent transcriptional upregulation than either a lone enhancer or G4.⁽⁸⁸⁾ This provides evidence that intergenic G4 sites may be of interest in the context of transcriptional regulation when considering the acquisition of resistance in HGSOC, especially considering the role enhancer sites play in regulating gene expression in this context. Pathway enrichment was conducted to investigate the biological processes associated with significantly upregulated genes correlated with G4 formation or accessibility gain. This would provide

a more comprehensive insight into the effect of G4 formation in PEO4 cells, as it accounts for the changes in gene expression which may contribute to the development of drug resistance.

3.2. The transcriptional profiles of PEO1 and PEO4 cells

To assess transcriptional differences between PEO1 and PEO4 cells, and associate this with G4 distribution, RNA was extracted from both cell lines for NGS. The differences in RNA abundance between the two cell lines gives a quantitative measure of gene expression activity, which can correspond to unique cellular behaviours between PEO1 and PEO4. Subsequently, this data can be used to probe how formation or loss of G4s is linked to gene expression changes between the two cell lines, and what biological function these genes are linked to. Total RNA was extracted from PEO1 and PEO4 cells using the Qiagen RNeasy Mini kit, with an additional on column DNase digestion step to remove contaminating DNA that can interfere with RNA sequencing. As the primary aim of this experiment was to assess differentially expressed genes between PEO1 and PEO4, the library preparation (conducted by Novogene) enriched for messenger RNA.

3.2.1. Assessment of RNA integrity

RNA was extracted from 3.5 million cells, per biological repeat, per cell line. A total of three biological replicates were prepared for both PEO1 and PEO4. Once the RNA was extracted, its integrity was assessed using the TapeStation. This acts as a quality control metric for how intact or degraded the RNA sample is. The software assigns the sample an RNA Integrity Number equivalent (RINe) that ranges from 1-10, with 1 corresponding to a degraded sample and 10 to an intact sample. This is calculated through evaluating the ratio between the 28S/18S rRNA present in the total RNA sample, with a shift towards the 18S peak indicating partial degradation. A value above 7 is considered to be acceptable for next generation sequencing, as advised by Illumina and Agilent. This means that the RNA is minimally degraded, and most of the original transcripts are intact to be accurately sequenced, to capture the full transcriptome. All RNA samples extracted

from PEO1 and PEO4 cells had an RINe score >9.5, indicating the successful extraction of RNA, with minimal degradation. The profiles obtained from the TapeStation show a clear distinction between the 18S and 28S rRNA peaks, with the 28S peak being the predominant one, highlighting the lack of degradation (**Figure 3.1 A&B**). As all samples surpassed the RINe threshold value of 7, they were sent for mRNA library preparation and paired end sequencing by Novogene.

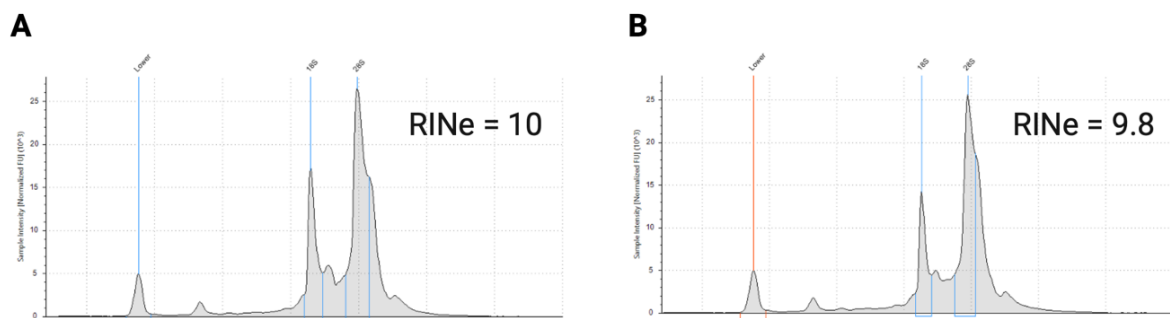


Figure 3.1 RNA integrity validation.

Example TapeStation profiles of RNA extracted from **A)** PEO1, and **B)** PEO4 with their calculated RINe values.

3.2.2. Analysis of RNA-seq data

The bioinformatic analysis of RNA-seq data initially follows a similar pathway to that of ChIP data (**Figure 3.2**). Adaptors were trimmed, reads aligned, and aligned reads indexed to allow visualisation on IGV. Unlike ChIP, reads were aligned using Spliced Transcripts Alignment to a Reference (*STAR*), which is an aligner specifically for RNA-seq data.(218) The premise is like *bowtie2*, with the read being mapped to the most similar genomic sequence.

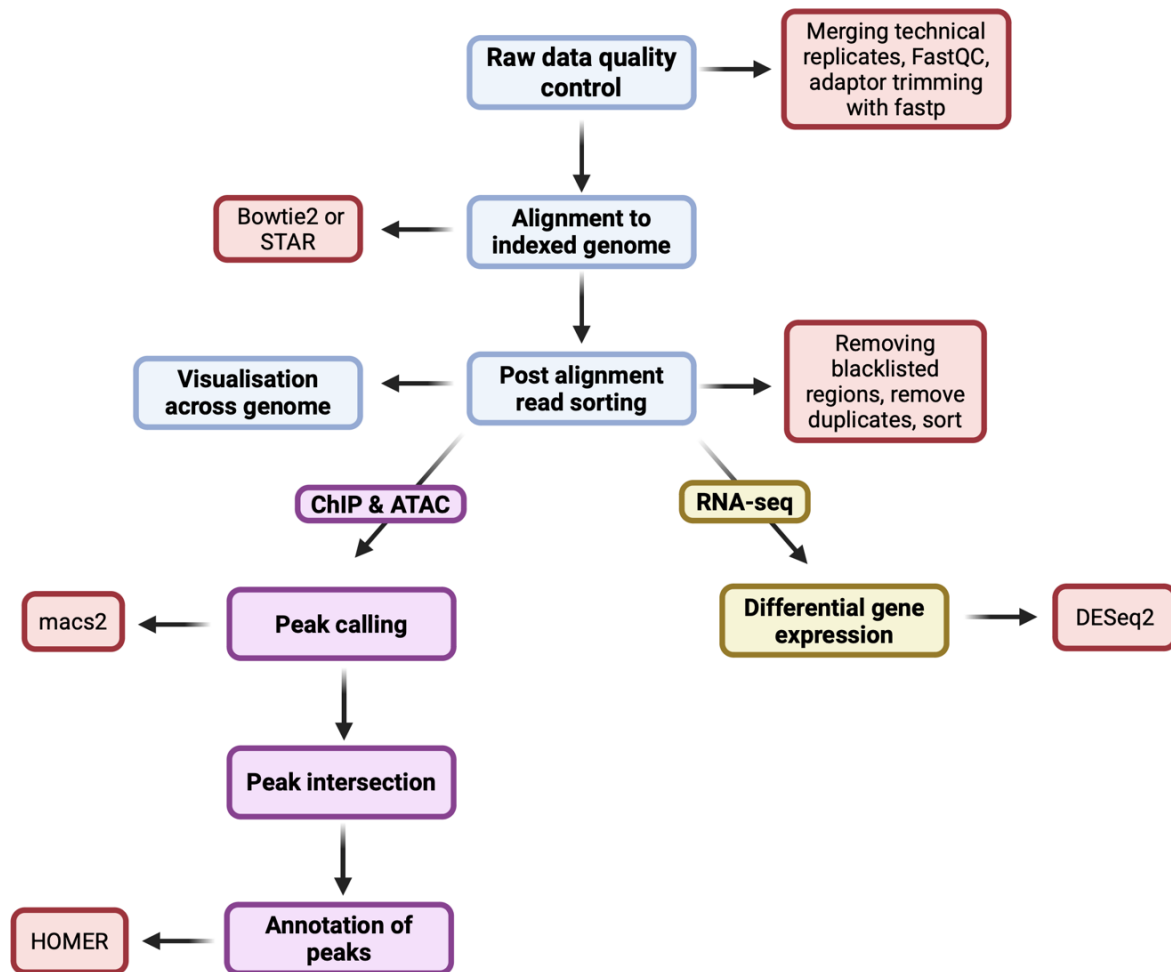


Figure 3.2 Bioinformatic analysis pipeline undertaken for NGS data.

After alignment, Principal Component Analysis (PCA) was conducted using iDEP.(219) This was performed to determine if there were any transcriptional differences occurring between biological replicates of the same cell line, and more widely between the PEO1 and PEO4 cell lines. The PEO1 and PEO4 cell lines were separated by PC1 meaning that the inherent transcriptional differences between the cell lines are responsible for the largest variance across all the samples (**Figure 3.3 A**). This was expected, as PEO1 and PEO4 are distinct cell lines from each other with key biological differences such as therapy response. Hence, the transcriptomes of the two cell lines will differ from each other through transcriptional adaptation, as exemplified by the PCA plot. Encouragingly, there was minimal variance between intra-cell line biological replicates, with the

maximal variance of PC2 being 3.3%. This highlights the reproducibility of cell culture and RNA extraction between biological replicates, as the transcriptome differences are negligible.

3.2.3. Differential expression

The key aim of RNA-seq analysis was to identify transcriptional variation between PEO1 and PEO4 cells, that could be associated with G4 formation, that may contribute to acquired drug resistance. To this end, differential expression analysis was conducted using DESeq2.(220) Genes are deemed to be differentially expressed between the two cell lines when the $\text{Log}_2(\text{fold-change})$ of gene expression is >1 , and the false discovery rate (FDR) was <0.05 . $\text{Log}_2(\text{fold-change}) >1$ means that there is a 2-fold increase in gene expression in one cell line over the other, and an FDR of 0.05 means that only 5% of these genes are expected to be a false positive. This means that differentially expressed genes that pass these thresholds are of statistical significance and biological relevance. In PEO4 cells, 2005 genes were identified as being significantly upregulated, and 2043 genes were significantly downregulated (**Figure 3.3 B**). This shows that there are key transcriptional differences between PEO1 and PEO4 cells, that could be implicated in the development of acquired drug resistance. The cause for some of these significantly differentially expressed genes could be due to alterations in copy number between the two cell lines. This would mean that the gene has not been actively upregulated or downregulated per se, but rather the basal number of the underlying gene has increased or decreased leading to an increase or decrease in available transcripts for detection. Hence, this could confound active epigenetic upregulation or downregulation of expression due to the acquisition of resistance. It would be beneficial to discern which genes copy number alteration applies to, in order to fully appreciate epigenetic effects occurring within the cell to drive this differential expression.

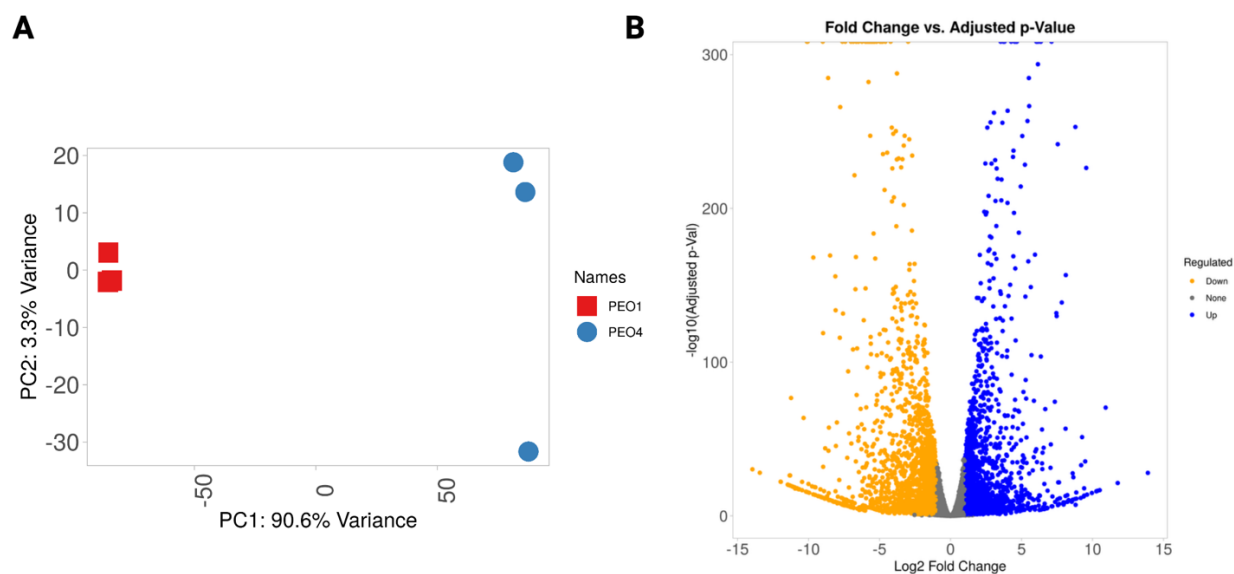


Figure 3.3 RNA-seq replicate reproducibility and differential expression.

A) PCA plot of RNA-seq data from three biological replicates extracted from PEO1 and PEO4 cells. **B)** Volcano plot of differentially expressed genes in PEO4 cells relative to PEO1 using DESeq2, defined as $\text{Log}_2(\text{fold-change})$ of gene expression >1 , and FDR was <0.05 . 2005 genes were significantly upregulated, and 2043 genes were significantly downregulated. Made on *iDEP*.(219)

The top 10 most significantly differentially expressed genes were determined from this generated dataset and compared to other datasets generated in the PEO1/4 cell lines from different studies (Table 3.1). Out of the top 10 genes differentially expressed from the RNA-seq data generated in this project, 4 are shared with one or more of the other datasets. Indeed, this is in comparison to a PEO1/4 pair in which the PEO1 cells have a different BRCA status, owing to a secondary mutation obtained during cell culture.(192) This overlap is in line with the overlap of Gallon et al., dataset, and the PEO1 BRCA+/PEO4 dataset, which both have five of the most significantly differentially expressed genes being present in on or more other datasets. This shows that the RNA-seq dataset produced in this project is in consensus with other datasets produced on the same cell lines and achieves a similar overlap to what is achieved between other datasets. It is expected that the exact same genes may not appear in the top 10

significantly differentially expressed genes of other datasets, owing to differing sources that the cells were obtained from, differing genetic backgrounds and potentially differing culturing conditions.

PEO1 BRCA ⁻ /4 from this project	PEO1 BRCA ⁺ /4 generated in the Di Antonio group	PEO1/4 from Gallon et al., BRCA status unknown
OAS2	GDA	ADAMTSL4
SLC16A7	ID4	IFI27
ADAMTSL4	COL4A2	GDA
STAT1	CHRD1	LY6E
PROM1	CSF1	COL4A2
NTN1	CDHH3	COL4A1
FLNC *	TGFB1	FLNC *
BCAT1	PLCB4	NTN1
CSF1	FLNC *	KRT81
FSCN1	ARHGEF7	CHRD1

Table 3.1 Top 10 most differentially expressed genes.

The top 10 most significantly differentially expressed genes found in the PEO1/4 cell line pair from the RNA-seq data generated in this study, in the PEO1/4 cell line pair in which PEO1 is BRCA⁺ generated within the Di Antonio group, and from Gallon et al., in which the BRCA status is unknown.(80) Genes that are significantly upregulated in PEO4 are in green, and those downregulated in red. Genes which are shared with one other data set are in bold, and those across all three studies bold and with a star.

Pathways associated with significantly upregulated or downregulated genes were then deciphered using KEGG pathway enrichment on *iDEP*. Notably, immune system and interferon signalling pathways were the most prominently upregulated in PEO4 cells (**Figure 3.4 A**). This could be due to the chronic stress that drug-resistant ovarian cancer cells are under from chemotherapy treatment, that has been shown to induce resistance in ovarian cancer mouse models.(221)(222) This contrasts with PEO1 cells, that are killed more efficiently by chemotherapy agents, and do not survive long enough to activate sustained stress responses. It has also previously been reported that ovarian cancer cells can evade immune detection via epigenetic reprogramming.(223)(224)(225) Notably, KEGG pathways classically associated with DDR are not significantly upregulated in PEO4 cells, relative to PEO1. This demonstrates that alternative mechanisms could also transcriptionally contribute to acquired drug resistance, and that it is not solely reliant on the difference in BRCA2 status or differential DDR capacity. Of the significantly downregulated KEGG pathways, neurotransmitter signalling and canonical survival pathways such as MAPK and calcium signalling are enriched (**Figure 3.4 B**). This could be reflective of an anti-apoptotic phenotype that drug-resistant ovarian cancer cells adopt to maintain survival in a chemotherapeutic environment.(226)

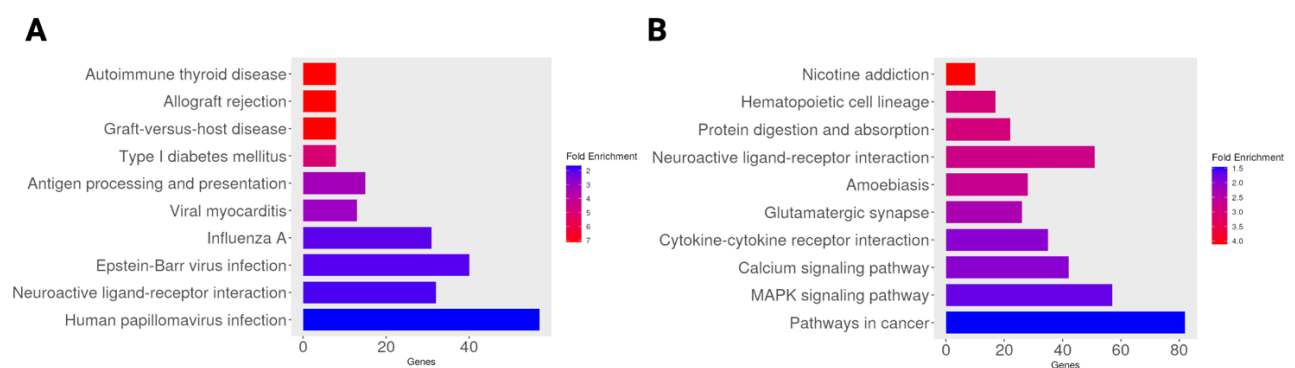


Figure 3.4 Differential gene expression pathway enrichment in PEO4. KEGG pathway enrichment of genes that were **A)** significantly upregulated or, **B)** significantly downregulated in PEO4. Made with *iDEP*.

To corroborate findings from KEGG pathway enrichment, the top 500 most significantly upregulated genes were also assessed using Gene Set Enrichment Analysis (GSEA), to find enriched hallmark gene sets. Similarly to KEGG pathway enrichment, there was a high enrichment in interferon and immune system pathways (**Table 3.2**). This contributes to the idea that PE04 cells exhibit an immune-like response, stemming from a chronic stress reaction to chemotherapeutic treatment, that could lead to a resistant phenotype. Additionally, the JAK/STAT pathway is the major signalling pathway activated via interferon signalling.(227) This pathway is implicated in cellular proliferation, survival, and development of resistance to chemotherapeutic agents.(228) Suppression of this pathway has been associated with increased sensitivity to treatment in ovarian cancer cells.(228)

Hallmark gene set	Gene set description	Adjusted p-value
Interferon alpha response	Genes up regulated in response to alpha interferon proteins.	2.73×10^{-25}
Interferon gamma response	Genes up regulated in response to IFNG.	1.96×10^{-24}
KRAS signalling down	Genes downregulated by KRAS activation.	2.54×10^{-11}
Inflammatory response	Genes defining inflammatory response.	2.23×10^{-10}
Myogenesis	Genes involved in development of skeletal muscle (myogenesis).	1.01×10^{-7}

Table 3.2 GSEA enrichment of differentially expressed genes in PE04.

The top 5 hallmark gene sets from GSEA of significantly upregulated genes in PE04 relative to PE01.

3.3. Profiling of chromatin accessibility

Chromatin accessibility was also determined in the PEO1 and PEO4 cells, as accessibility is a key mediator in G4 formation, and acquired drug resistance in ovarian cancer.(97)(80) Regions of open chromatin were mapped using ATAC-seq, in PEO1 and PEO4 cells. The rationale behind this, was that changes in chromatin accessible site distribution can affect transcription, and have been linked to epigenetic regulation of drug resistance in ovarian cancer.(80) Moreover, G4s are known to predominantly form in open chromatin regions.(97) Therefore, comparing the impact of G4 formation to an already established epigenetic regulator of drug resistance can provide a greater insight into the role of G4s in this context. ATAC-seq employs a Tn5 transposase enzyme that cleaves DNA at open chromatin regions, and inserts adaptors thus fragmenting the DNA.(229) After sequencing, regions of high read density are identified as open chromatin regions, and can be coupled with transcriptomic data to assess how changes in chromatin accessibility are linked to changes in gene expression between PEO1 and PEO4 cells.

3.3.1. Assessment of ATAC profiles

Prior to sequencing, ATAC-seq samples were subjected to quality control checks using the TapeStation. These controls aimed to visualise fragment sizes generated by the transposase, and whereabouts the peaks are located in relation to each other. ATAC-seq fragments typically are between the 200-1000 bp range. The peaks appear due to fragment build up at that fragment length and should be 160-200 bp apart from each other. This is because 147 bp of DNA are coiled around histones to form nucleosomes, which are separated from each other by 20-90 bp of linker DNA.(230) Subsequently, the peaks on the TapeStation signifying successful transposition are positioned 160-200 bp apart from each other, due to the periodicity of nucleosomes. This nucleosomal laddering is indicative of a high-quality library. For ATAC-seq samples from both PEO1 and PEO4 cells, this characteristic nucleosomal laddering was identified (**Figure 3.5 A&B**). This gave confidence that the ATAC-seq libraries were suitable for sequencing, and thus three

biological replicates from each cell line were subjected to paired end sequencing at The Francis Crick Institute.

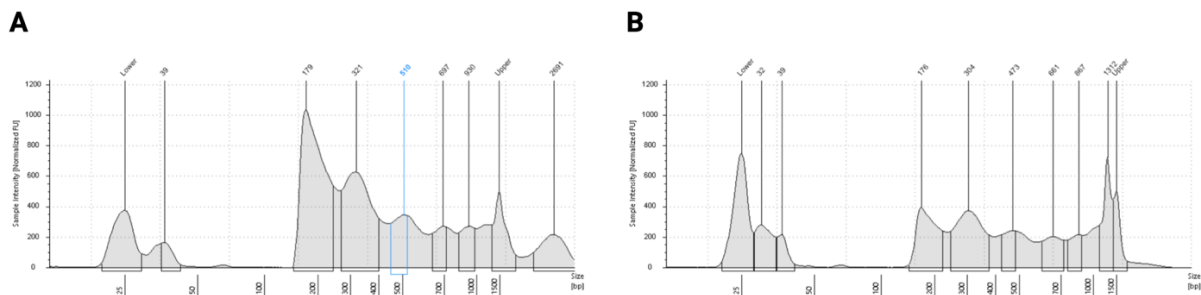


Figure 3.5 ATAC-seq sample validation.

Example TapeStation profiles of ATAC-seq samples from **A)** PEO1, and **B)** PEO4, showing a characteristic nucleosomal ladder indicative of successful ATAC-seq library generation.

3.3.2. Analysis of ATAC-seq data

Once sequenced, raw ATAC-seq data was processed following the same bioinformatic pipeline as was used for ChIP-seq. The aim was to identify regions of high read build up that were classed as chromatin accessible sites. Changes in the locations of these sites between PEO1 and PEO4 cells, particularly at intergenic loci, have previously been associated with ovarian cancer drug resistance.(80) Subsequently, ATAC NGS data was analysed to identify peaks unique to PEO1 or PEO4. These peaks could then be coupled with gene expression data to assess how these changes in chromatin accessibility are associated with transcriptional changes, in comparison to changes in G4 distribution.

ATAC NGS raw read files were quality assessed, the adaptors trimmed and aligned to the hg38 reference genome. The overall raw read count, uniquely mapped reads and read duplication rate is detailed in Table 3.3. Overall, the number of uniquely mapped concordant reads for each sample was sufficient for further downstream analysis, as it provides high enough depth to call accessible regions of over closed regions of chromatin. However, the duplication rate is higher than that obtained for the previous G4 ChIP

samples. This could be due to overamplification during the ATAC experiment, and a few cycles less of PCR amplification may reduce the read duplication rate in this instance.

Condition	Rep	Total no. of mapped reads	Overall read alignment	No. of uniquely mapped reads	Uniquely and concordantly aligned	Read duplication
PEO1 ATAC	1	72,642,289	97%	38,864,397	53.51%	28.37%
	2	88,299,461	94.53%	42,252,758	47.85%	37.9%
	3	92,582,121	96.28%	44,782,553	48.37%	37.9%
PEO4 ATAC	1	74,476,472	97.07%	46,247,462	62.10%	25.69%
	2	64,293,951	92.84%	35,842,818	55.75%	29.29%
	3	67,198,464	94.25%	40,379,281	60.09%	16.43%

Table 3.3 Alignment quality of ATAC samples.

Table of total number of reads, overall alignment rate, uniquely and concordantly aligned read number and percentage of total mapped reads and read duplication percentages for PEO1 and PEO4 input and ATAC samples.

After post-alignment read sorting and filtering, sample correlation was assessed using *deepTools plotCorrelation* using the Spearman method. Similarly to G4 ChIP, the intra-cell line samples correlated most strongly to each other (>0.94 PEO4, >0.92 PEO1), whilst correlation remained high between inter-cell line samples (0.82-0.86) (**Figure 3.6 A**). This indicates that relatively high reproducibility was achieved between biological replicates of the same cell line. Additionally, whilst it hints to conservation of accessible regions between PEO1 and PEO4 cell lines, there are still differences between the two aligned read files. These differences may be epigenetic adaptation mechanisms which differentiate one cell line from the other, as has previously been reported.

ATAC peaks were then called using a variation of macs2 that allows for the absence of an input file. In this instance, a uniform background is assumed, and regions of high density are determined relative to the uniform background. To obtain finalised ATAC peak maps in PEO1 and PEO4 cells, peaks were identified that were present in 2/3 of the biological replicates, the same as with G4 ChIP peaks. PEO4 cells had more accessible sites (103,846) than PEO1 cells (89,449) (**Figure 3.6 B**). Around half of those found in each cell line were shared with each other, showing a 50% conservation of accessible regions (**Figure 3.6 C**). There were 54,877 sites gained in PEO4 and 40,480 lost which were of most interest in the context of development of ovarian cancer drug resistance.(80) ATAC-seq data has previously been published in PEO1 and PEO4 cells, and 92% of PEO1 ATAC peaks and 82% of PEO4 ATAC peaks previously published overlapped with the ATAC data generated in this project.(80) However, the total number of ATAC peaks generated in each study varied, with the ATAC peaks ascertained in this project being almost double the previously reported figures for both PEO1 and PEO4. Owing to the high overlap, it may be that one study underestimates the number of accessible sites for each cell line, whilst the other overestimates. ENCODE does suggest that >50,000 peaks may be acceptable, whilst a peak number of >70,000 is recommended.(231) This figure is more illustrative of the ATAC peaks identified in this project, however they may not all be true peaks. The overlapping peaks detected could be deemed to be of high confidence between the cell lines, with the remaining peaks identified in this study either true accessible sites that have been previously missed, or technical artefact from PCR amplification. This highlights the variation in peak number that can occur through the completion of the same experimental method in different environments, and why comparison between different datasets must be done with caution.

The genomic locations of ATAC peaks were determined using *homer*. The primary site of ATAC peaks were at intron and intergenic loci (**Figure 3.6 D**). Interestingly, a similar pattern occurred to that of the G4 genomic distribution data. When considering unique peaks to PEO1 and PEO4, there was a decrease in those localised at promoters, which corresponded in an increased ATAC peak abundance at intron and intergenic regions. This highlights the importance of non-promoter sites in acquired drug resistance in ovarian cancer.

Additionally, a check of the tracks and determined peaks was conducted through visualisation of the genomic tracks in IGV (**Figure 3.6 E**). This showed a clear and distinct peak located in the RPA3 gene promoter, at the same location that a G4 peak had been previously detected. This strengthens the validity of the G4 determined at this location, as it it coincides with an accessible peak, a requisite of formation.

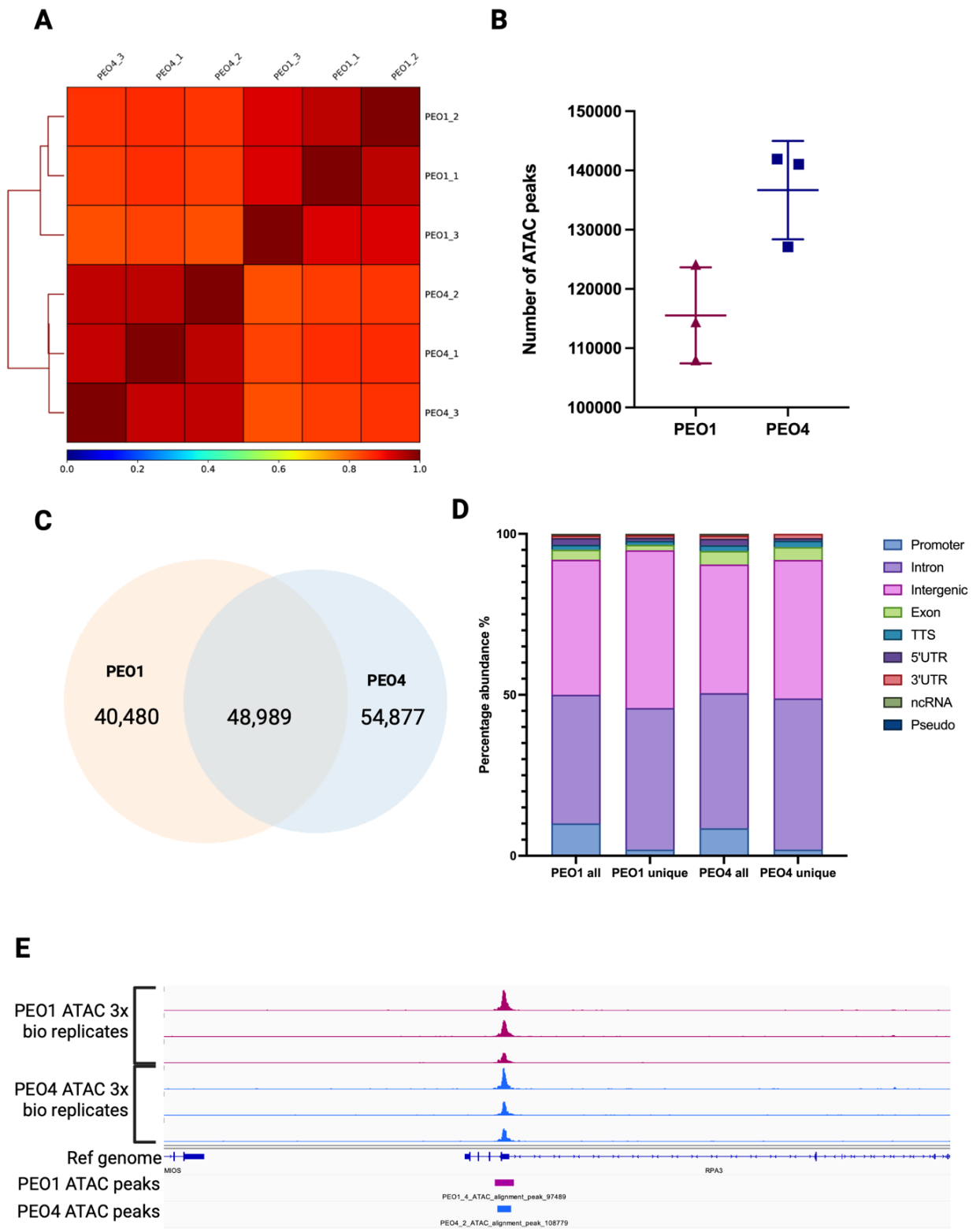


Figure 3.6 ATAC-seq data correlation, peaks and genomic distribution.

A) Heatmap Spearman correlation plots of PEO1 and PEO4 ATAC-seq samples, showing high correlation between PEO1 and PEO4 samples (0.82-0.86) and high reproducibility between biological replicates of the same cell line (>0.94 PEO4, >0.92 PEO1). Created using *deepTools plotCorrelation*. **B)** ATAC peaks called in three biological replicates, with 108,026-124,133 peaks called in the three PEO1 biological replicates and 127,118-141,926 called in the three PEO4 biological replicates. **C)** ATAC peaks unique to PEO1, PEO4 or conserved between the two cell lines, showing roughly 50% of accessible sites are conserved between the two cell lines. Made with BioInfoRx. **D)** Genomic location distribution of ATAC peaks, with the majority of accessible peaks residing in intron and intergenic regions in both PEO1 and PEO4. **E)** IGV plots of 3 replicates of PEO1 and PEO4 ATAC alignment files, showing a distinct peak present in all replicates in the RPA3 gene promoter.

3.4. Integration of G4 ChIP, ATAC and RNA-seq data

To assess the impact on gene expression associated with a gain or loss of G4 peaks in PEO4 cells, the expression change of the nearest gene was calculated for G4 peaks unique to PEO1 or PEO4. This was presented as the $\text{Log}_2(\text{FC})$ of expression, with a positive value equating to upregulated gene expression in PEO4 cells, and a negative value corresponding to downregulation. The gain or loss of G4 peaks was compared to the gain or loss of chromatin accessible sites, identified from ATAC-seq. To control for the inherent sequence bias associated with GC-rich regions, in which G4s form, ATAC peaks were intersected with OQS sites (95) to specifically select for GC-rich accessible sites. This was done to enable a fair comparison, and ensure that the gene expression changes associated with G4 formation could be attributed to G4 structure formation, and not just the high GC content found in that region.

3.4.1. Global gain and loss of G4s and ATAC in PEO4 cells

First, the impact of global gain or loss of G4 peaks on gene expression in PEO4 cells was analysed to assess if G4 formation correlates to increased gene expression, as previously reported in different cell lines.⁽²¹⁶⁾⁽¹⁶²⁾⁽¹¹³⁾ In line with expectations, a gain in G4 was associated with an average increase in gene expression ($\text{Log}_2\text{FC}=+0.45$), whilst a loss in G4 resulted in an average decrease in gene expression ($\text{Log}_2\text{FC}=-0.34$) (**Figure 3.7A**). This observation further validated the successful generation of viable G4 maps in the PEO1 and PEO4 cell lines, as they correlate with expected and validated transcriptomic effects.

By assessing the global effect of accessibility gain on gene expression in PEO4 cells, it was found that there was a negligible impact on the average fold change ($\text{Log}_2\text{FC}=+0.08$) (**Figure 3.7B**). This contrasts significantly with a gain in G4 ($\text{Log}_2\text{FC}=+0.45$, $p<0.0001$), highlighting the relevance of G4s in gene expression modulation comparatively to changes in chromatin accessibility, on average. In limited instances, a gain in accessibility is concomitant with a gain in G4. This results in a synergistic effect on gene expression upregulation ($\text{Log}_2\text{FC}=+0.86$), which is highly significant in comparison to a simple gain in accessibility ($p<0.0001$), and significant when compared to a gain in G4 ($p=0.0013$). Whilst this only occurred in 85 cases, this highlights the ability of G4 formation to act synergistically with other epigenetic mechanisms. Additionally, it substantiates the idea that gene expression changes associated with G4 formation is attributed to the G4 structure itself, as in most cases G4 formation is occurring in an already accessible region.

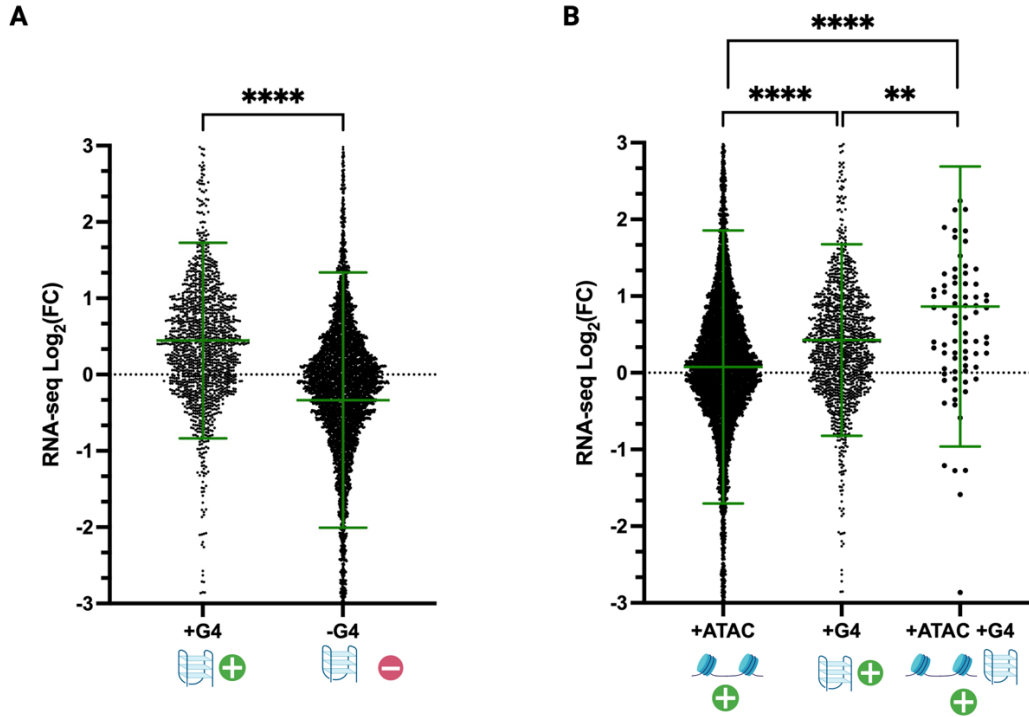


Figure 3.7 Differential gene expression of genes that gain or lose a G4.

Change in gene expression between PEO1 and PEO4 of genes associated with **A)** a gain or loss in G4 peak in PEO4 cells and **B)** a gain in ATAC peaks (overlapping with OQS), G4 peaks, or both an ATAC and G4 peak at the same time in PEO4. Statistical significance assessed by Mann-Whitney U-test, **= $p < 0.01$, ****= $p < 0.0001$. Error bars show standard deviation of gene expression data generated by RNA-seq in triplicate.

3.4.2. Changes in gene expression at different genomic loci

Analysis was then expanded to looking at gained PEO4 G4 and ATAC peaks mapped specifically to either promoter, intron, or intergenic regions to evaluate if genomic location influenced expression changes. When looking specifically at promoter peaks, a gain in either a G4 or ATAC peak resulted in an average increase in expression of proximal genes (**Figure 3.8 A**). This indicates the impact of both increased accessibility and G4 formation in transcriptional upregulation at promoter sites. A gain in an accessible peak in a promoter region warranted a larger expression increase ($\text{Log}_2\text{FC} = +0.66$) over a gain in promoter G4 peak ($\text{Log}_2\text{FC} = +0.37$), however this difference was not significant ($p = 0.26$). This observation contrasts with findings from previous genomic studies comparing healthy and immortalized cell lines and suggests that a gain in promoter G4 in

PEO4 cells does not confer a greater transcriptional advantage than a gain in promoter accessibility, at least when ovarian cancer cells evolve and develop drug resistance.(97)

Regional analysis was then extended to intron and intergenic loci. These regions were chosen as they were the most abundant sites of G4 formation in PEO4 cells after promoters, as well as for their implication in distal transcriptional regulation in other studies.(141)(232)(150) Moreover, non-promoter sites have been reported to be of importance for epigenetic regulation of drug-resistant ovarian cancer.(80)(78) For both intron and intergenic regions, a gain in G4 was associated with highly statistically significant increased gene expression, in comparison to a gain in chromatin accessibility ($p < 0.0001$) (**Figure 3.8B and 3.8C**). The largest expression increase was at gained intron G4 sites ($\text{Log}_2\text{FC} = +0.77$), which was 6-fold higher than a gain in an intron ATAC peak ($\text{Log}_2\text{FC} = +0.13$) (**Figure 3.8 D**). At intergenic sites, G4 peaks were associated with an average Log_2FC 5-fold higher than a gain in intergenic accessible sites ($\text{Log}_2\text{FC} = +0.45$ and $\text{Log}_2\text{FC} = +0.09$, respectively) (**Figure 3.8 D**). Indeed, it is also possible that the transcriptional impact of intergenic G4 formation is underestimated, due to the complexity of associating intergenic regions with target genes. Whilst promoters and introns can often be attributed to single genes, intergenic regions can regulate genes thousands of base pairs away.(217) Hence, the true regulatory power of intergenic G4s may be greater than indicated here. To further probe this, experiments that capture the chromatin conformation to visualise 3D looping interactions, to associate distal intergenic G4s with their gene targets, should be performed to determine their true transcriptional potential.(233)

These results indicate that G4s forming in non-promoter regions may be of importance in epigenetic reprogramming in drug-resistant ovarian cancer cells. Not only do they confer a larger increase in gene expression over a gain in chromatin accessibility, but they also correlate with larger gene expression changes than a gain in promoter G4s. Thus, these observations suggest that investigation into the effect of G4 formation should not be limited to proximal promoter sites, but extended to non-promoter regions that may also be leveraged for increased gene expression.

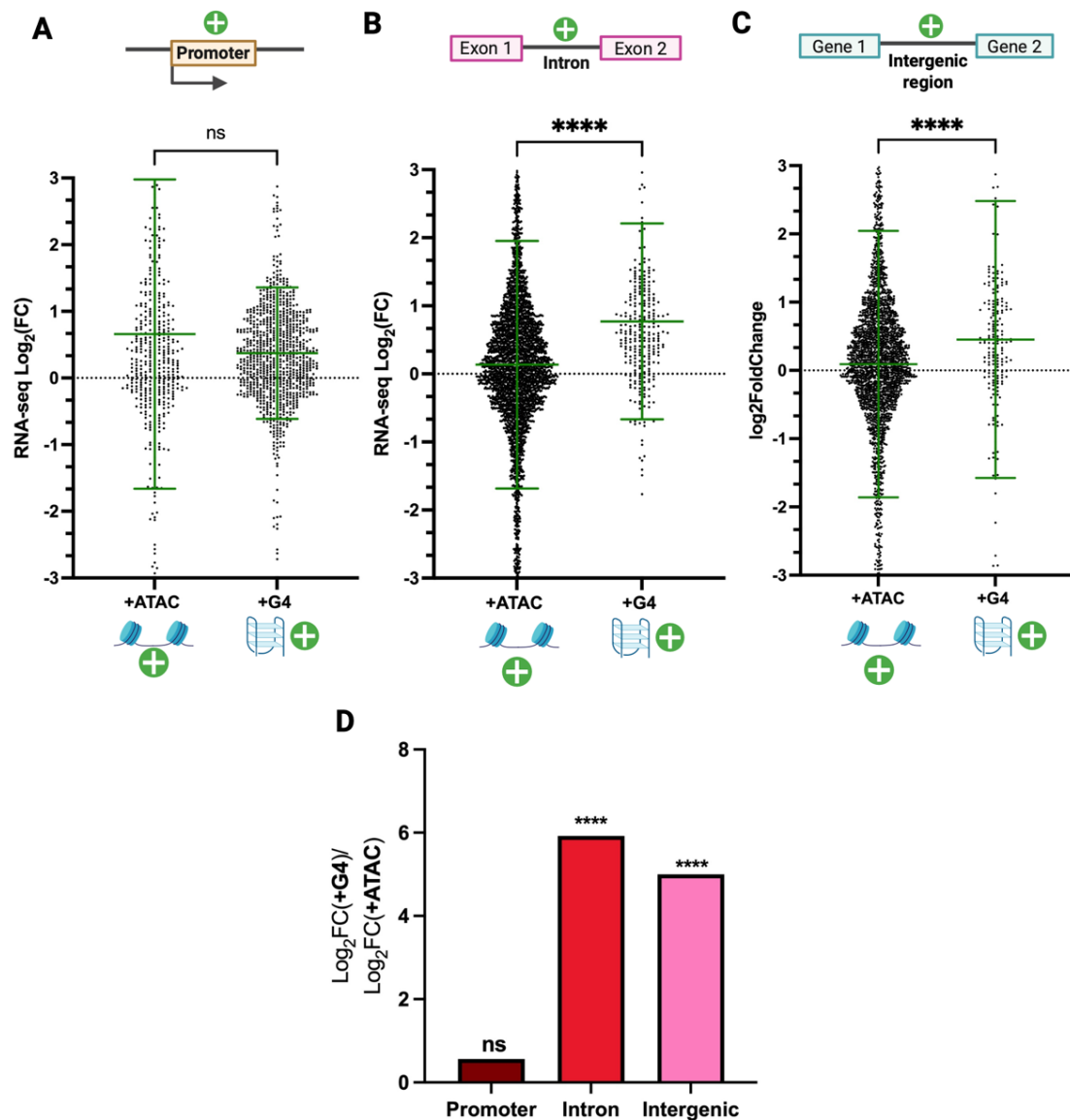


Figure 3.8 Differential expression of genes that gain a G4 at different locations. Change in gene expression between PEO1 and PEO4 of genes associated with a gain in ATAC peaks (overlapping with OQS) or G4 peaks at **A**) promoters, **B**) introns or **C**) intergenic regions in PEO4 cells. **D**) The fold-difference in Log₂FC between +G4 and +ATAC from promoter, intron and intergenic regions. Statistical significance assessed by Mann-Whitney U-test, ns=not significant, ****= $p < 0.0001$. Error bars show standard deviation of gene expression data generated by RNA-seq in triplicate.

To validate the conclusions determined from the data above, peak calling of G4 maps was conducted using three alternative parameters, to yield more stringent results. This was done to ensure the trends in gene expression associated with G4 gain were not due to false positive peaks from noisy genomic regions, which is a challenge associated with genomic mapping of DNA structures. Peaks were called by either; i) using only peaks that intersected with *in vitro* experimentally validated G4 sequence (OQS) (95), ii) calling peaks using a more stringent q-value of 0.01 instead of $q < 0.05$ (as is standard when using macs2 for peak calling), or iii) considering peaks present in three biological replicates instead of the two originally used. When using these more stringent parameters, the numbers of peaks may vary, which is to be expected when calling peaks with more rigorous parameters. This highlights that regions of low coverage may generate false positive peaks, which is an inherent issue with antibody based genomic mapping techniques.(234)(103) Noisy regions of the genome can appear highly variable, and result in the loss of a peak upon more rigorous peak calling.

Gratifyingly, regardless of the peak procedure calling implemented, the overall trend and correlation with gene expression remained unchanged in comparison to default macs2 peak calling parameters (**Figure 3.9 A-F**). The gain of G4s at non-promoter regions was still associated with a significantly greater increase in gene expression when compared to a gain in accessibility, for all three peak calling conditions (**Figure 3.9 B,D,F**). This demonstrates that the conclusions made from maps generated using default peak calling parameters remain unaffected when using a series of much more stringent parameters. This underscores the validity of the conclusions made, and highlights that the trends observed are not a result of variable and noisy regions.

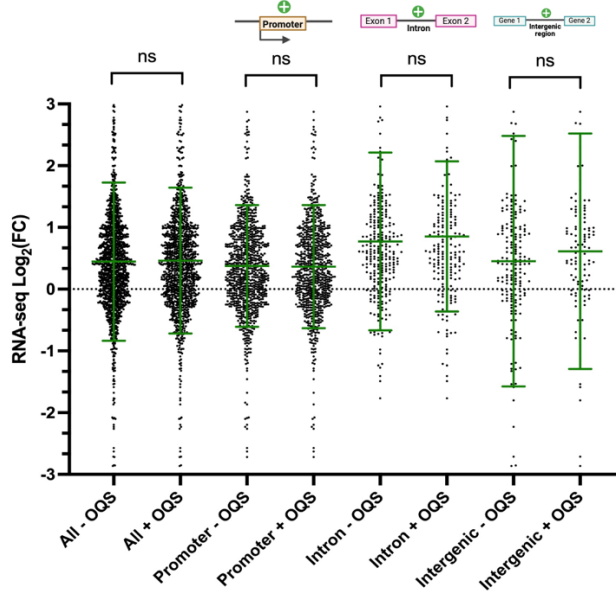
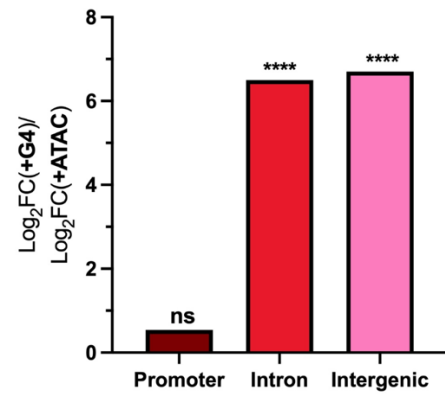
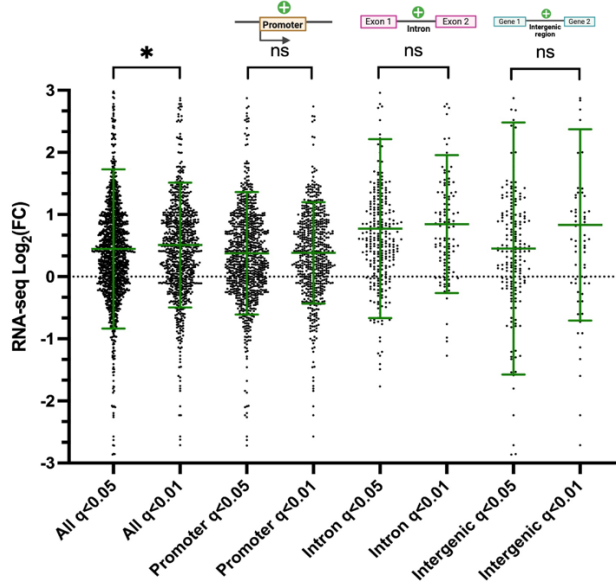
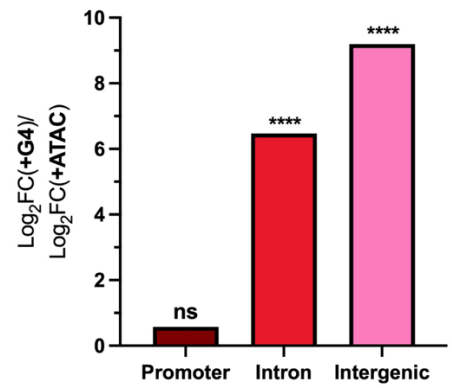
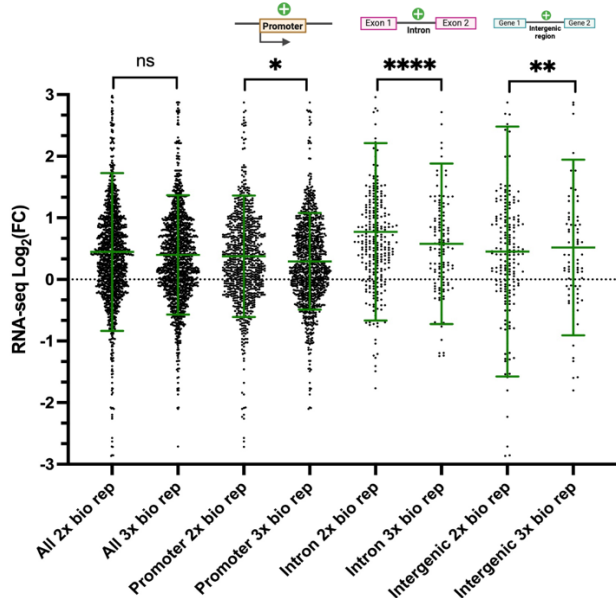
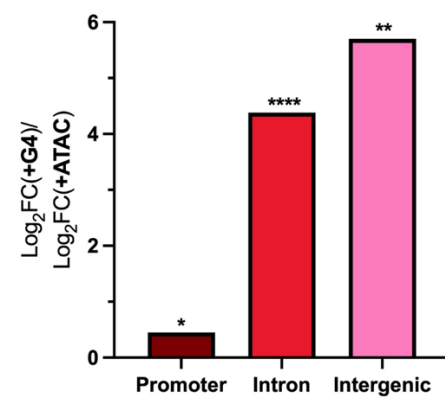
A**B****C****D****E****F**

Figure 3.9 Validation of G4 maps with alternative peak calling parameters. Change in gene expression between PEO1 and PEO4 of genes associated with a gain in G4 peaks in PEO4 cells called with default macs2 parameters from two biological replicates versus varying parameters (i-iii) and the fold-change in expression of the more stringent peaks compared to gained chromatin accessible sites. **A)** Default macs2 peaks compared to G4 peaks overlapping with OQS and, **B)** fold-change in gene expression of G4 peaks overlapping with OQS compared to gain in ATAC peak. **C)** Default macs2 peaks ($q < 0.05$) compared to peaks with $q < 0.01$ and **D)** fold-change in gene expression of $q < 0.01$ G4 peaks compared to gain in ATAC peak. **E)** Default macs2 G4 peaks from two biological replicates versus those present in all three biological replicates and **F)** fold-change in gene expression of G4 peaks present in three biological replicates compared to gain in ATAC peak. Statistical significance assessed by Mann-Whitney U-test, ns=not significant, ****= $p < 0.0001$. Error bars for the dot plots show standard deviation of gene expression data generated by RNA-seq in triplicate.

3.4.3. G4-related effect not observed with peak loss

Analysis into the impact of a loss of a G4 peak or chromatin accessible site on gene expression in PEO4 cells was also conducted, to assess if the observed trends for gain of sites were reversed. In promoter sites, the loss of either a G4 peak or ATAC peak was associated with a decrease in the average Log₂FC value (-0.35 and -1.17, respectively) (**Figure 3.10 A**). However, while the gain in either peak did not have a significant transcriptional impact over the other, the loss of a promoter accessible site caused a highly significant downregulation of gene expression compared to the loss of a promoter G4 ($p < 0.0001$) (**Figure 3.10 D**). This suggests that the loss of promoter accessibility could be key for epigenetic regulation in PEO4 cells, due to the considerable gene expression downregulation. The effect is much more prominent than a loss of G4 in the same loci, which gives equivalent up or down regulation of expression depending on gain or loss, respectively.

Moreover, at intron and intergenic sites, the loss of G4 structures is not associated with significantly downregulated gene expression, relative to the loss of chromatin accessibility in these regions (**Figure 3.10 B,C,D**). This supports the notion that the loss in chromatin accessibility has a stronger regulatory effect on gene expression than the gain, similarly to promoter sites. However, this is in contrast to the trend that a gain in G4 at non-promoter sites is associated with highly significantly increased gene expression relative to accessibility gain. This could be because G4 formation is more transcriptionally functional as opposed to gain in chromatin accessibility and subsequently utilised by PEO4 cells to upregulate gene expression at these loci. G4 gain was associated with a $+\text{Log}_2\text{FC}$ value equivalent to the $-\text{Log}_2\text{FC}$ associated with a G4 loss. This suggests that G4 gain and loss has the same transcriptional impact, and that one is not more transcriptionally altering than the other. However, whilst G4 loss had a similar downregulatory effect in comparison to G4 gain, it did not exceed the impact of losing accessible chromatin sites.

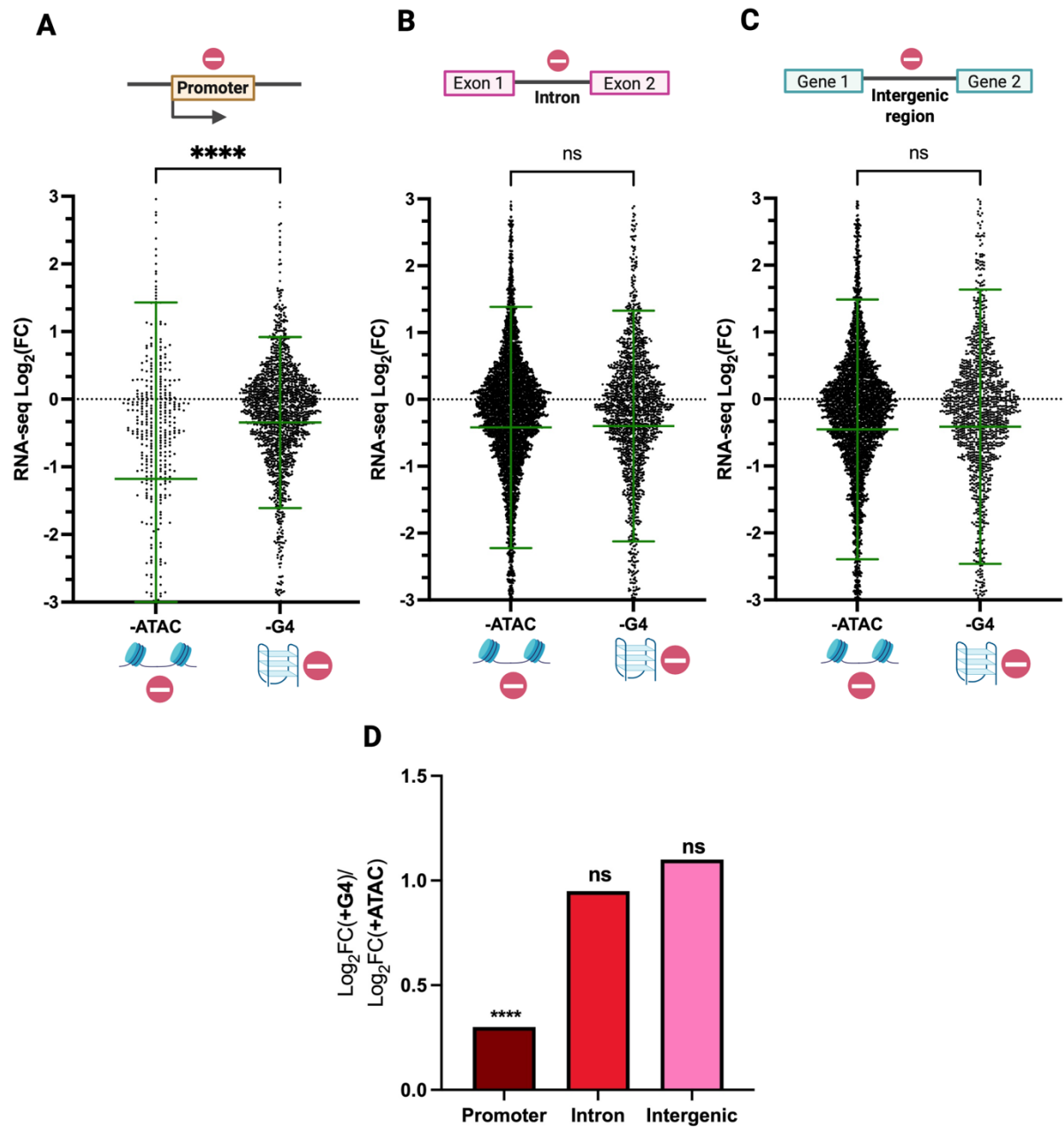


Figure 3.10 Differential expression of genes that lose a G4 in PEO4. Change in gene expression between PEO1 and PEO4 of genes associated with a loss in ATAC peaks (overlapping with OQS) or G4 peaks at **A**) promoters, **B**) introns or **C**) intergenic regions in PEO4 cells. **D**) The fold-difference in Log₂FC between +G4 and +ATAC from promoter, intron and intergenic regions. Statistical significance assessed by Mann-Whitney U-test, ns=not significant, ****= $p < 0.0001$. Error bars for the dot plots show standard deviation of gene expression data generated by RNA-seq in triplicate.

These results indicate that G4s have an independent transcriptional impact in PEO4 cells, relative to changes in chromatin accessibility. Whilst a gain in a non-promoter G4 generates a highly statistically significant increase in gene expression over gain in chromatin accessibility, the loss does not result in statistically significant downregulation. In promoter regions, the loss of accessibility causes the highest change in gene expression, highlighting the importance of accessible promoters in transcriptional capacity. Overall, PEO4 cells seem to leverage both mechanisms of epigenetic control to upregulate and downregulate gene expression.

3.5. G4s transcriptional association is independent of BRCA status

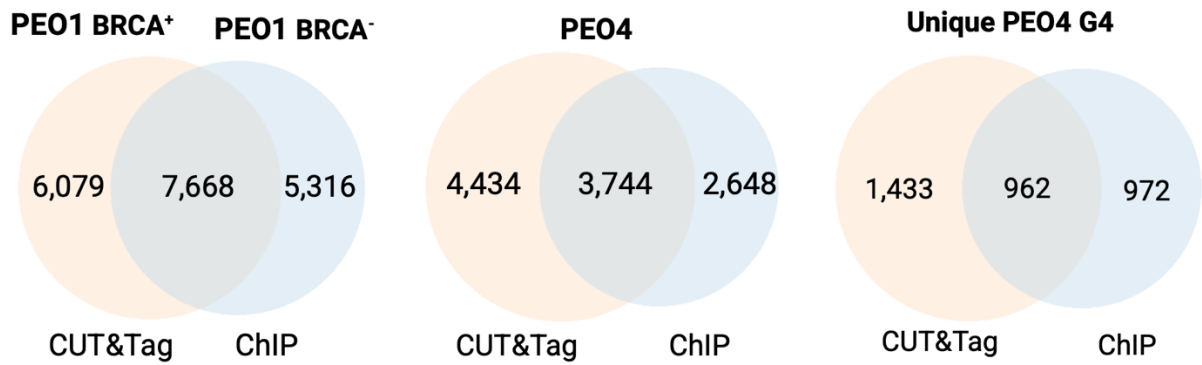
To differentiate the impact that BRCA status has on the acquired drug-resistant state of PEO4 cells, as opposed to epigenetic adaptation, G4 maps generated in the original PEO1/4 paired cell line were compared to maps generated in a pair in which the PEO1 cells gained an additional mutation that functionally reinstates BRCA2.(192) These BRCA⁺ PEO1 cells have a similar ability for homologous recombination to PEO4 cells, and so can be used to independently investigate how epigenetic adaption contributes towards acquired drug resistance. Importantly, the PEO1 BRCA⁺ cells exhibited a 6-fold greater sensitivity to cisplatin than PEO4 cells, with the experiment being conducted by Dr Jenna Robinson, a previous PhD student in the Di Antonio group. This is in line with the difference between the original PEO1 BRCA⁻ and PEO4 cell line pair. This highlights, that the functional restoration of BRCA2 in PEO4 cells is not the sole contributor to acquired drug resistance, and that epigenetic adaptation and transcriptional mis-regulation may also be key orchestrators.

G4 maps in the PEO1 BRCA⁺/PEO4 pair were generated by Dr Jenna Robinson, a previous PhD student in the Di Antonio group, using G4 CUT&Tag. This was to ensure that the maps generated were not biased by G4 mapping technique. Interestingly, more G4 peaks were detected in PEO1 BRCA⁺ than in PEO4 cells, at 13,745 and 8,176 G4s respectively. This is similar to the trend seen in original PEO1 BRCA⁻/PEO4 pair, and as CUT&Tag does not

rely on enrichment of G4 regions over non-G4 regions, indicates that the observed difference in peak numbers seen in the G4 ChIP cell lines could be of true biological relevance. The difference in overall G4 numbers between PEO1 and PEO4 is an interesting observation, which could be due to a variety of factors including the different epigenetic states of the two cell lines. The fact that PEO4 cells have more accessible sites than PEO1, yet a lower number of G4s, highlights how they can be independent epigenetic mechanisms and represent a difference in epigenetic state of the two cell lines. However, the molecular mechanism that underpins the difference in overall G4 numbers cannot be speculated based on the generated genomics data alone, and additional experiments would have to be undertaken to derive this.

The overlap of peaks detected in PEO1 and PEO4 cells from ChIP or CUT&Tag was approximately 59%, which is almost identical to the 60% overlap reported from previous studies (**Figure 3.11 A**).⁽¹⁰⁴⁾ This validates the consistency of the G4 maps generated by independent techniques, and between the two cell line pairs. Of those G4s gained in PEO4 cells, 962 of those were conserved between the two mapping techniques. When intersecting these conserved gained PEO4 G4s with RNA-seq data, there was no significant difference of the gene expression change associated with gained G4s detected by G4 ChIP alone, compared to those detected by both ChIP and CUT&Tag in promoter, intron and intergenic regions (**Figure 3.11 B**). These results indicate that whilst PEO4 cells have a global reduction in G4s, the elevated gene expression associated with gained G4s is independent of BRCA status and G4 mapping technique used.

A



B

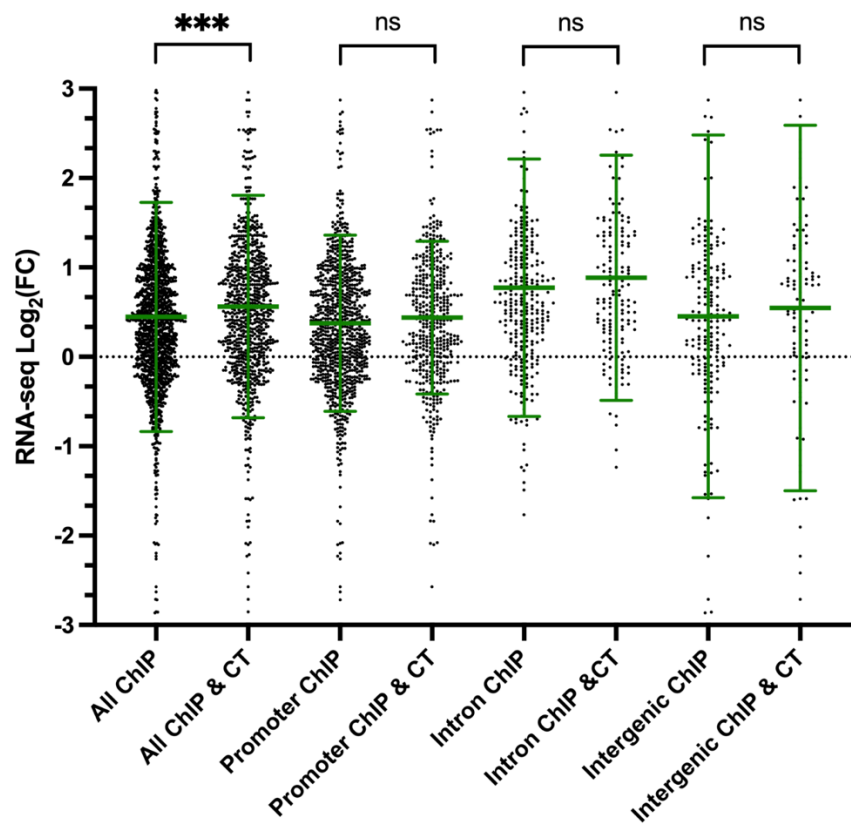


Figure 3.11 Differential expression of genes associated with a gain in G4, detected by ChIP and CUT&Tag, in two PEO models.

A) G4 peaks unique to and conserved between PEO1 BRCA⁻ and PEO1 BRCA⁺, PEO4, and unique peaks to PEO4 detected by either G4 ChIP or CUT&Tag. Made with BioInfoRx. **B)** Change in gene expression between PEO1 and PEO4 of genes

associated with a gain in G4 peaks detected in G4 ChIP compared to those detected by both ChIP and CUT&Tag in PEO4 cells.

3.6. Pathway enrichment of G4 associated genes

To assess whether the G4 associated transcriptional changes were implicated in drug resistance, pathway enrichment of significantly differentially expressed genes linked to G4 formation was performed. This was conducted similarly to global RNA-seq pathway enrichment that was described previously in this chapter, but with a prior selection step to analyse only genes that were associated with either a gain or loss in G4 in PEO4 cells. Additionally, genes correlated to peaks in promoters, introns, and intergenic regions were also investigated, to evaluate whether the pathways differed depending on peak loci.

3.6.1. Global pathway enrichment

Biological pathways enriched from significantly upregulated genes associated with a gain in G4 in PEO4 cells was first investigated to assess the global biological impact correlated with G4 gain. KEGG pathway enrichment (235) was performed using the ShinyGo gene-set enrichment tool (211), and the most enriched pathways evaluated (

Figure 3.12). Notably, pathways such as Wnt signalling and mTOR signalling are enriched and have previously been reported to play important roles in cell survival, growth, DDR, and proliferation in ovarian cancer.(236) Aberrant activation of such pathways can promote upregulated cellular proliferation and decreased apoptosis, contributing to acquired drug resistance.(38) In relation, Wnt and mTOR signalling plays a crucial role in DDR and maintenance of cancer stem cells, and signalling pathways regulating pluripotency of stem cells is also an enriched pathway associated with G4 gain.(237)(238) This suggests that upregulated genes correlated with G4 gain in PEO4 are promoting cancer stem cell characteristics and DDR. The acquisition of these stem-like traits through EMT has been highly associated with acquired ovarian cancer drug resistance.(32)(38)(239)(240)

Indeed, when performing a literature search using NDEx iQuery (241), that seeks for gene set similarities across >80,000 publications, the publication that had the highest gene set likeness was discussing the targeting of drug transport and the regulation of critical signalling networks in breast cancer.(242) The publication highlighted how the regulation of TGF-beta and p53 (amongst others) influence drug transport, cellular proliferative, and apoptotic networks that impact therapy response. Overall, the most similar published gene set is in concordance with the enriched KEGG pathways associated with gained G4s. This suggests that the global gain in G4, and associated increased gene expression, is enriched in pathways known for their importance in drug resistance in many cancers, including ovarian.

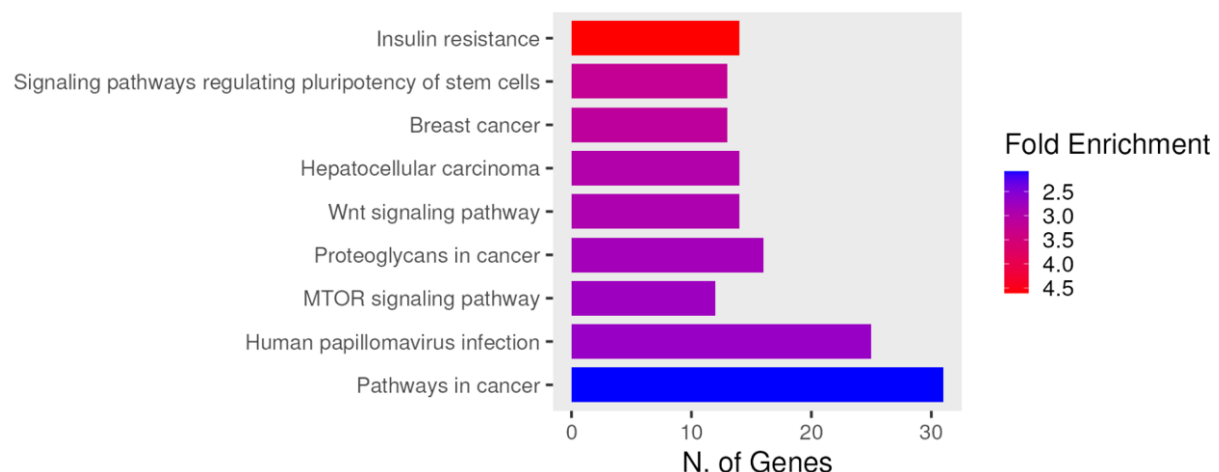


Figure 3.12 Pathway enrichment of upregulated genes linked with a gain in G4s. KEGG pathway enrichment of significantly upregulated genes associated with a gain in G4 in PEO4. Wnt, mTOR and pluripotency are all implicated in CSC, that are associated with acquired drug resistance in HGSOC. Made with *ShinyGO*.(211)

Pathway enrichment analysis was performed on significantly downregulated genes associated with G4 loss in PEO4. Fewer KEGG pathways were enriched compared to G4 gain and were of reduced relevance to drug resistance (

Figure 3.13). The exception to that was the enrichment of the calcium signalling pathways which can play a complex biological role. The up and downregulation of calcium signalling pathways can promote proliferation, cancer cell survival, and drug resistance, whilst also sensitising cancer cells to therapy, and inducing apoptosis in other instances.(243)(244)(245) For instance, some calcium channels, and associated pathways, are downregulated in chemotherapy resistant cells, whilst others are upregulated and contribute to the cancer stem cell like properties of a drug-resistant ovarian cancer cells.(245)(246)

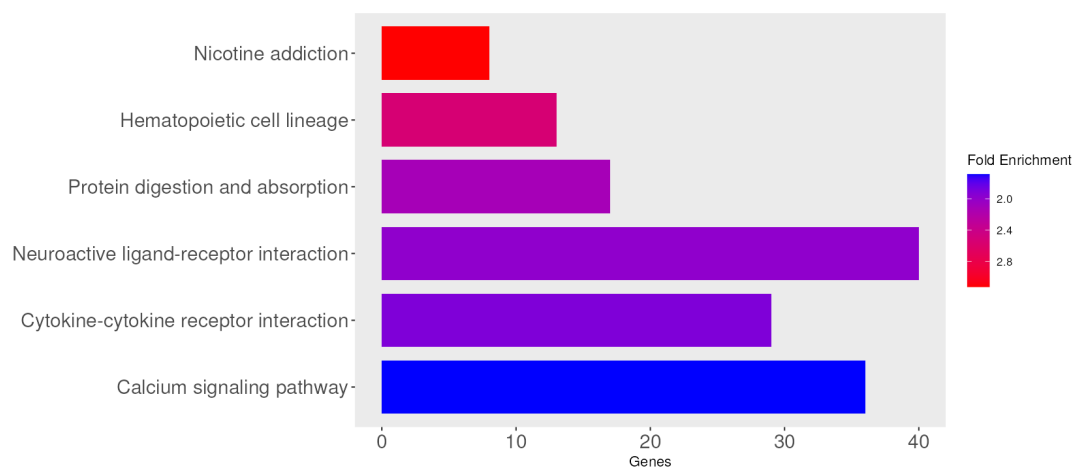


Figure 3.13 Pathway enrichment of downregulated genes linked with a loss in G4. KEGG pathway enrichment of significantly downregulated genes associated with a loss of G4 in PEO4. Fewer pathways are enriched relative to a gain in G4. Made with *ShinyGO*.(211)

Global pathway enrichment of significantly upregulated genes associated with a gain in G4 in PEO4 unveil multiple pathways previously associated with drug resistance in ovarian cancer.(238)(237)(38)(236)(21)(22) This contrasts with the loss of G4 in PEO4, and the associated significantly downregulated genes, in which limited downregulated pathways could be linked to drug resistance.(246)(243)(245) This implies that the gain in G4, and subsequent transcriptional upregulation, could be primarily involved with the

epigenetic reprogramming of PE04 cells to their drug-resistant state, as opposed to the loss of G4 and associated gene downregulation.

3.6.2. Regional pathway enrichment

To investigate whether the enriched pathways differed based on genomic loci of the G4 peak, KEGG pathway enrichment was completed on significantly upregulated genes associated with either promoter G4s, or non-promoter G4s (intron and intergenic combined).(235) For non-promoter regions, no KEGG pathways were enriched, most likely due to the lower levels of peaks and associated gene expression changes which can make the enrichment threshold harder to reach.(247) For promoter regions, the singular pathway of human papillomavirus infection was enriched.

As an alternative, Gene Set Enrichment Analysis (GSEA) was conducted to find enriched hallmark gene sets corresponding to these regions.(248) Upregulated genes associated with G4 formation in intron and intergenic regions are enriched in gene sets that have implicated in drug resistance (**Table 3.4**). The TNF signalling via NFkB gene sets is associated with ovarian cancer development and progression, and maintenance of cancer stem cells, which have all been linked to the development of chemoresistance.(249)(37) Similarly, it has been reported that targeting oestrogen response is a promising therapeutic avenue to overcoming multi-drug-resistant ovarian cancer.(250)(251) Likewise, glycolysis is linked to ovarian cancer drug resistance via a plethora of mechanisms, including the rapid production of ATP, maintenance of redox balance, and increased drug efflux even under drug pressure.(252)(253)(254)

Hallmark gene set	Gene set description	Adjusted p-value
Late oestrogen response	Genes defining late response to oestrogen.	5.35×10^{-6}
KRAS signalling and downregulation	Genes downregulated by KRAS activation.	5.35×10^{-6}
TNF signalling via NFKB	Genes regulated by NF-kB in response to TNF.	3.63×10^{-5}
Early oestrogen response	Genes defining early response to oestrogen.	2.2×10^{-4}
Glycolysis	Genes encoding proteins involved in glycolysis and gluconeogenesis.	1.18×10^{-3}

Table 3.4 GSEA enrichment of genes linked with a gained non-promoter G4.

The top 5 hallmark gene sets from GSEA of significantly upregulated genes associated with the formation of G4s in intron and intergenic regions in PEO4. Gene sets are identified by how much the genes present in the hallmark gene set are over-represented in the gene list corresponding to upregulated genes associated with the formation of G4s in intron and intergenic regions in PEO4.

GSEA was also performed on gained promoter G4s in PEO4 cells. The enriched gene sets associated with these G4s are linked with DNA damage response (UV), apoptosis (p53) and immunity (interferons) (

Table 3.5). Genes altered by UV damage can be in response to a DNA damaging effect, of which many anti-cancer drugs elicit. Subsequently, an upregulation of DDR mechanisms occurs to counteract these effects and contribute to therapy resistance.(255) The p53

pathway is also implicated in DNA damage, through regulating cell fate and the entrance into apoptosis after DNA damage has occurred. Aberrant upregulation of the p53 pathway has previously been connected to drug resistance for this reason.(256) However, the majority of HGSOC have a missense *TP53* mutation, PE01/4 cell included.(257) This suggests that upregulation of genes in this pathway is not the *TP53* gene itself, but rather downstream elements involved in the p53 pathway that contribute to drug resistance.

Hallmark gene set	Gene set description	Adjusted p-value
UV response upregulation	Genes up regulated in response to ultraviolet (UV) radiation.	1.53×10^{-5}
p53 pathway	Genes involved in p53 pathways and networks.	1.13×10^{-4}
Interferon alpha response	Genes up regulated in response to alpha interferon proteins.	1.29×10^{-4}
UV response downregulation	Genes downregulated in response to ultraviolet (UV) radiation.	2.66×10^{-4}
Interferon gamma response	Genes up regulated in response to IFNG.	5.31×10^{-3}

Table 3.5 GSEA enrichment of genes associated with promoter G4s.

The top 5 hallmark gene sets from GSEA of significantly upregulated genes associated with the formation of G4s in promoter regions in PE04. Gene sets are identified by how much the genes present in the hallmark gene set are over-represented in the gene list corresponding to upregulated genes associated with the formation of G4s in promoter regions in PE04.

Both promoter and non-promoter regions in PEO4 cells form G4s associated with significantly upregulated genes that are enriched for hallmark gene sets involved in drug resistance. Whilst the gene sets they enrich for are distinct, drug resistance in ovarian cancer can occur via many mechanisms that work cooperatively.(21)(22)(258) It can thus be interpreted that it is the formation of G4s present in all regions of the genome, that are correlated with transcriptional upregulation, that are enriched in pathways linked to drug resistance.

3.7. Chapter summary and conclusions

In this chapter, genomic G4 maps were intersected with RNA-seq data to assess the impact of a gain or loss of G4 on gene expression changes, in drug-resistant ovarian cancer cells. This was compared to changes in chromatin accessibility, which have previously been reported for their relevance to drug resistance.(80) The global gain in G4 was associated with an average increase in gene expression, whilst the loss in G4 corresponded to a decrease, supporting previous studies.(259)(216) Investigation into the impact G4 formation had at different genomic loci revealed that there was no transcriptional advantage in gaining a G4 peak over an accessible peak in promoter regions. However, a gain in G4 at intron and intergenic regions resulted in a large increase in gene expression relative to a gain in accessibility. Interestingly, this trend was not mirrored when G4s were lost in PEO4 cells, in which the difference over losing an accessible site was minimal. This indicates that it is the formation of G4s in PEO4 cells that is implicated in the transcriptional upregulation of target genes, and non-promoter G4s should also be considered. This observed trend was independent of the BRCA status of the PEO1/PEO4 pair, and G4 mapping technique used, strongly indicating the potential role of G4s in the epigenetic regulation of genes altered in cells with acquired drug resistance, regardless of their genetic background.

Pathway enrichment analysis was conducted to associate the gain or loss of G4s with transcriptional regulation in PEO4 cells in particular pathways. It was determined that G4 formation in PEO4 cells was linked to the upregulation of genes in pathways linked to acquired drug resistance in ovarian cancer, regardless of if the gain in G4 occurred in promoter or non-promoter regions. These include Wnt, TNF and mTOR signalling pathways, and UV response. Notably, upregulation of genes associated with UV response can be associated to increased DDR, a key pathway in the development of acquired drug resistance in HGSOC. Upregulation of DDR is readily employed by resistant HGSOC to evade apoptosis initiated by the unstable accumulation of DNA damage.(255) Moreover, mTOR signalling is implicated in the DNA damage response, and cisplatin resistant tumour cells often have constitutive activation of mTORC1 signalling. This is via the close association with FANCD2, which is a DNA interstrand crosslink repair protein.(50)

Upregulated mTOR signalling results in increased FANCD2 expression, that is associated with a decreased response to platinum therapy. This suggests that the formation of G4s within the mTOR pathway could be leveraged by PEO4 cells to sustain increased gene expression of this pathway and contribute to PEO4 cisplatin resistance. The enrichment of Wnt and TNF signalling associated with significantly upregulated genes associated with a G4 gain is also of high relevance to acquired drug resistance in HGSOC. These signalling pathways regulate EMT associated genes and promote CSC properties.(38)(37) EMT and CSC are two tightly linked mechanisms of acquired drug resistance in HGSOC, where EMT promotes the acquirement of stem-like features, and CSCs have an inherent tendency to resist apoptosis. Aberrant Wnt and TNF signalling has been linked to chemotherapy resistance, and downregulation of these pathways has been shown to increase sensitivity of resistant cells to chemotherapy.(40) Therefore, the gain in G4s in PEO4 cells are associated with significantly increased expression of pathways validated for their involvement in acquired drug resistance, in HGSOC. This could be a consequential formation because of upregulation of these pathways, or the formation of G4s may be leveraged by the cells to maintain a heightened level of gene expression in pathways promoting resistance.

It is important to highlight that the most enriched pathway associated with G4 formation in PEO4 cells is insulin resistance. Indeed, this is a pathway that is resistance related, just to a different source. It would be a beneficial expansion of the work presented here to probe deeper into which genes are implicated in enriching this pathway, and if they are of relevance to the acquisition of drug resistance. It may well be that the genes implicated in this pathway are also enriched in other pathways such as mTOR. At present, the current analysis lacks sufficient depth into the individual genes that are being enriched in these pathways and would be an important development to truly understand the key genes that G4s may be regulating in these key pathways. By ignoring the most enriched pathways as they are not deemed relevant initially, there is a risk of overlooking key G4 biology which may be implicated in drug resistance.

4. Perturbation of G4s transcriptional impact

4.1. Chapter overview

To directly probe the transcriptional effects exerted by G4s, PEO1 and PEO4 cells were exposed to G4 targeting ligands, with the aim of investigating the transcriptional perturbation associated with G4 targeting. This would help uncover the relevance of G4-mediated transcriptional control on drug resistance pathways in PEO4 cells. It is important to reiterate that G4 ligands often have the opposite biological impact that an endogenous G4 does. Hence, stabilisation of G4s is generally associated with transcriptional repression. Subsequently, this was selected as the method to investigate the direct link between G4s and gene expression changes associated with drug resistance.⁽²⁶⁰⁾⁽¹¹³⁾ Two G4 ligands were used to avoid any ligand biased observation. Cells were treated with G4 ligands for 2 or 6 hours prior to RNA extraction, to investigate the transcriptional impact of G4 targeting over time. Previous studies have shown that treatment of cells with G4 ligands causes an induction of DNA damage at 24 hours.⁽¹⁸³⁾⁽²⁶⁰⁾ To prevent enrichment of these pathways, and to instead see the primary impact of transcriptionally repressing G4s, 2- and 6-hour time points were chosen.

4.1.1. G4 ligand dose optimisation

24-hour cell viability assays were conducted with the G4 ligands Pyridostatin (PDS) and PhenDC-3 to establish concentrations for use in RNA-seq experiments, to directly probe the impact of G4 formation on acquired drug resistance pathways. For both ligands, a concentration of 5 μ M was chosen, as indicted in **Figure 4.1 A** and **Figure 4.1 B**. This concentration was selected as at 24-hours a marginal effect on cell viability was seen. This indicates that this dose is low enough to avoid global toxicity yet is still biologically active

within the cells. Subsequently, transcriptional changes linked with G4 targeting should still be identified at 2- and 6-hour timepoints, without observing secondary effects, such as cell death and stress responses, which would mask such observations.

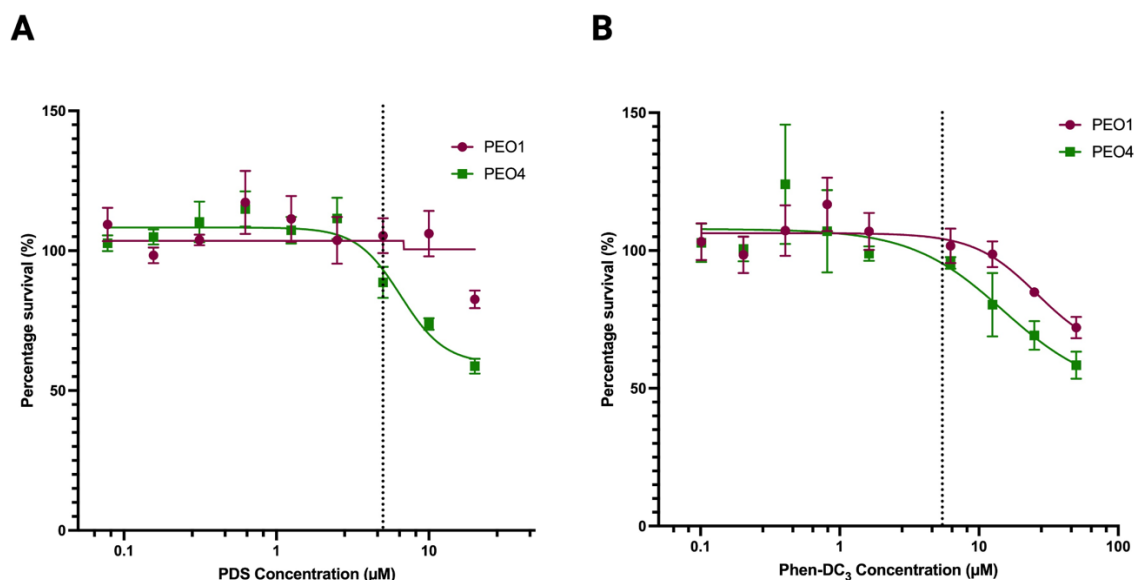


Figure 4.1 Cytotoxicity of PDS and PhenDC-3 in PEO1 and PEO4.

MTS viability assays after 24-hours of PEO1 and PEO4 cells in response to increasing concentrations of **A)** PDS and **B)** PhenDC-3. Dotted lines show the concentrations chosen for 2- and 6-hour treatment followed by RNA-seq. Experiments were performed in triplicate, with error bars showing the standard deviations.

4.1.2. Transcriptional effect of targeting G4s over time

To visualise any general transcriptional differences arising from cell line, ligand, or treatment time, Principal Component Analysis (PCA) was conducted using iDEP.(219) PEO1 and PEO4 cell lines are separated by PC1, meaning that inherent cell line differences are the biggest cause of variance across all samples (**Figure 4.2**). PC2 separates the samples based on whether the cells had been treated with a G4 ligand or not, meaning drug treatment is the cause of the second most prominent variation. The G4 ligand treated samples cluster on top of each other, irrespective of ligand used, or treatment

time, highlighting similar transcriptional effects of G4 ligands in both the PEO1 and PEO4 cells.

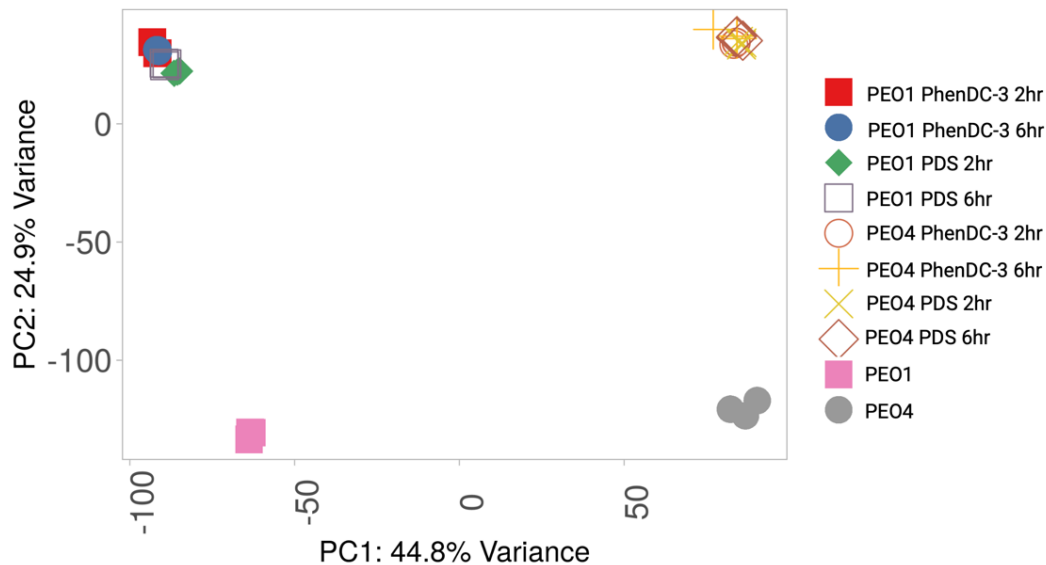


Figure 4.2 PCA plot of PEO1 and PEO4 RNA-seq with G4 perturbation.

PCA plot of native PEO1 and PEO4, and after treatment for 2 or 6 hours with PDS or PhenDC-3. RNA-seq was performed as n=3 for each sample, and data presented is the average of those three replicates. Made on iDEP.(219)

Drug treated RNA-seq data was then intersected with G4 maps to assess how drug treatment impacted transcription in different regions. In PEO1 cells, the average downregulatory effect of either PDS or PhenDC-3 at 2 or 6 hours was negligible, at only Log_2FC -0.05 to -0.08 (**Figure 4.3 A**). This contrasts with genes associated with G4s in intron and intergenic regions which saw a more pronounced average decrease in expression, particularly when using PhenDC-3 (**Figure 4.3 B&C**). However, the downregulatory effect was not temporal, as there was no significant decrease in gene expression when using PDS or PhenDC3 between 2 hours and 6 hours. With PhenDC-3, there was even a significant increase in the average gene expression associated with genes correlated with intronic G4s between the 2-hour and 6-hour timepoints ($\text{Log}_2\text{FC}=-$

0.30 and -0.21, respectively). Furthermore, for genes associated with intergenic G4s there was a non-significant increase in gene expression, from $\text{Log}_2\text{FC} = -0.27$ to -0.22. This could mean that the G4s present in these regions in PEO1 cells are involved in crucial processes needed for cell survival. Consequently, the cells work to counteract the transcriptional repression that stabilising the G4s causes, via upregulation of expression through alternative mechanisms. To determine whether this is the case, or if it is experimental artefact, additional timepoints would need to be generated in between 2 and 6 hours, and post-6 hours. This would show whether 5 μM PhenDC-3 causes a primary decrease in gene expression, that is followed by upregulation over time, as the cells transcriptionally adapt to overcome the effects of G4 ligand treatment. Given this differential expression observed here, it is somewhat surprising that not more variation is not observed in the PCA plot presented before. This could be that the downregulation and alteration in expression occurs within the same genes regardless of drugging condition, which in turn does not alter the PCA plot substantially. Alternatively, it could be that as PC2 appears to segregate the G4 stabilised cells from the naïve cells, that a PC3 would be needed to differentiate variation caused by the different G4 ligands used in this experiment.

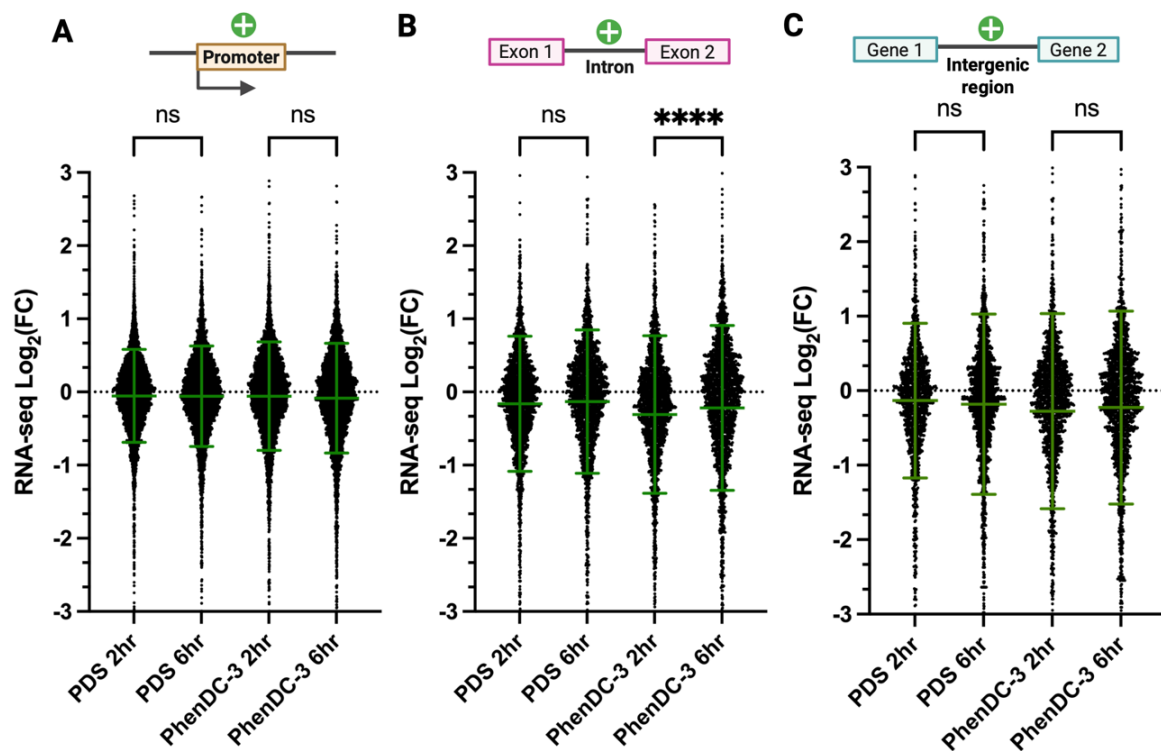


Figure 4.3 Differential expression of genes linked with PEO1 G4s after G4 perturbation. Change in gene expression between native PEO1 and PEO1 treated with either PDS or PhenDC-3 for 2 or 6 hours of genes associated with PEO1 G4 peaks at **A)** promoters, **B)** introns or **C)** intergenic regions. Statistical significance assessed by Mann-Whitney U-test, ns=not significant, ****= $p < 0.0001$ comparing the difference in gene expression between 0 hours and 2 hours and 0 hours and 6 hours drug treatment. This was done to assess the statistical significance of downregulation of gene expression of each time point and if they differed to each other. Error bars show standard deviation of gene expression data generated by RNA-seq in triplicate.

In PEO4 cells, 2-hour G4 ligand treatment caused a negligible effect on the change in average gene expression across all regions (**Figure 4.4 A-C**). In promoter loci, this observation persisted after 6 hours of PDS treatment, similarly to in PEO1 cells (**Figure 4.4 A**). When using PhenDC-3, a significant increase in the average gene expression was seen after 6 hours of treatment, relative to at 2 hours ($\text{Log}_2\text{FC} = +0.08$ and $+0.01$, respectively, $p < 0.0001$). For non-promoter regions, the average expression change decreased after 6-hours of G4 ligand treatment (**Figure 4.4 B&C**). This decrease was not significant in relation to the 2-hour timepoint for PDS treated cells, but highly significant for PhenDC-3 treated cells ($p < 0.0001$). This trend could have biological significance with how G4 targeting in drug-resistant cells has temporal dependence for its transcription altering effects. As PEO4 cells are drug-resistant, it may take longer for the G4 ligands to enter the cells due to the acquired adaptation mechanisms to counteract drug effects, such as reduced influx.(261)(262)(258) Subsequently, additional timepoints after 6-hours should be collected to gain a more comprehensive view into the transcriptional impact of G4 targeting in PEO4 cells.

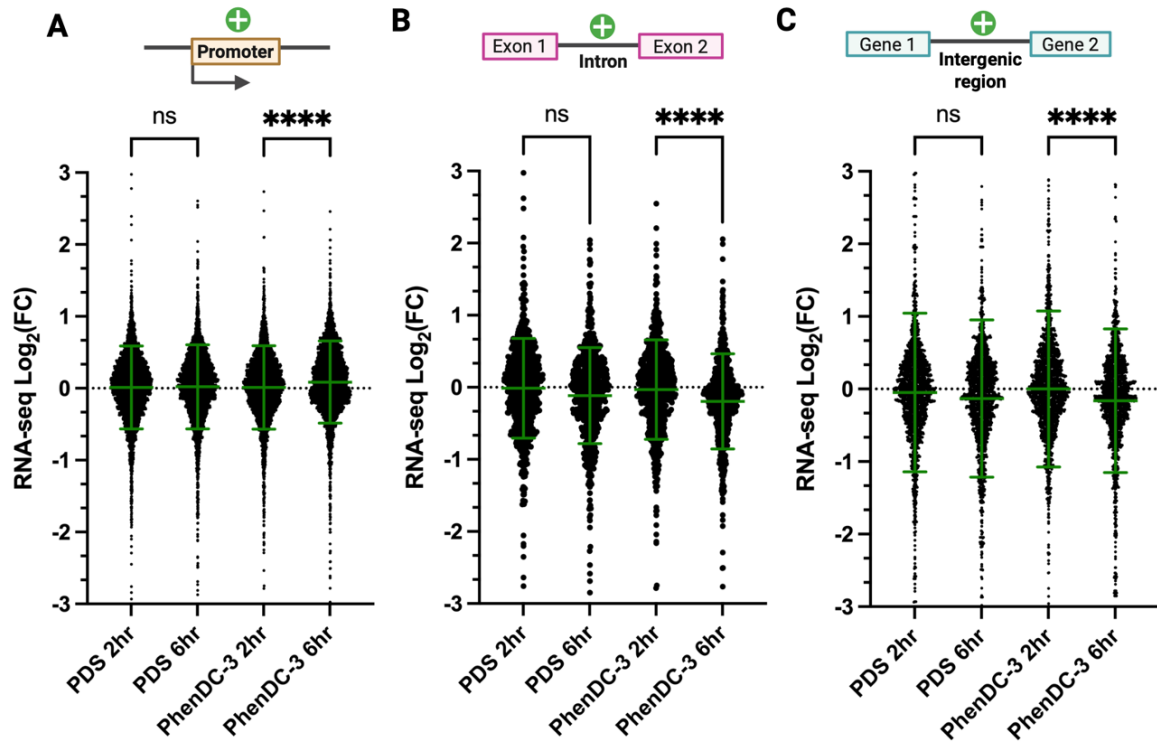


Figure 4.4 Differential expression of genes linked with PEO4 G4s after G4 perturbation. Change in gene expression between native PEO4 and PEO4 treated with either PDS or PhenDC-3 for 2 or 6 hours of genes associated with PEO4 G4 peaks at **A)** promoters, **B)** introns or **C)** intergenic regions. Statistical significance assessed by Mann-Whitney U-test, ns=not significant, ****= $p < 0.0001$. Error bars show standard deviation of gene expression data generated by RNA-seq in triplicate.

A further expansion of this analysis would be to isolate genesets from previously identified enriched pathways associated with G4 formation and assess whether they are preferentially downregulated upon G4 formation. These pathways could include insulin resistance, which was the most enriched pathway, and mTOR and Wnt, two pathways highly reported in the literature to be implicated in acquired HGSOC drug resistance. This would show if genes in these genesets are indeed downregulated upon G4 stabilisation, the extent of the downregulatory effect, and if all or just some of the genes are downregulated. This would provide a deeper insight into how G4 formation may be implicated in transcriptional regulation of these pathways, and at which parts of the

pathway do they have the most prominent effect. Moreover, this analysis would be a more targeted approach into probing the impact of G4 formation and the potential mechanism of G4 stabilisation as a therapeutic avenue of preferentially targeting certain pathways implicated in resistance.

4.2. Pathway enrichment associated with G4 targeting

Pathway enrichment analysis was performed on significantly downregulated genes associated with G4s following G4 ligand treatment, to identify pathways directly impacted by G4 stabilisation. For the following analysis, data from PhenDC-3 treatment was used due to the greater effect on gene expression observed, relative to PDS.

4.2.1. Global pathway enrichment of significantly downregulated genes

Complimenting the trends observed in PEO1 cells regarding gene downregulation over time, there were less enriched KEGG pathways after 6 hours of PhenDC-3 treatment than after 2 hours (**Figure 4.5 A&B**). Relative to native PEO1 cells, after 2 hours of PhenDC-3 treatment, oestrogen, oxytocin and cAMP signalling were significantly downregulated. However, after 6 hours of treatment these pathways were no longer enriched. As alluded to previously, this could be a result of cellular adaptation to overcome the pathway downregulation initiated by G4 stabilisation. For instance, oestrogen and cAMP signalling pathways are linked to ovarian cancer growth and metabolism.(263)(264) Hence, active upregulation of these pathways via alternative regulatory mechanisms may be initiated, to counteract the effect of G4 stabilisation, and promote cell survival and cancer progression.

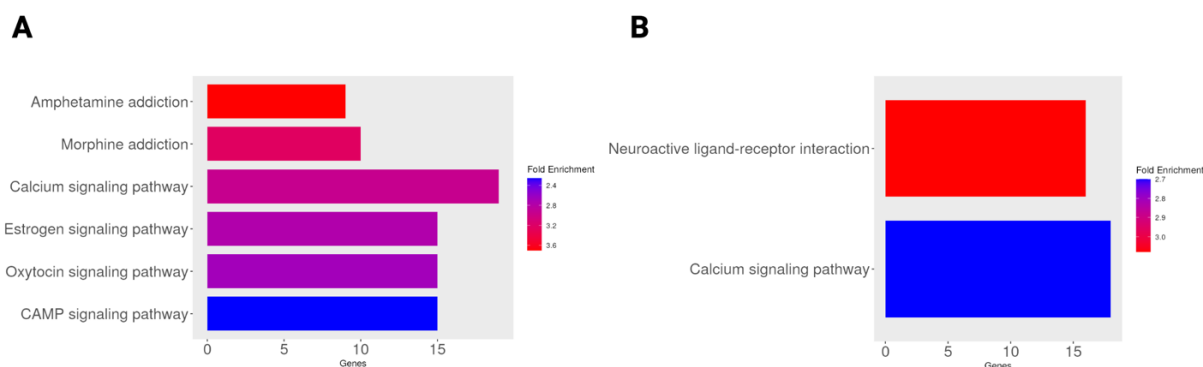


Figure 4.5 Pathway enrichment of downregulated genes in PEO1 after G4 perturbation. KEGG pathway enrichment of significantly downregulated genes associated with G4s in PEO1 cells after **A**) PhenDC-3 treatment for 2 hours or, **B**) PhenDC-3 treatment for 6 hours. Fewer pathways are enriched after 6 hours relative to at 2 hours of PhenDC-3 treatment. Made with iDEP.(219)

No KEGG pathways were enriched when investigating significantly downregulated genes associated with G4s in PEO4 after 2- or 6-hour PhenDC-3 treatment. This could be because there were much fewer genes significantly downregulated after 6-hour PhenDC-3 treatment in PEO4 cells relative to PEO1 (234 and 733, respectively). Potentially, this could be because PEO4 cells have half the amount of G4s, and so it would be expected to have a reduced number of downregulated genes linked to G4s, in comparison to PEO1. Additionally, it could be due to the full potential of G4 stabilisation and gene downregulation having not been achieved at the 6-hour timepoint, as previously speculated. As a result, the lower amount of significantly downregulated genes could have made the enrichment threshold harder to reach.(247)

Nonetheless, GSEA was conducted to investigate enriched hallmark gene sets related to significantly downregulated genes after 6-hours of PhenDC-3 treatment (**Table 4.1**). Notably, enrichment of cell cycle regulating gene sets (G2M checkpoint and E2F targets) suggests that G4 formation in PEO4 cells aids DDR that contributes to drug resistance.(265)(266) Specifically, these genes and transcription factors promote DNA

damage response, cell stemness, and apoptosis evasion, which can contribute to drug resistance in ovarian cancer cells. Cell cycle checkpoint genes are often overexpressed in drug-resistant ovarian cancer to prevent irreparable accumulation of DNA damage that induces apoptosis.(30) Additionally, MYC targets and mTORC1 signalling are two other hallmark gene sets that are significantly downregulated upon PhenDC-3 treatment. Both pathways are well reported for their involvement in acquired drug resistance in ovarian cancer.(267)(268)(46)(212) These findings corroborate the hypothesis that the formation of G4s in PEO4 cells are actively implicated in the upregulation of genes involved in the development drug resistance.

Hallmark gene set	Gene set description	Adjusted p-value
mTORC1 signalling	Genes up regulated through activation of mTORC1 complex.	6.71×10^{-12}
MYC targets	A subgroup of genes regulated by MYC.	5.73×10^{-10}
G2M checkpoint	Genes involved in the G2/M checkpoint, as in progression through the cell division cycle.	3.74×10^{-7}
Androgen response	Genes defining response to androgens.	1.34×10^{-7}
E2F targets	Genes encoding cell cycle related targets of E2F transcription factors.	2.7×10^{-7}

Table 4.1 GSEA enrichment of downregulated genes in PEO4 after G4 perturbation.

The top 5 hallmark gene sets from GSEA of significantly downregulated genes associated with G4s in PEO4, after 6-hour PhenDC-3 treatment. Gene sets are identified by how much the genes present in the hallmark gene set are over-represented in the gene list corresponding to significantly downregulated genes associated with G4s in PEO4, after 6-hour PhenDC-3 treatment.

4.2.2. Regional pathway enrichment in PEO4

To further investigate if the genomic location of G4s in PEO4 cells impacted the association to drug resistance, GSEA was completed for genes correlated to either promoter G4s, or intron and intergenic G4s, collectively. Similarly to global pathway enrichment, the 6-hour PhenDC-3 timepoint was chosen, as this resulted in the greatest effect on gene expression. The hallmark gene sets associated with significantly downregulated genes upon PhenDC-3 treatment in promoter regions have all been previously discussed in their relation to acquired drug resistance in ovarian cancer (**Table 4.2**).⁽³⁷⁾⁽²⁶⁹⁾⁽²⁵⁶⁾⁽⁴⁶⁾ Interestingly, p53 pathways and the interferon responses were significantly upregulated in PEO4 cells, in respect to PEO1, when gaining a G4 in promoter regions. The subsequent downregulation upon G4 ligand treatment supports the notion that G4 formation in promoter regions of PEO4 cells is implicated in the upregulation of these pathways.

Hallmark gene set	Gene set description	Adjusted p-value
Interferon gamma response	Genes up regulated in response to IFNG	7.32×10^{-11}
TNF signalling via NFKB	Genes regulated by NF-kB in response to TNF.	6.49×10^{-10}
Interferon alpha response	Genes up regulated in response to alpha interferon proteins.	8.26×10^{-9}
mTORC1 signalling	Genes up regulated through activation of mTORC1 complex.	1.97×10^{-6}
p53 pathway	Genes involved in p53 pathways and networks.	1.97×10^{-6}

Table 4.2 GSEA enrichment of downregulated genes linked with promoter G4s after G4 perturbation.

The top 5 hallmark gene sets from GSEA of significantly downregulated genes associated with promoter G4s in PEO4, after 6-hour PhenDC-3 treatment. Gene sets are identified by how much the genes present in the hallmark gene set are over-represented in the gene list corresponding to significantly downregulated genes associated with promoter G4s in PEO4, after 6-hour PhenDC-3 treatment.

When investigating hallmark gene sets associated with significantly downregulated genes correlated to non-promoter G4s, the p53 pathway and TNF signalling via NFkB were also enriched (**Table 4.3**). In addition, the early oestrogen response was enriched, which was also significantly upregulated in correlation to the formation of G4s in intron and intergenic regions, relative to PEO1. Genes involved in the response to hypoxia were also significantly downregulated upon G4 ligand treatment. Hypoxia, and the subsequent transcriptional upregulation in response, have been reported to contribute to drug resistance through signal transduction and increased drug efflux.(270)(271) The combined enrichment of these pathways in relation to G4 targeting, and the downstream effects at intron and intergenic relative to promoter sites, highlights the importance of non-promoter G4s in the development of drug resistance.

Hallmark gene set	Gene set description	Adjusted p-value
TNF signalling via NFkB	Genes regulated by NF-kB in response to TNF.	7.81×10^{-11}
Early oestrogen response	Genes defining early response to oestrogen.	6.76×10^{-9}
Myogenesis	Genes involved in development of skeletal muscle (myogenesis).	3.05×10^{-6}
P53 pathway	Genes involved in p53 pathways and networks.	3.05×10^{-6}
Hypoxia	Genes up regulated in response to low oxygen levels (hypoxia).	1.14×10^{-4}

Table 4.3 GSEA enrichment of genes associated with non-promoter G4s after G4 perturbation.

The top 5 hallmark gene sets from GSEA of significantly downregulated genes associated with intron and intergenic G4s in PEO4, after 6-hour PhenDC-3 treatment. Gene sets are identified by how much the genes present in the hallmark gene set are over-represented in the gene list corresponding to significantly downregulated genes associated with intron and intergenic G4s in PEO4, after 6-hour PhenDC-3 treatment.

G4s formed in both promoter and non-promoter regions were associated with significantly downregulated pathways linked to drug resistance upon G4 ligand treatment. These findings reiterate the idea that all G4s, regardless of genomic location, can contribute to transcriptional regulation, contributing to acquired drug resistance.

4.3. Re-sensitisation of PEO4 to cisplatin

To investigate whether the targeting of G4s could impact the cisplatin sensitivity of the PEO cell line pair, Dr Jenna Robinson, a previous PhD student in the Di Antonio group, co-treated the PEO1 BRCA⁺/PEO4 pair with PDS and cisplatin to assess differences in toxicity in the presence or absence of G4 stabilisation. PEO1 BRCA⁺/PEO4 cells were incubated with a single dose of 0.5 μ M PDS, with increasing concentrations of cisplatin over a 72-hour period, to investigate whether cisplatin toxicity changed in the presence of a G4 stabilising molecule. The PDS concentration used was 10-fold below the GI₅₀ value of PDS toxicity within these cell lines, meaning no toxicity is induced from the G4 binder alone, yet only in combination with cisplatin. Notably, there was no significant change in sensitivity of the PEO1 cells to cisplatin in the presence of PDS (**Figure 4.6**). However, PEO4 cells gained a 3-fold increase in sensitivity to cisplatin upon incubation with the G4 stabiliser. PDS has previously been described to induce DNA damage at G4 sites, and so it would be expected that PDS would synergise with a cross-linking drug such as cisplatin.⁽¹⁶⁷⁾⁽¹⁸²⁾ Nevertheless, given that no synergistic effect was observed in the PEO1 cell line, which contained twice as many G4s than in PEO4 cells, the DNA damaging function of PDS does not explain the selective re-sensitisation of PEO4 cells to cisplatin. Therefore, it could be reasoned that it is the stabilisation of the unique G4s in the PEO4 cells, which form in the signalling pathways associated with a drug-resistant cell state, that are responsible for this differential sensitivity. This corroborates findings from the transcriptional changes associated with G4 perturbation, that are also enriched in acquired drug resistance pathways. Therefore, the targeting of G4s in drug-resistant HGSOC cells could offer a selective therapeutic avenue to re-sensitise resistant cells to cisplatin.

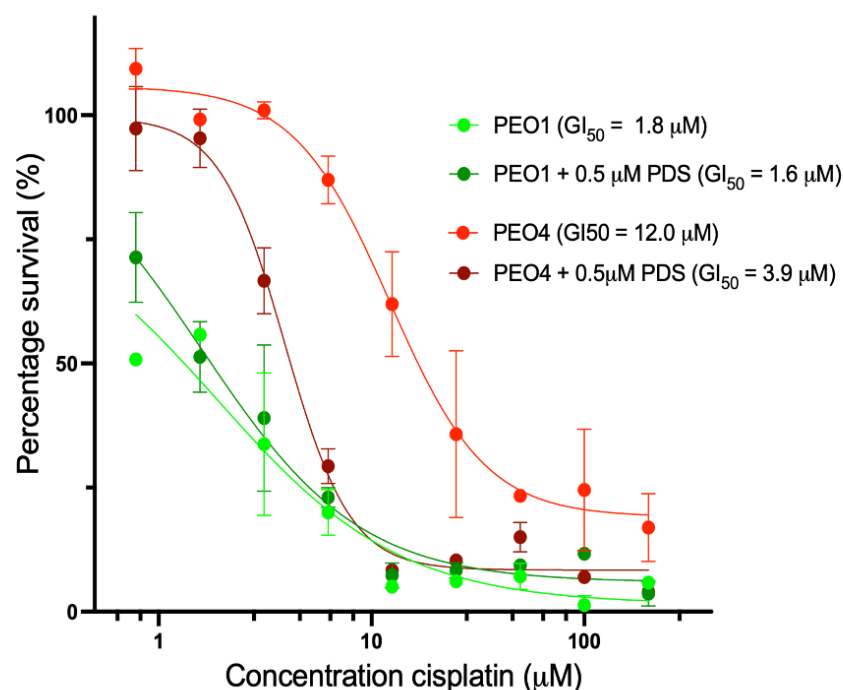


Figure 4.6 Partial re-sensitisation of PEO4 to cisplatin.

Cell viability of the PEO1^{BRCA+}/PEO4 pair to increasing concentrations of cisplatin in the presence or absence of 0.5 μM PDS, showing a 3-fold increase in sensitivity of PEO4 cells but no sensitivity change is observed in PEO1 cells. Error bars are standard deviation of results collected in biological triplicate. Figure adapted from (88).

Whilst this cell viability study was conducted in the PEO1 BRCA⁺/PEO4 pair, obtained from a different source to the PEO1 BRCA⁻/PEO4 pair, it could be reasoned that a similar outcome could be observed in the original PEO1 BRCA⁻/PEO4 pair. This is because the gain in G4 in the PEO4 cell of both PEO pairs had very similar transcriptional effects. Moreover, as it is the PEO1 cell line that has inherent genetic differences, and not the PEO4 cell line, it could be hypothesised that the effect in PEO4 cells would remain the same. Maybe an increase in sensitivity may also be seen in the PEO1 BRCA⁻ cell line, due to the DNA damaging potential of G4 stabilisation, in collaboration with the homologous recombination deficiency of the cell line. Potentially, PhenDC-3 may induce a greater effect in the PEO1 BRCA⁻/PEO4, owing to the greater transcriptional downregulation in PEO4 cells that was observed after PhenDC-3 treatment, than when PDS was used. Cell viability experiments should be conducted in the PEO1 BRCA⁻/PEO4 pair, with increasing

cisplatin concentrations in combination with both PDS and PhenDC-3 to see if the same re-sensitisation effect is observed, and if this is ligand dependent.

4.4. Chapter summary and conclusions

PEO1 and PEO4 cells were treated with G4 stabilising ligands to directly perturb the transcriptional impact that G4 formation has, and to investigate whether this perturbation impacts pathways associated with acquired drug resistance. Using PhenDC-3 resulted in a larger transcriptional change than PDS, with maximal decrease in gene expression change detected after 2-hours in PEO1 cells in non-promoter regions. Contrastingly, minimal change in the average gene expression was measured after 2-hours in PEO4 cells, whilst this was significantly lower after 6-hours. This suggests that the effects of G4 stabilisation require a longer time to be observed, potentially due to the adaptive mechanisms of the cell to resist drug effects.(21) Additional time points after 6-hours should be completed to offer a more comprehensive insight into the transcriptional potential of G4 targeting in PEO4 cells. Encouragingly, treating PEO4 cells with PhenDC-3 results in the significant downregulation of genes associated with gained G4s enriched in pathways canonically associated with acquired drug resistance in ovarian cancer cells. This was observed regardless of whether the G4s were gained in promoter or non-promoter sites, further corroborating the idea that the gain in G4 in PEO4 cells can lead to the transcriptional upregulation of pathways associated with resistance, regardless of the G4 loci. Moreover, in the PEO1 BRCA⁺/PEO4 pair, Dr Jenna Robinson observed that PDS can selectively re-sensitise PEO4 cells relative to PEO1. This offers a potential new therapeutic strategy to combat drug resistance in HGSOC.

5. Validation in an alternative paired HGSOC cell line

5.1. Chapter overview

To assess whether G4-mediated transcriptional regulation of drug resistance was more widely attributed to other paired ovarian cancer cells, or whether it was limited to the PEO1/4 cell line pair, G4 maps were generated in an alternative paired ovarian cancer cell line and intersected with available RNA and ATAC-seq data. The rationale behind this was that epigenetic adaptation, leading to acquired drug resistance, is not a unique mechanism to PEO1/4 cells, and so the role of G4s in this context may also not be conserved to just PEO1/4 cells. Consequently, should the transcriptional regulation associated with gained G4s persist in pathways associated with acquired drug resistance in an alternative cell line pair, this would indicate the fundamental role that G4s may play in gene regulation of resistance related genes. Furthermore, G4s could represent a universal epigenetic regulator of chemotherapy resistance, irrespective of the patient the cells are derived from.

5.2. PEA1 and PEA2 paired ovarian cancer cell lines

To achieve this aim, the patient-derived paired cell lines PEA1 and PEA2 were selected. Similarly to the PEO1/4 pair, PEA1 was taken from a single patient suffering from HGSOC, prior to the development of resistance to platinum-based therapy, and PEA2 was taken after resistance had arisen.(272) However, unlike PEO1, PEA1 are treatment naïve, and were established prior to the patient receiving any chemotherapy treatment (**Figure 5.1**). PEA2 were derived from the patient 6 months later, once platinum-resistance had developed.(273) Additionally, the PEA1/2 pair have no *BRCA1/2* mutations. Hence, they have comparable capacity for homologous recombination and DDR, unlike the original PEO1/4 pair used in this study. Overall, the PEA1 and PEA2 cell lines offered a unique

insight into the G4 landscape between a treatment naïve and a resistant cell line, with no variation in BRCA status, established from a different patient. Hence, the PEA1 and PEA2 cell lines were selected to generate endogenous G4 maps, to see how the distribution of G4s changed between different ovarian cancer cell line pairs, established from different individuals under different circumstances. Additionally, there was publicly available RNA-seq and ATAC-seq data for the PEA cell line pair, which could be intersected with G4 maps, similarly to the analysis conducted with the PEO pair.⁽⁸⁰⁾ This would enable a comprehensive insight into whether the formation of G4s in drug-resistant ovarian cancer cells, and subsequent association with the enrichment of upregulated drug resistance pathways, is limited to the PEO cell line pair, or more widely attributed to other patient derived drug-resistant cell lines.

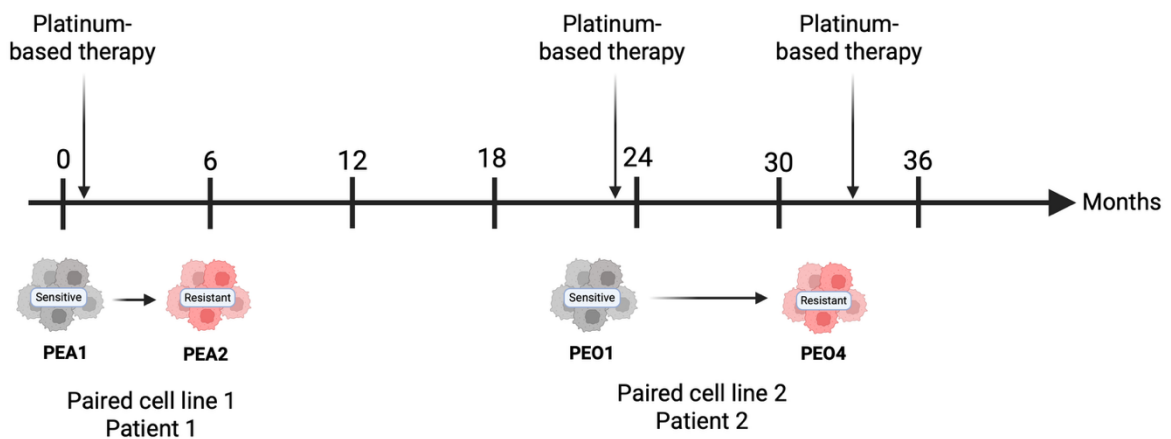


Figure 5.1 Establishment of PEO1/4 and PEA1/2 cell line pairs.

Schematic of the treatment timelines and cell line establishment of the paired patient derived cell lines PEO1 and PEO4, and PEA1 and PEA2.

5.3. Generation of genomic G4 maps in PEA1 and PEA2

Genomic G4 maps were then generated in the PEA1/2 cell line pair, in order to assess if the G4 landscape differed from the PEO1/4 pair. Based on the understanding gained from the generation of G4 maps in PEO1/4 cells, successful production of chromatin fragments

between 100-600 bp was essential to ensure translation of the protocol into PEA1/2 cells. To this end, 2, 4, 6, and 8 million PEA1 cells were cross-linked and sonicated for 15 cycles, before visualising fragment size distribution on the TapeStation. By using a range of cell numbers, it enabled visualisation of the fragment size distribution at different sample concentrations to determine optimal cell number for successful sonication. It was determined that 4 million PEA1 cells was indeed the optimal cell number for cross-linking and sonication for 15 cycles. This is because it had a larger peak within the 100-600 bp range compared to 2 million cells, with minimal high bp tail fragments like with 6 and 8 million cells (**Figure 5.2 A-D**). Additionally, the concentration of the 4 million cell sample was 186 ng/ μ L, which is within the desired range. As the size of the PEA1 and PEA2 cells were similar, 4 million PEA2 cells were also cross-linked and sonicated to produce fragments within the desired 100-600 bp range.

G4 ChIP was performed on the optimally sonicated PEA1 and PEA2 chromatin, and enrichment success determined using qPCR. For both cell lines, optimal enrichment of G4 regions over non-G4 regions was detected, at 6.9-fold for PEA1 and 8.6-fold for PEA2 (**Figure 5.2E&F**). The high enrichment of validated G4-forming regions suggested that the G4 ChIP protocol had been successfully translated into an additional paired ovarian cancer cell line. Two additional biological replicates were then cross-linked, sonicated, and immunoprecipitated to >5-fold enrichment prior to submitting all 3 biological replicates per cell line for library preparation and paired end Illumina NGS at the Francis Crick Institute. One PEA1 biological replicate was generated by Chelsie Li, a PhD student in the Di Antonio group.

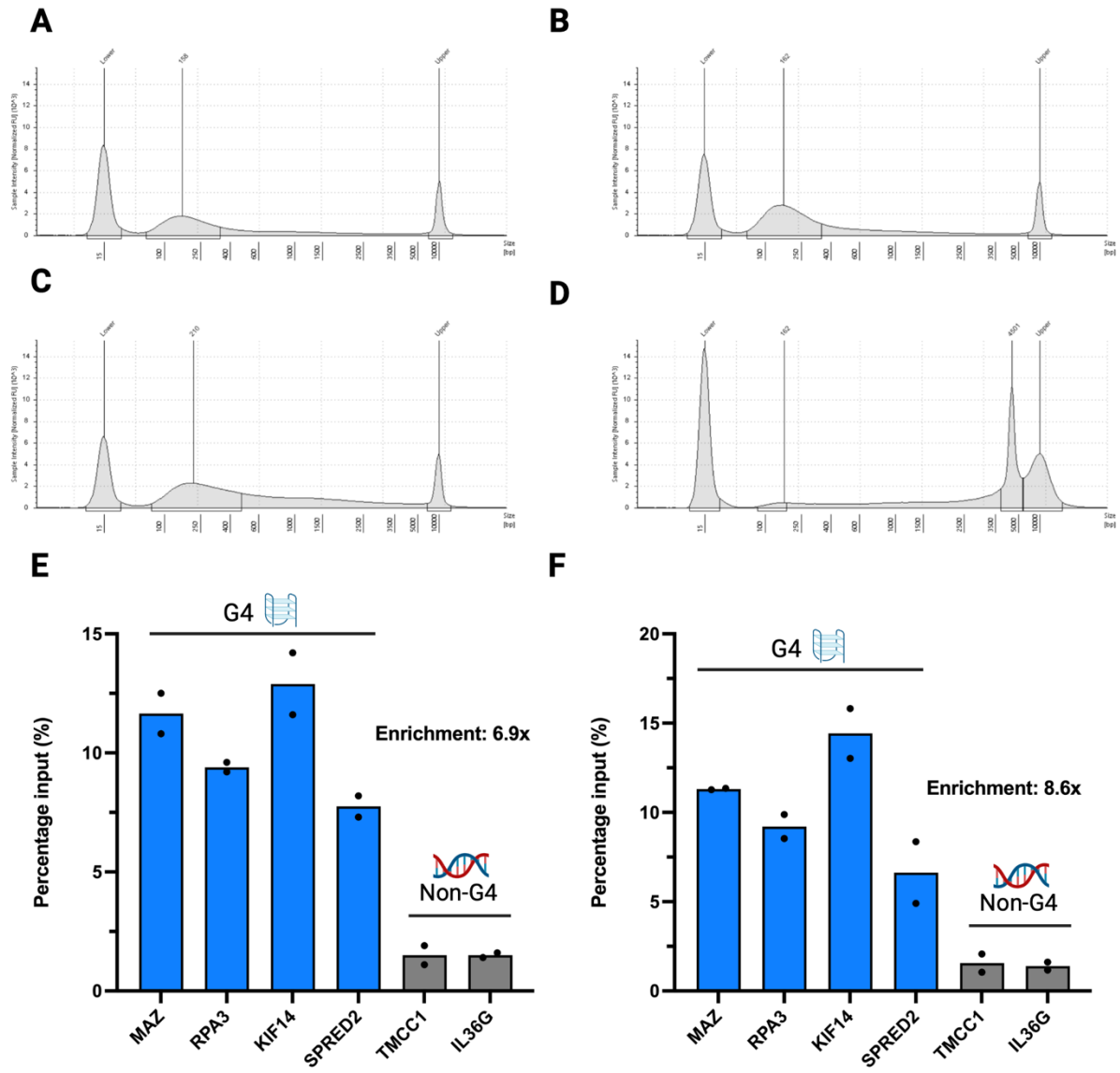


Figure 5.2 Chromatin preparation ChIP-qPCR of PEA1 and PEA2 cells.

TapeStation D5000 profile of **A)** 2 million, **B)** 4 million, **C)** 6 million and **D)** 8 million PEA1 cells after 15 cycles of sonication. **E)** BG4 ChIP-qPCR of the sonicated PEA1 chromatin 6.9-fold enrichment of G4 regions (blue) and non-G4 regions (grey) and **F)** BG4 ChIP-qPCR of the sonicated PEA2 chromatin 8.6-fold enrichment of G4 regions (blue) and non-G4 regions (grey). Percentage input values plotted for qPCR as an average of $n=2$ technical replicates, with individual values for each genomic region represented as a dot.

Raw G4 ChIP NGS *fastq* files were bioinformatically processed as described in chapter 2.5. The alignment details are present in Table 5.1. The quality of the data generated was of a high standard, with minimal read duplication and a sufficient number of uniquely mapped concordant reads for downstream processing.

Condition	Rep	Total no. of mapped reads	Overall read alignment	No. of uniquely mapped reads	Uniquely and concordantly aligned	Read duplication
PEA1- Input ChIP	1	19,514,291	99.22%	16,108,615	82.55%	6.29%
	2	18,748,924	99.25%	15,440,333	82.35%	7.21%
	3	17,189,857	99.19%	14,171,338	82.44%	5.45%
PEA1- G4 ChIP (tech reps merged)	1	38,026,078	96.75%	29,437,489	77.41%	5.73%
	2	39,861,048	96.90%	31,419,188	78.82%	6.13%
	3	35,216,511	98.30%	28,158,064	79.96%	5.45%
PEA1- Input ChIP	1	18,128,098	99.17%	14,910,688	82.25%	6.08%
	2	21,916,604	99.19%	17,972,463	82%	6.02%
	3	17,849,600	99.12%	14,639,600	82.02%	6.12%
PEA2- G4 ChIP (tech reps merged)	1	37,297,530	97.29%	28,875,653	77.42%	5.91%
	2	37,349,927	96.64%	28,525,889	76.37%	5.42%
	3	37,110,503	97.60%	28,704,661	77.35%	5.81%

Table 5.1 Alignment quality of ChIP samples.

Table of total number of reads, overall alignment rate, uniquely and concordantly aligned read number and percentage of total mapped reads and read duplication percentages for PEA1 and PEA2 input and G4 ChIP samples.

Peaks were called using standard *macs2* parameters, and those found in 2/3 biological replicates were taken forward as the finalised G4 peak map. Similarly to the PEO1/4 pair, more peaks were identified in the drug sensitive ovarian cancer cell line, with PEA1

having 7,269 peaks and PEA2 having 6,215 (**Figure 5.3 A**). Both drug-resistant HGSOC cell lines, PEA2 and PEO4, have a similar number of G4 peaks to each other. However, the drug sensitive PEA1 cell line has 5,715 peaks fewer than the other drug sensitive cell line investigated: PEO1. Around 52% of the peaks detected in PEA2 were unique, and termed gained G4s, whilst 4,309 peaks were lost from the drug sensitive PEA1 cell line (**Figure 5.3 B**). This differs to the PEO1/4 pair, in which only 30% of the G4s identified in PEO4 cells were unique compared to PEO1. This could suggest that PEA2 cells have greater epigenetic adaptations from its drug sensitive counterpart (PEA1), than PEO4 does to PEO1, due to the larger abundance of unique peaks. Alternatively, the variation in peaks observed could partly be due to technical variation. This can arise due to differences in sequencing depth and coverage, and hence it is important to generate ChIP samples of a comparable sequencing quality. In the instance of the PEO1/4 and PEA1/2 pairs, the sequencing quality to seems to be similar, with comparable numbers of uniquely mapped concordant reads being obtained in all the cell lines. However, the sample is still reliant on the quality of chromatin sonication and G4 region enrichment, regardless of the comparable quality of the sequencing data. Hence, the technical variation may arise through differences relating to sample preparation. To validate the credibility of the PEA1/2 G4 maps, peaks were intersected with OQS sites that are G4s that had been identified in the human genome *in vitro*.⁽⁹⁴⁾ In all cases, the overlap of PEA1/2 G4 ChIP peaks and OQS sites was 79-85% (**Figure 5.3 C**). This high overlap gave further confidence of the successful immunoprecipitation of G4 structures, and subsequent bioinformatic processing. To further corroborate this, the most enriched motif detected from the PEA1 and PEA2 peaks were identified using MEME suite. A highly G-rich motif was found in 99% of both PEA1 and PEA2 peaks, further authenticating the genomic G4 maps generated in PEA1 and PEA2 cells (**Figure 5.3 D**). Additionally, alignment maps and peak files were viewed in IGV to confirm G4 peak location in a region previously validated in qPCR to enrich for a G4 structure (top), and for a negative region that did not form a G4 (bottom) (**Figure 5.3 E**). Indeed, a peak is visible in all three replicates from both PEA1 and PEA2 in the G4-positive RPA3 promoter, and no in the G4-negative IL36G promoter, and this is reflected in the peak calling.

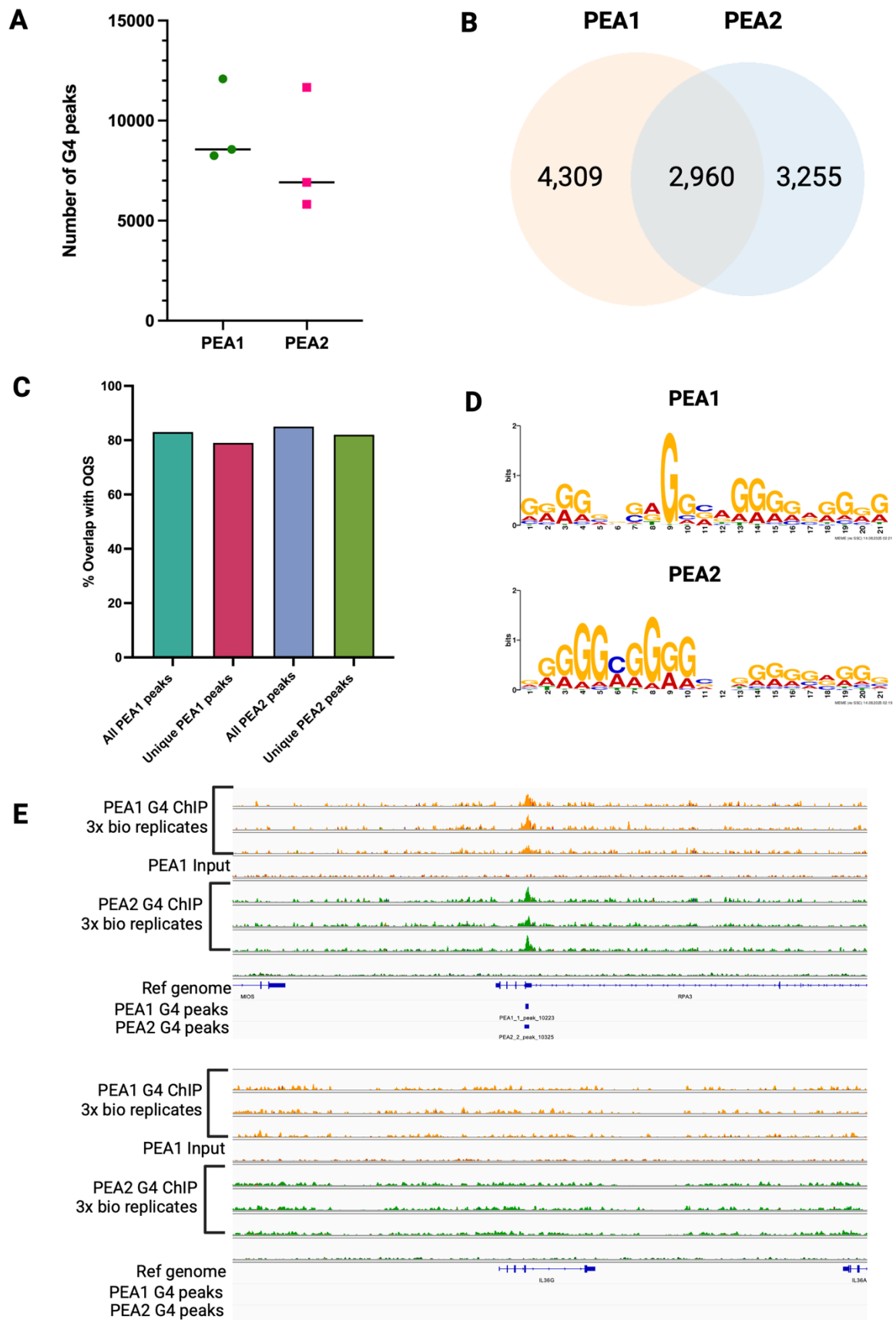


Figure 5.3 G4 peaks and validation in PEA1 and PEA2 cells.

A) G4 peaks called in all 3 biological replicates for PEA1 and PEA2. 8,246-12,090 peaks were identified in the three PEA1 biological replicates and 5,817-11,668 in the PEA2 biological replicates. **B)** G4 peaks unique to PEA1, PEA2, or conserved between the two cell lines- Made with BioInfoRx, **C)** Percentage overlap of G4 peaks with OQS sites of all PEA1 and PEA2 peaks and those unique to each cell line and, **D)** the most enriched motif in the PEA1 (top) and PEA2 (bottom) G4 peaks, showing G-rich enriched motifs. Generated by *MEME suite*. **E)** IGV alignment maps and peak files of PEA1 and PEA2 G4 ChIP and input samples at a G4 forming region RPA3 promoter, and G4 negative region IL36G.

Validated PEA1/2 G4 peaks maps were then annotated using *homer annotatePeaks* to determine their genomic locations. Interestingly, the pattern of distribution of G4 peaks follows a similar trend to that of the PEO1/4 cell lines. The most abundant site of G4 formation was at promoter regions, with 45% of PEA1 G4s, and 54% of PEA2 G4s residing in these loci (**Figure 5.4**). However, when considering peaks unique to either PEA1 or PEA2, this promoter localisation drops by 10% and 7%, respectively. In turn, the abundance of G4s found in intron and intergenic regions increases by 8% in PEA1, and 7% in PEA2. This observation mirrors the G4 distribution seen in the PEO1 and PEO4 cell lines. Whilst promoters are the most abundant sites of G4 formation, when considering unique G4 peaks the distribution slightly shifts towards intron and intergenic regions. These similarities between cell line pairs could indicate that G4s are functioning in a similar capacity, regardless of the patient the cell lines are established from.

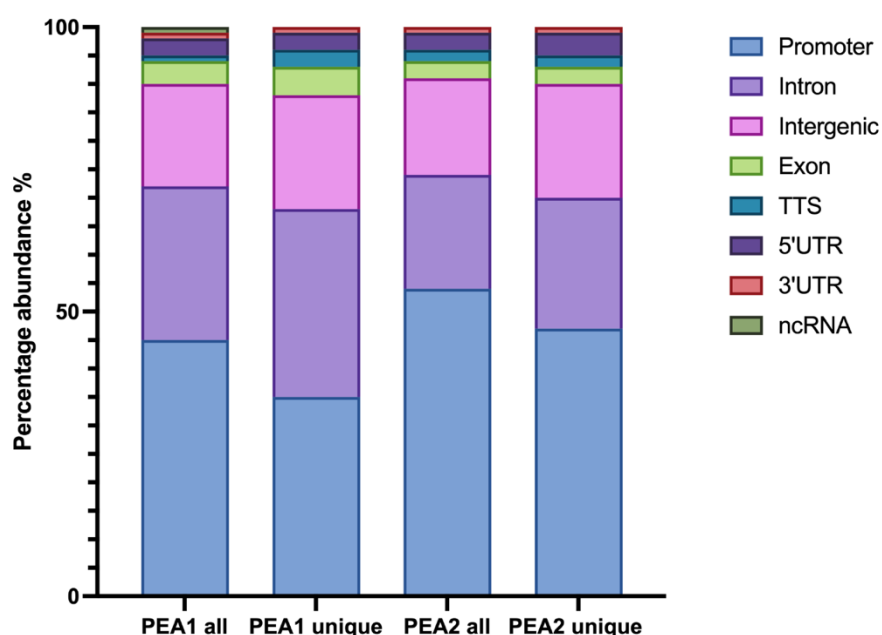


Figure 5.4 Genomic location distribution of PEA1 and PEA2 G4 peaks.

Promoter loci are the predominant site of G4 formation in all those identified in PEA1 and PEA2 cells. This localisation reduces when considering those unique to either PEA1 or PEA2.

5.3.1. Evaluation of the transcriptional effect of a gain or loss of G4 in PEA2

To assess the associated transcriptional effect of gaining or losing a G4 in the drug-resistant PEA2 cells, genomic G4 maps were intersected with publicly available RNA-seq data from the PEA1/2 cell line pair.⁽⁸⁰⁾ Similarly to in chapter 3.2, raw RNA-seq files were bioinformatically processed to identify differentially expressed genes, before being correlated to the nearest unique G4. Differentially expressed genes were also associated with a gain or loss in chromatin accessible site, obtained via publicly available ATAC-seq data from the same publication.⁽⁸⁰⁾ The availability and integration of all these datasets allows comparable insight into how G4 formation is correlated to gene expression changes, relative to changes in chromatin accessibility, like had been previously investigated in the PEO1/4 pair. This would provide insight into whether the associations

observed in the PEO1/4 pair were unique to those cell lines or also detected in an alternative ovarian cancer drug resistance cell line pair.

At promoter loci, the gain in an accessible site in PEA2 cells was associated with a large increase in gene expression ($\text{Log}_2\text{FC}=+2.20$), that was highly statistically significant over a gain in G4 ($\text{Log}_2\text{FC}=+0.77$, $p<0.0001$) (**Figure 5.5 A**). This contrasts with what was observed in PEO4 cells, in which gaining one epigenetic mechanism did not increase gene expression significantly over the other. This suggests that the gain of promoter accessibility has a greater upregulatory effect on gene expression levels than the gain in G4 at these sites, in PEA2 cells. However, G4s at this site are still associated with gene expression upregulation, more so than in PEO4 cells even, and so should not be disregarded for their impact at these loci. Upon expanding the analysis to intron and intergenic regions, this regulatory capacity was reversed. At both intron and intergenic sites, the gain of a G4 in PEA2 cells resulted in a highly statistically significant increase in gene expression, relative to a gain in accessibility (**Figure 5.5 B-D**). This trend mirrors what was observed in PEO4 cells, in which G4s gained in non-promoter sites were also associated with a greater increase in gene expression compared to a gain in accessible site. Moreover, the gain in a PEA2 G4 at intron or intergenic regions also correlates with larger gene expression changes than a gain in promoter G4s (intron $\text{Log}_2\text{FC}=+1.20$, intergenic $\text{Log}_2\text{FC}=+1.28$). This further highlights the importance of non-promoter sites in drug-resistant ovarian cancer cells for transcriptional regulation. Encouragingly, this result suggests that the transcriptional regulatory effect associated with a gain in G4 in drug-resistant ovarian cancer cells is not limited to the PEO1/4 cell line but also can be attributed to the PEA1/2 pair also.

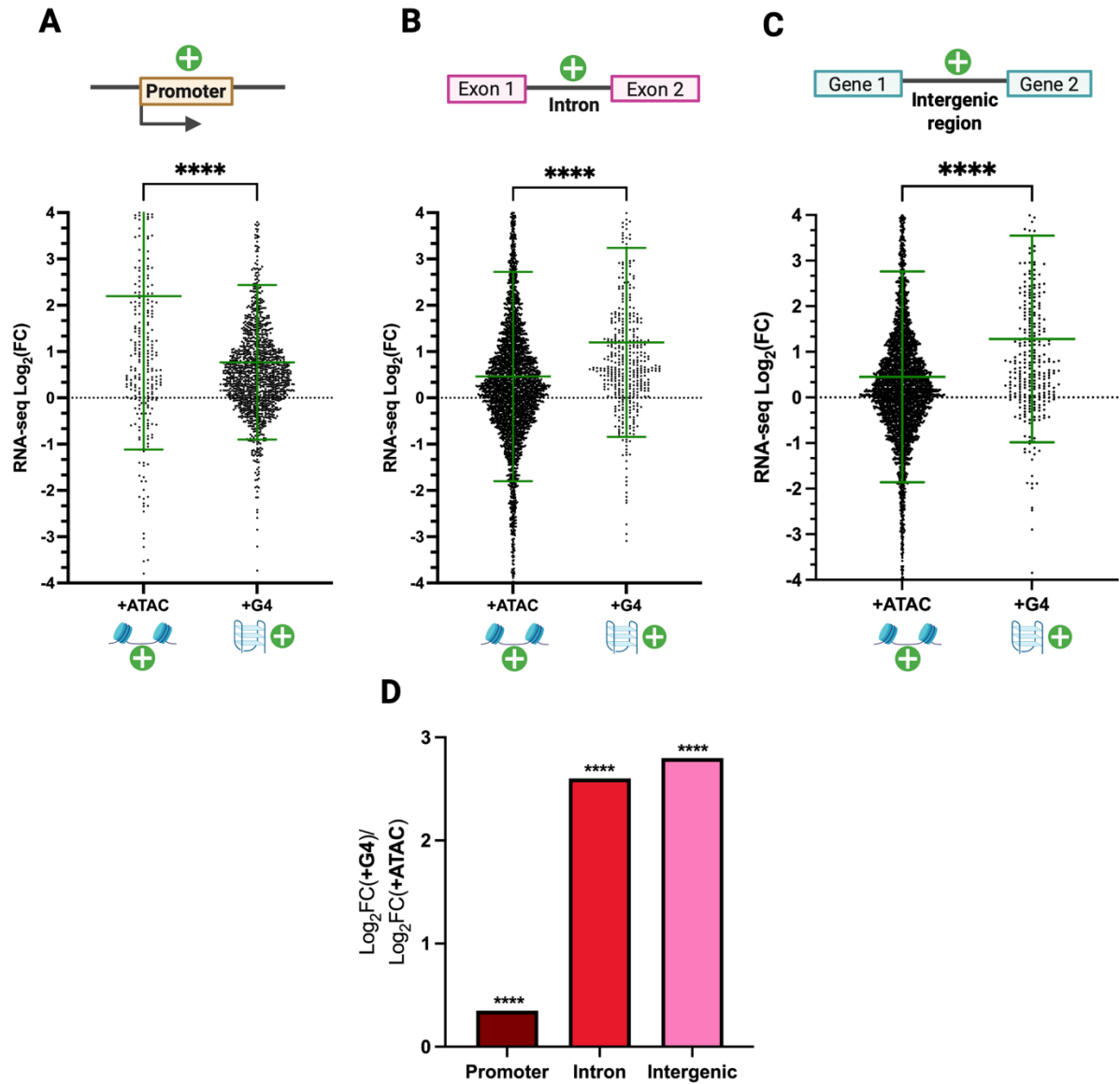


Figure 5.5 Differential expression of genes associated with a G4 gain in PEA2 cells. Change in gene expression between PEA1 and PEA2 of genes associated with a gain in ATAC peaks or G4 peaks at **A)** promoters, **B)** introns or **C)** intergenic regions in PEA2 cells. **D)** The fold-difference in Log_2FC between +G4 and +ATAC from promoter, intron and intergenic regions. Statistical significance assessed by Mann-Whitney U-test, ****= $p < 0.0001$. Error bars show standard deviation of gene expression data in triplicate.

Conversely, the loss of G4 at promoter, intron, or intergenic sites in PEA2 cells did not confer a greater downregulatory effect on gene expression than the loss of an accessible site, at any of these regions (**Figure 5.6 A-D**). This contrasts the effect of gaining either of these epigenetic regulatory mechanisms, in which one always conferred a significantly greater transcriptional impact than the other. In addition, the loss of a G4 at these sites was associated with a lower downregulatory effect than the up regulatory effect correlated with a gain. For instance, in intronic regions, a gain in G4 is associated with $\text{Log}_2\text{FC}=+1.20$ whilst a loss of G4 at these loci is correlated with $\text{Log}_2\text{FC}=-0.30$. This is an observation that can also be extended to promoter sites ($\text{Log}_2\text{FC}=+0.77$ vs $\text{Log}_2\text{FC}=-0.48$) and intergenic regions ($\text{Log}_2\text{FC}=+1.28$ vs $\text{Log}_2\text{FC}=-0.12$). This reinforces the idea that it is the gain in G4 in drug-resistant cells that is contributing more strongly to the transcriptional reprogramming of the cells, rather than the loss of G4 and subsequent gene expression downregulation.

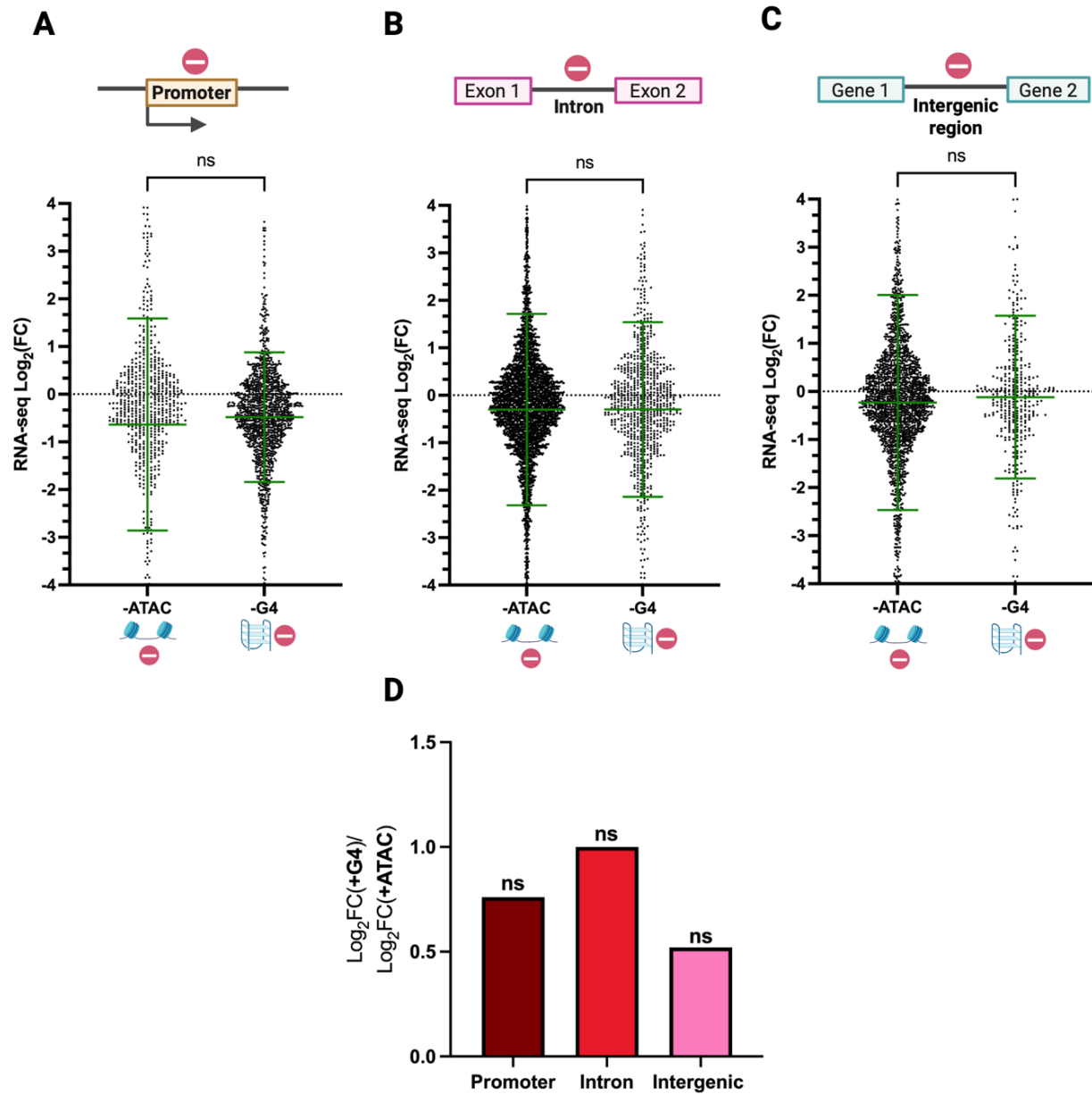


Figure 5.6 Differential expression of genes associated with a G4 loss in PEA2 cells. Change in gene expression between PEA1 and PEA2 of genes associated with a loss in ATAC peaks or G4 peaks at **A)** promoters, **B)** introns or **C)** intergenic regions in PEA2 cells. **D)** The fold-difference in Log₂FC between +G4 and +ATAC from promoter, intron and intergenic regions. Statistical significance assessed by Mann-Whitney U-test, ns=not significant. Error bars show standard deviation of gene expression data in triplicate.

Together with the observations made from the PEO1/4 data, these results indicate that G4s forming in non-promoter regions of drug-resistant ovarian cancer cells harness greater transcriptional regulatory power than a gain in chromatin accessibility, on average. This is not the case upon losing a G4 in these regions, in which there is no significant difference compared to losing a chromatin accessible site. On the contrary, promoter regions in drug-resistant ovarian cancer cells appear to be more variable. Whilst in PEO4 cells the gain in G4 did not result in significantly increased expression, relative to a gain in accessibility, a loss of accessibility resulted in highly significant downregulation compared to a loss in G4. This result was the converse for PEA2 cells, whilst a gain in promoter accessibility results in highly statistically significant gene expression upregulation compared to a gain in G4, the loss of one over the other was not significant. It can thus be determined that G4s represent a novel epigenetic regulatory mechanism, harnessed by drug-resistant ovarian cancer cells to regulate gene expression. Moreover, the importance of non-promoter G4s in gene regulation is further confirmed in this cell line. The trends observed in both the PEO1/4 and PEA1/2 pair indicate that the regulatory capacity of G4s is not limited to just the PEO1/4 pair, but perhaps more widely implicated in drug-resistant ovarian cancer cells.

5.3.2. Pathway enrichment of gained PEA2 G4s

KEGG pathways enriched from significantly upregulated genes associated with a gain in G4 in PEA2 cells were then identified to assess whether the unique G4 enrichment was associated with pathways linked to acquired drug resistance, similarly in PEO4 cells. Interestingly, no KEGG pathways were enriched for significantly upregulated genes associated with a global gain in PEA2 G4, or a gain in promoter PEA2 G4. However, KEGG pathways were enriched for G4s gained in intron and intergenic regions (**Figure 5.7**). Notably, the mTOR signalling pathway is enriched, which was also enriched in PEO4 cells, and has previously been reported to play important roles in cell survival, growth, and proliferation in ovarian cancer.(236) Moreover, the PI3K-Akt signalling pathway is closely associated with mTOR, and is highly dysregulated in ovarian cancer cells, leading to uncontrolled cellular growth and development of chemotherapy resistance.(274) Indeed, targeting the PI3K-Akt/mTOR pathway has been proposed as a mechanism to

increase sensitivity of platinum-resistant ovarian cancer to platinum therapy.(47) This KEGG pathway enrichment analysis further highlights the significance of non-promoter G4s in transcriptional upregulation of pathways associated with acquired drug resistance in ovarian cancer. Even more so due to the lack of enriched KEGG pathways when considering gained promoter PEA2 G4s.

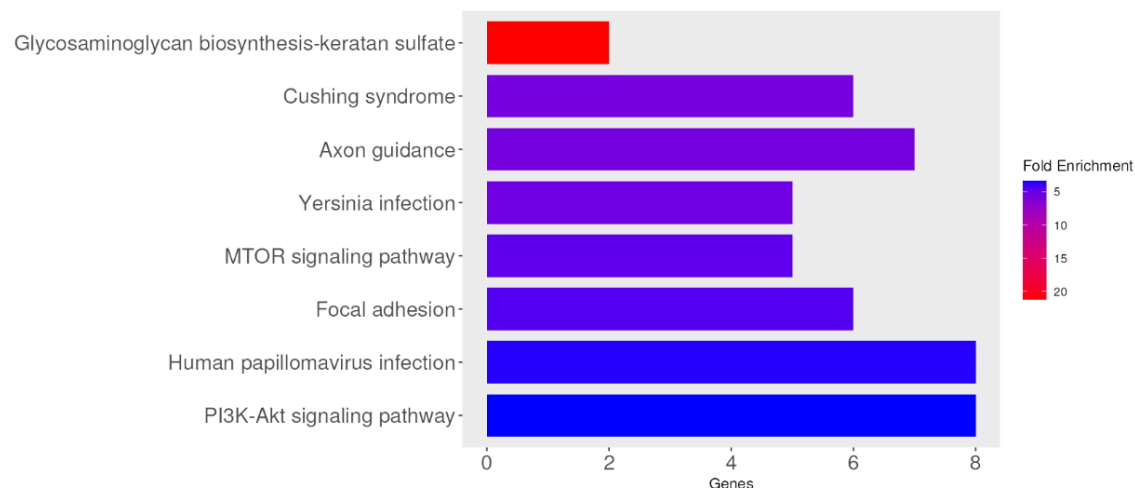


Figure 5.7 Pathway enrichment of genes linked with non-promoter PEA2 G4s.

KEGG pathway enrichment of significantly upregulated genes associated with gained intron and intergenic G4s in PEA2 cells. Made with *iDEP*.(219)

A literature search was also conducted using NDEx iQuery (241), that assesses for gene set similarities across >80,000 publications. When considering significantly upregulated genes associated with a gain in PEA2 G4, in any genomic loci, the publication that had the highest gene set likeness was about TGF- β signalling that contributes to the progression of EMT and metastasis.(275) Even more encouragingly, the next most similar published dataset was from a paper reporting on DNA methylation and transcriptome changes associated with cisplatin resistance in ovarian cancer.(185) This publication also persists amongst the most similar for the gene sets of significantly upregulated genes associated with a gained promoter G4, or a gained intron and intergenic G4 in PEA2 cells. This suggests that the gain in G4, and correlated gene expression upregulation, occurs in

pathways previously reported to be associated with cisplatin resistance. This further supports the hypothesis that drug-resistant ovarian cancer cells are using G4s as a mechanism to leverage a state of elevated gene transcription, to contribute to validated dysregulated signalling of pathways implicated in acquired drug resistance.

5.4. Comparison of acquired G4s in drug-resistant ovarian cancer cells

The gain in G4 in either PEO4 or PEA2 drug-resistant ovarian cancer cells was associated with transcriptional upregulation of pathways linked to acquired drug resistance. To probe if the regulatory effect in each drug-resistant cell line was similar, and if the G4s were conserved, the association with gene expression data and genomic locations were directly compared.

5.4.1. Acquired PEA2 G4s are associated with a greater transcriptional upregulation

To investigate whether gained G4s were associated with similar levels of transcriptional upregulation in drug-resistant ovarian cancer cells, the gene expression changes linked to a gain in either a PEO4 or PEA2 G4 were directly compared. In each genomic loci, a gain in a PEA2 G4 was associated with a significantly greater increase in gene expression than a gain in PEO4 G4 (**Figure 5.8 A**). However, when considering the relative increase in expression associated with a gain in G4 compared to a gain in chromatin accessibility, the gain in G4 in PEO4 cells had a greater relative transcriptional impact than it did in PEA2 (**Figure 5.8 B&C**). This indicates that whilst the gain in G4 in PEA2 cells is linked to higher levels of expression upregulation, other epigenetic mechanisms such as gain of chromatin accessibility also warrant greater levels of transcriptional upregulation than in PEO4 cells. Nonetheless, the transcriptional upregulation linked to gain in G4s in intronic or intergenic regions in either PEA2 or PEO4 cells is statistically significant compared to a gain in chromatin accessibility. This underlines the significance of non-promoter sites in epigenetic transcriptional regulation within drug-resistant ovarian cancer cells.

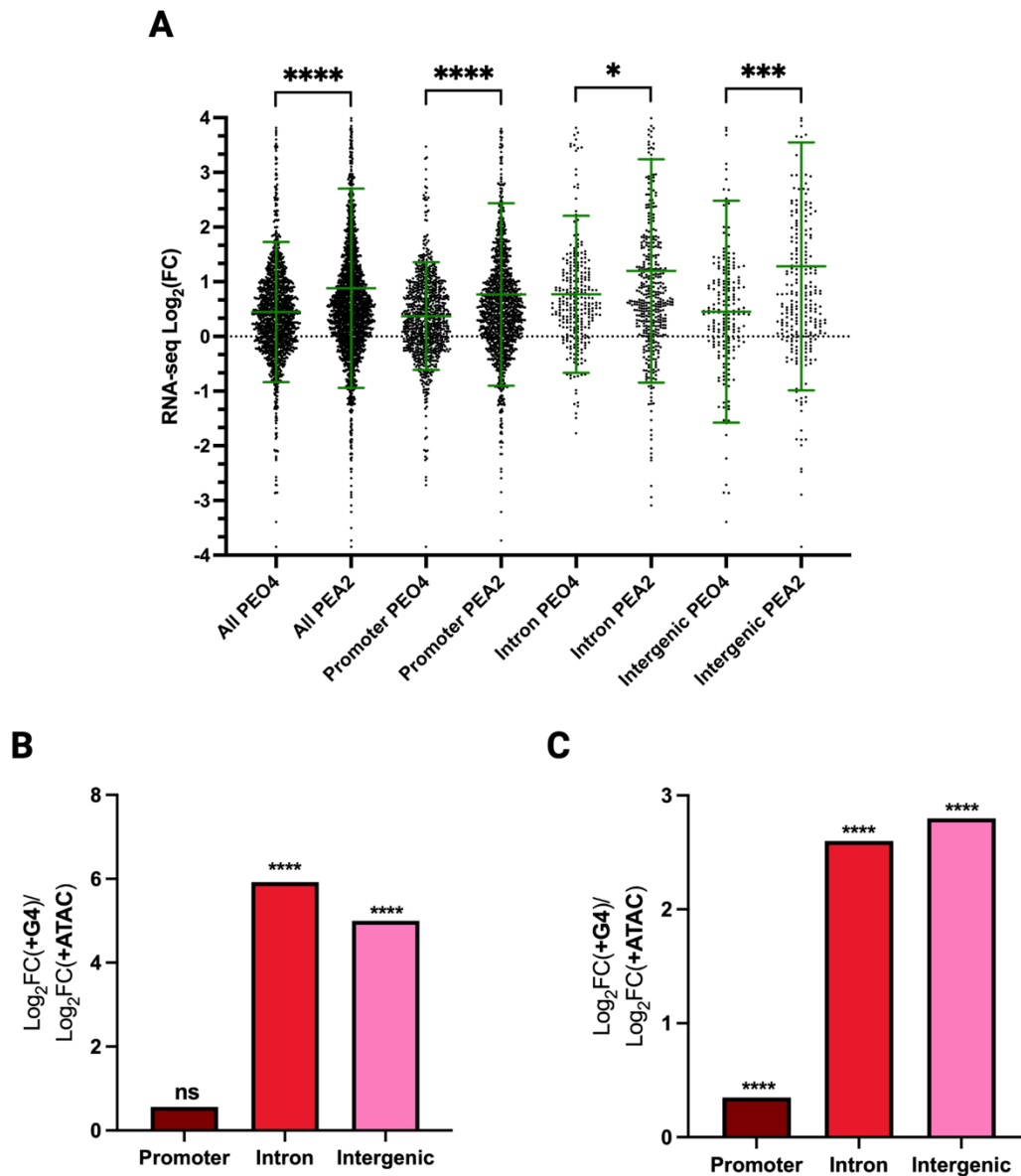


Figure 5.8 Differential expression of genes associated with G4 gain in drug resistance.

A) Change in gene expression between PEO1 and PEO4 or PEA1 and PEA2 of genes associated with a gain in G4 peaks in PEO4 or PEA2 cells globally or at promoters, introns or intergenic regions. **B)** The fold-difference in Log₂FC between +G4 and +ATAC from promoter, intron and intergenic regions in PEO4 cells or **C)** PEA2 cells. Statistical significance assessed by Mann-Whitney U-test, ns=not significant, *=p<0.05, ***=p<0.0001, ****=p<0.0001. Error bars show standard deviation of gene expression data in triplicate.

5.4.2. Conservation of G4s between PEA1/2 and PEO1/4

Similarities between the genomic G4 landscapes were first investigated by comparing read alignment files between the two cell line pairs, using *deepTools plotCorrelation* and the Spearman method. The alignment profiles of the drug-resistant cell lines were not deemed to be anymore similar than they were to the drug sensitive cell line of their opposite pair (**Figure 5.9 A**). For instance, PEA1 and PEO4 had a correlation coefficient of 0.85 on average, PEO1 and PEA2 had 0.80, and PEA2 and PEO4 were 0.83. This suggests that the G4 ChIP read profiles of drug-resistant cell lines are no more like each other than a drug sensitive and resistant profile from two distinct cell line pairs. It was thought that due to the similar transcriptional effects seen by gained G4s in drug-resistant cell lines that the read profile may be more like each other, and the G4s may form in similar regions. However, as this is not the case from the correlation plot, the effects seen could be due to distinct regions of G4 formation.

To probe this idea, G4 peak files from PEO1 were intersected with those found in PEA1, PEO4 with PEA2, and those unique to PEO4 or PEA2. This was conducted to investigate whether the G4s forming in the drug sensitive cell lines were conserved, or distinct from each other, and the same for the drug-resistant cell lines. Additionally, unique G4 peaks that were identified in either PEO4 or PEA2 were intersected to assess whether these G4s acquired in the drug-resistant state were conserved between drug-resistant cell lines. In the case of all G4s found in either PEA1 or PEA2 cells, around 50% were conserved with the drug sensitive or drug-resistant PEO counterpart (**Figure 5.9 B**). However, when considering G4s unique to either PEO4 or PEA2 (compared to their drug sensitive cell line counterpart), only 8% of these peaks found in PEA2 and 13% of those found in PEO4 were conserved between the two drug-resistant cell lines (**Figure 5.9 B**). This is surprising, owing to the higher overlap associated with global PEO4 and PEA2 G4s. However, IGV tracks looking at key validated regions look like each other (see Figure 3.6 and Figure 5.3), meaning there is not a significant error in data generation for these two cell lines that can be determined via the validated control regions to account for this. One explanation could be that the G4s are indeed in similar regulatory areas to each other but not be at overlapping genomic loci. For instance, PEO4 and PEA2 may both have a G4 in

the same promoter, yet with no overlap in base pair location and hence not be called as a conserved gained peak. This could mean that the G4 is indeed having a similar transcriptional impact in the same genes between the two cell lines, yet this is not detected using the current methodology to determine overlapping peaks. To assess whether this is the case, peaks can be extended either side by a defined base pair number, such as 150 base pair, to extend the peak coordinates. The same intersection can then be conducted, and this time will count proximal peaks within the same regulatory element as conserved peaks. This should be conducted as a way to assess whether the overlap between PEO4 and PEA2 gained G4s is indeed low, or if the G4s forming are simply proximal to each other and so not currently detected. If the number remains low, it could mean that it is simply a small number of G4s that is key for the acquisition of drug resistance, and the other G4s which are not conserved are random formation as opposed to a gene specific driver.

Despite this, a similar enrichment of G4s occurs in drug resistance associated pathways and is coupled with an up regulatory transcriptional effect in both drug-resistant cell lines. This suggests that it might not be the gain of a G4 in an exact genomic location, but rather an accumulation across different locations, that leverage a higher level of gene expression, particularly enriched in pathways that contribute to acquired drug resistance. This is like other epigenetic marks which tend to be dynamic in response to environmental cues and further strengthens the idea that G4s can be considered as epigenetic regulators within cells.

The genomic locations were determined for the G4s conserved between the drug-resistant PEO4 and PEA2 cells using homer annotatePeaks. Interestingly, >70% of the G4s conserved between the two were located in promoter regions (**Figure 5.9 C**). On the contrary, only 20% of the conserved G4s were in intronic and intergenic regions. This distribution remained constant when considering all conserved G4s, or just those unique to the drug-resistant cell lines. This suggests that the sites of formation of non-promoter G4s are more variable than at promoters across different drug-resistant cell lines. However, the formation of these non-promoter G4s is associated with the largest

transcriptional upregulation. Whilst the locations of them may not be conserved in the drug-resistant cell lines, the correlated transcriptional impact and enrichment in reported drug resistance pathways maintained.

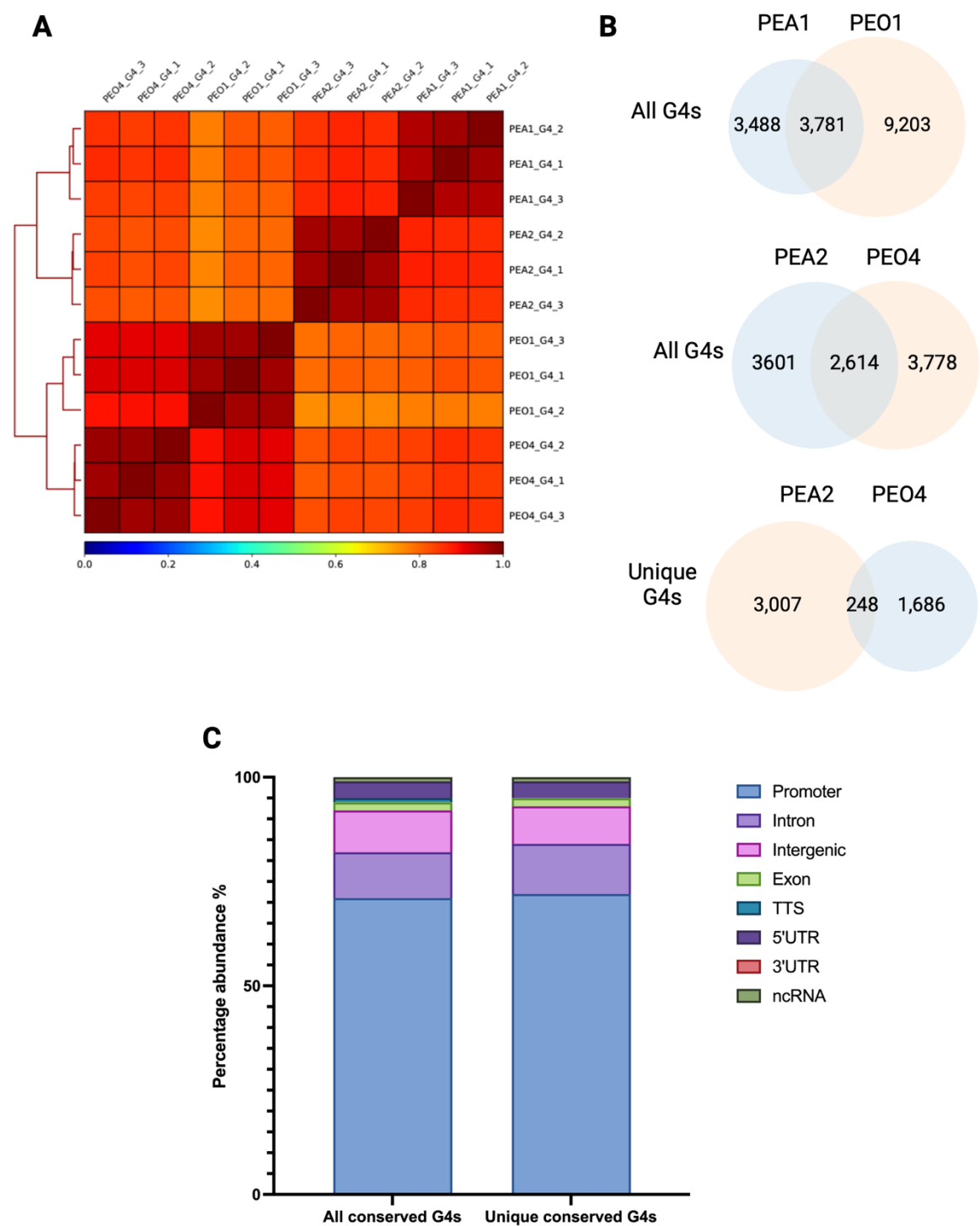


Figure 5.9 Correlation, peak overlap and genomic distribution of G4s.

A) Heatmap Spearman correlation plots of PEO1, PEO4, PEA1 and PEA2 G4 ChIP-seq samples. Created using *deepTools plotCorrelation*. **B)** G4 peaks unique to PEO1, PEA1 or conserved between the two cell lines (top), PEO4 and PEA2 (middle), or unique to PEO4 and PEA2 (bottom) **C)** Genomic location distribution of G4 peaks conserved between PEO4 and PEA2 or unique to both PEO4 and PEA2 relative to their drug sensitive counterparts.

5.4.3. Pathway enrichment of conserved PEO4 and PEA2 G4s

Despite the low number of conserved gained G4s between the two drug-resistant cell lines, pathway enrichment analysis was conducted to assess whether these conserved G4s were indeed enriched in drug resistance pathways. No KEGG pathways were enriched, potentially owing to the low number of gained G4s conserved between PEO4 and PEA2. Instead, GSEA was performed to identify any enriched hallmark gene sets that these G4s may be located in. The p53 pathway, mTORC1 signalling, and hypoxia were all identified as enriched hallmark gene sets (**Table 5.2**). These were all pathways that were associated with significantly downregulated genes after G4 ligand treatment in PEO4 cells, and that have been reported to be linked with drug resistance in ovarian cancer. Of note, mTOR and mTORC1 signalling were consistently enriched through KEGG and GSEA hallmark gene sets for gained G4s in PEO4 and PEA2 and conserved between the two. These signalling pathways have been well reported for their involvement in ovarian cancer cell proliferation, DDR, autophagy suppression, and metabolism, which can influence cellular sensitivity and resistance development to cisplatin.(46)(276) Inhibition of mTORC1/2 signalling has been reported as a novel strategy for re-sensitisation of platinum resistant ovarian cancer cells.(212)(277) Potentially, regulation of this pathway through the suppressive targeting of G4s enriched in these pathways could provide an additional therapeutic strategy for drug-resistant ovarian cancer. This concept would need to be further probed to assess how large a contribution G4s are having to the upregulation of these pathways.

Hallmark gene set	Gene set description	Adjusted p-value
P53 pathway	Genes involved in p53 pathways and networks.	1.67×10^{-5}
Unfolded protein response	Genes up regulated during unfolded protein response, a cellular stress response related to the endoplasmic reticulum.	4.04×10^{-5}
Glycolysis	Genes encoding proteins involved in glycolysis and gluconeogenesis.	1.3×10^{-4}
mTORC1 signalling	Genes up regulated through activation of mTORC1 complex.	1.3×10^{-4}
Hypoxia	Genes up regulated in response to low oxygen levels (hypoxia).	8.87×10^{-3}

Table 5.2 GSEA enrichment of conserved G4s in PEO4 and PEA2.

The top 5 hallmark gene sets from GSEA of genes that gained a G4 in both PEO4 and PEA2 cells. Gene sets are identified by how much the genes present in the hallmark gene set are over-represented in the gene list corresponding to genes that gained a G4 in both PEO4 and PEA2 cells.

5.5. Chapter summary and conclusions

In this chapter, the role of G4s in acquired drug resistance was investigated in an alternative ovarian cancer cell line pair to PEO1/4. Indeed, the same trends observed in PEO4 were mirrored in PEA2. G4s gained in non-promoter regions were associated with a statistically significant increase in gene expression, relative to a gain in chromatin accessibility in the drug-resistant cell lines. This association was not replicated upon G4 loss in PEA2, in which there was no significant difference between losing a G4 or accessible site. Additionally, the gain in G4s were enriched in significantly upregulated pathways previously linked to drug resistance. These observations were identical to that seen with PEO1/4, highlighting that the association of gained G4s with drug resistance pathways isn't limited to just that cell line pair.

To investigate whether the gained G4s were conserved across drug-resistant ovarian cancer cell lines, unique G4s to PEO4 or PEA2 were intersected. Interestingly, few gained G4s were conserved between the two. This may suggest that the exact location of G4 formation in drug-resistant ovarian cancer cells is not conducive to transcriptional upregulation of genes in drug-resistant pathways, but rather the enrichment of them in such pathways. Alternatively, it could suggest that a few select G4s are vital for the transcriptional regulation of a few key genes implicated in resistance, similarly to how methylation of a few key genes can lead to the acquisition of drug resistance. This further strengthens the idea that G4s can be considered to be emerging epigenetic marks in the context of drug-resistant ovarian cancer. Particularly given the low overlap of gained G4s between the cell lines, yet similar association with gene expression changes. Therefore, it is difficult to ascertain whether it is the global shift in gene expression pattern associated with a gain in G4 in a drug resistant HGSOC cell line, or the gain in a subset of specific G4s at particular loci that can be implicated in the acquisition of resistance. Moreover, it could be possible that G4s are indeed being gained within the same genes, and just that the exact base pair loci are not identical between the two cell lines. For instance, a G4 may be gained in the same promoter, but not at the same loci and hence not deemed to be a shared peak whilst still forming in the same regulatory element. To investigate whether this is the case, or if there are just few gained G4s conserved between the cell lines,

additional analysis to extend the genomic length of the peaks should be conducted to capture any proximal G4 formations in the same regulatory elements. This will help to investigate the true conservation of G4s in regulatory elements of genes implicated in acquired drug resistance, to further probe whether it is a small subset of G4s that could be implicated, or a more global gain across these pathways in which exact genomic loci is not important.

Whilst there are similarities in the cell line pairs gene expression profiles associated with a gain in G4 across promoter, intron, and intergenic sites, there were also differences in G4 distribution observed. The drug-resistant cell lines, PEO4 and PEA2, had similar numbers of G4s, however the drug sensitive cell lines, PEO1 and PEA1 had a difference in total G4 prevalence of around 5,000. This could be a technical artefact associated with the G4 ChIP-seq profiling, or it could be of true biological relevance and PEA1 cells indeed have less G4s. PEA1 is a treatment naïve HGSOC cell line, whilst PEO1 was derived from a patient after chemotherapy treatment, whilst the cells were still responding to therapy. Based on the G4 numbers obtained, it could be plausible that chemotherapy treatment could increase the number of G4s detected due to the DNA damaging nature of the therapy. This could be one biological explanation for the difference in G4 cell number. However, to fully probe whether this is a true biological difference, or a technical artefact, profiling of G4s in these cell lines using G4 CUT&Tag would be beneficial to negate the potential inconsistencies in enrichment and chromatin sonication. Furthermore, PEA2 cells had more unique G4s compared to PEO4 cells. This could be because PEA2 cells are more epigenetically diverged from PEA1, compared to PEO4 relative to PEO1. Moreover, it could be that PEA2 has more unique peaks that are proximal to each other and in the same regulatory regions and hence are seemingly unique, whilst PEO4 could have more overlapping peaks. This difference in peak number is also intriguing given the low overlap between unique PEA2 and PEO4 G4 peaks, especially when considering that the overlap between all PEA2 and PEO4 peaks is considerably higher. Likewise, this could be as G4 peaks are forming proximally, and so are not considered to be overlapping, or it could well be that PEA2 and PEO4 share few unique peaks, which could be interesting from a biological perspective.

It was also noted that particular pathways were consistently enriched between the PEO4 and PEA2 cells lines. The mTOR signalling pathway emerged when considering gained PEO4, PEA2, or conserved gained G4s between the two cell lines, and also was amongst significantly downregulated pathways upon G4 ligand treatment in PEO4 cells. This pathway is well documented for its involvement in acquired HGSOC drug resistance, and suggests that gained G4s in drug-resistant ovarian cancer cells are occurring in specific pathways to maintain elevated levels of transcription.(278)(277)(212) As such, they can be considered for their therapeutic potential as targets to regulate these pathways and increase sensitivity of drug-resistant HGSOC to treatment.

6. Conclusions and future outlook

HGSOC is the most common and aggressive form of EOC, with 75% of patients being diagnosed at the late stage.(1) Whilst patients initially respond well to first line treatment, roughly 75% of patients will develop resistance to therapy.(21) The acquired resistance, and subsequent disease relapse, is the major cause of fatality at this stage, and as such remains a significant challenge in the treatment of HGSOC.(23) Resistance can develop via multiple molecular mechanisms, and can occur over a relatively short period of time from weeks to months after treatment is started. It can also develop in the absence of major genetic alteration.(21) Thus, non-genetic factors are thought to play a key role in the acquisition of drug resistance in HGSOC patients.(70) These epigenetic modifications act to alter the expression of genes contributing to signalling pathways associated with acquired drug resistance, and include DNA methylation, histone modifications, and changes in chromatin accessibility.(80)(59)(67) These epigenetic regulatory effects can occur at both proximal and distal sites relative to the target gene, through promoter regions, but also distal sites, such as enhancer regions, that can alter the 3D organisation of the genome to bring regulatory elements into proximity.(78)(80)

One emerging epigenetic mark, that has not yet been considered in the context of acquired drug resistance in HGSOC, are DNA G4s.(83) G4s are a non-canonical secondary nucleic acid structure, which form primarily in open chromatin regions, from G-rich DNA sequences.(85) Subsequently, G4s are a change in the DNA structure and not a change to the underlying DNA sequence. G4s have been highly implicated in transcriptional control by acting as transcription factor binding hubs in promoter regions, as well as aiding chromatin looping events to bring together enhancer and promoter regions, to elevate gene expression.(87)(232) Furthermore, G4s are highly dynamic structures, which have the capacity to fold and unfold quickly within cells, as well as being regulated by G4 resolvases (helicases) to remove the endogenous structure. The transient nature of a G4 suggests that it forms for a purpose, before dissipating to reduce interference with regular cellular processes. In the context of epigenetic regulation, this could be to elevate

gene transcription of a target gene to a specific stimulus, such as drug treatment. Whether the G4 preludes elevated transcription, or whether it is leveraged by the cells to maintain a heightened level of transcription is highly debated. However, the formation of G4s in active and gene elevated transcription is utilised by the cell, and so it's role as an emerging epigenetic regulator should not be overlooked. It should also be considered that G4s can work cooperatively with other epigenetic regulators, such as chromatin accessibility and enhancer elements, and so the true impact of an individual regulator can be hard to ascertain. Instead, G4s may contribute to the epigenetic network that regulates altered transcriptional levels in response to drug treatment.

Moreover, G4s are globally enriched in cancer cells relative to healthy cells, and the perturbation of G4s in BRCA-deficient cancer cells has been an effective method of treating cancer cells resistant to cisplatin and PARP inhibitors.(97)(149)(171) However, it has not previously been considered how the G4 landscape may change as a cancer cell evolves. More specifically, how the G4 distribution alters as acquired drug resistance is developed. It could be hypothesised that the G4 landscape will change as the cancer cells acquire resistance, as the G4s are leveraged to maintain an elevated level of gene expression of pathways implicated in drug resistance. Given that acquired drug resistance is a major challenge associated with HGSOC, and that it is highly governed by epigenetic adaptation, HGSOC was deemed to be the ideal model to investigate if alterations in the G4 landscape occur in the development of cancer cell drug resistance.(70)(80)(83) In particular, the project aimed to investigate if G4 formation may be used by HGSOC cells to epigenetically contribute to drug resistance.

First, endogenous G4 maps were produced in the paired HGSOC cell line; PEO1 and PEO4. PEO1 was derived from a patient prior to the acquisition of drug resistance, whilst PEO4 was derived from the same patient after the evolution of drug resistance.(187) The G4 sequencing revealed that whilst promoter loci were the most prominent site of G4 formation in PEO4 cells, those gained at intron and intergenic sites were associated with the largest upregulation of gene expression. Furthermore, the upregulation of gene expression associated with a gained intronic or intergenic G4 was significantly greater

when compared to a gain in chromatin accessibility at these loci, which are reported regions of importance in acquired HGSOC drug resistance.(80) Additionally, this trend was not replicated when a G4 was lost in the drug-resistant PEO4 cells, highlighting that it is the gain in G4 in these cells which is correlated to the largest transcriptional changes, and not the loss and subsequent decrease in gene expression. However, there are limitations associated with the G4 maps generated in this study using G4 ChIP-seq. Firstly, the reliance on optimal sonication to achieve sufficient enrichment of G4-forming regions over non-G4-forming regions can create variability within the G4 maps. The success of the ChIP is tightly linked to the success in generating optimally sheared fragments of a desired concentration. Hence, samples sonicated to within a similar size distribution, and with a G4 enrichment should be generated to ensure a comparison that is more uniform and to compare like for like. Alternatively, profiling G4s using G4 CUT&Tag instead of ChIP could be a viable option to overcome this limitation, as it does not use sonication and enrichment, but instead allows direct tagmentation of DNA at G4 sites. Therefore, this should reduce any bias that arises through the generation of G4 maps using G4 ChIP-seq and generate more robust maps for comparison. Moreover, additional technical replicates can be generated ($n>2$) to account for any technical variation that can occur when conducting G4 ChIP-seq. In this study, two technical immunoprecipitated replicates were generated per biological replicate, as is the standard across published G4 ChIP papers. However, the protocol is reliant on small volumes of enzymes and starting material, and variation in these volumes can generate technical variation between the ChIP samples. Generating additional replicates would negate for these experimental errors, and improve the robustness of the G4 maps.

It is possible that the full extent of intergenic G4s regulatory effects is being underestimated in this study, due to the complexity of associating intergenic regions with target genes. Whilst promoters and introns can often be attributed to single genes, intergenic regions can regulate genes thousands of base pairs away.(217) The annotation used in this study attributes the G4 peak with the closest gene, whilst this may not be the gene that intergenic G4s are regulating via chromatin looping mechanisms. To probe the true impact of intergenic G4 formation, 3D chromatin looping studies to capture the chromatin conformation should be performed to uncover the full transcriptomic effects

intergenic G4s play in drug-resistant HGSOC.(233) This could be investigated via Hi-C, which detects genome wide interactions, in combination with generated G4 maps to see if the sites of contact are at the same loci as G4 formation. This would provide an insight into whether intergenic G4s are at sites that are in contact with distal promoter regions and uncover which genes they may be regulating. Subsequently, this would provide a more accurate idea of the transcriptional impact of intergenic G4s, rather than assuming the transcriptional effect intergenic G4s elicit is at the nearest gene. Moreover, performing 3D chromatin looping studies after G4 perturbation would uncover whether the stabilisation of the intergenic and promoter G4s prevents chromatin looping events that brings the two distal regulatory sites together, and where these interactions are lost.

The observations that G4 gain in PEO4 were associated with increased gene expression was irrespective of the BRCA background of the PEO pair, and the observed trends were recapitulated in a PEO pair in which the PEO1 cell line had functional BRCA.(192) This highlights that the BRCA status of the cells is not impacting the trend in elevated gene expression associated with G4 gain. Indeed, the genetic and epigenetic adaptations are distinct from each other. Despite the prominent transcriptomic effects associated with a gained G4 at non-promoter sites, the gain in G4 at all genomic sites is correlated with a significant upregulation of genes in pathways linked to drug resistance. These include pathways such as mTOR, Wnt, TNF and DDR, which have been reported to be of relevance in CSC, EMT, and reduction of genomic instability that can otherwise trigger cancer cell apoptosis.(46)(49)(38)(255)

To assess whether the formation of a G4 was merely associated to an increase in gene expression in these pathways, or if G4 formation in PEO4 cells plays an active role in maintaining elevated gene expression, perturbation experiments were performed using G4 stabilising ligands. Indeed, the stabilisation of G4s by PDS and PhenDC-3 resulted in an average decrease in gene expression over a 6-hour incubation period. Of note, treating PEO4 cells with PhenDC-3 resulted in the significant downregulation of genes associated with gained G4s enriched in pathways canonically associated with acquired drug resistance in ovarian cancer cells. This suggests that the gain of G4 in these genes is not

just a passenger adaptation but rather being leveraged by the PEO4 cells to sustain increased levels of gene expression of these genes to contribute to the acquisition of drug resistance. To investigate whether G4 perturbation, and the subsequent reduction of expression of pathways associated with drug resistance, could alter the response of PEO4 cells to cisplatin, Dr Jenna Robinson performed a cytotoxicity combination study between cisplatin and PDS in the PEO1 BRCA⁺/PEO4 cell line pair. Whilst the drug-sensitive PEO1 BRCA⁺ cell line showed no difference in cisplatin response when in the presence of the G4 stabiliser, the drug-resistant PEO4 cells showed a 3-fold increase in sensitivity to cisplatin in combination with PDS. As no effect is observed in PEO1 cells, which have 2-fold higher G4 content, this suggests that it is not the DNA-damaging effects of PDS that explains this partial-re-sensitisation. Therefore, it is hypothesised that it could be the formation of G4s in unique genomic locations linked to acquired drug resistance relative the drug-sensitive PEO1 cells that could be responsible. Thus, the transcriptional changes associated with a gain in G4 in PEO4 can be exploited therapeutically to increase sensitivity of resistant cells to chemotherapy agents. Moreover, this combination would be selective for resistant cells, owing to the unique G4 landscape they acquire during the development of drug resistance.

To further validate G4 targeting as a mechanism of re-sensitising resistant cells to therapy, it would be interesting to combine G4 perturbation with other treatments in which resistance has developed. For example, with PARP inhibitors, which have been previously reported to synergise with G4 stabilisers in BRCA-deficient PARP inhibitor resistant cancer cells.(171) In addition, given the observations that show combining cisplatin and PDS in a BRCA-proficient resistant cell line elicits a partial-re-sensitisation to therapy, it could be hypothesised that this combination could be effective in other BRCA-proficient drug resistant cell states. This would exemplify that it is indeed the unique G4 landscape adopted by resistant HGSOc cells that can be exploited to increase sensitivity to therapy, and that a BRCA-deficient background is not a requisite for success. Thus, it implicates the G4s epigenetic regulatory capacity, and not just the DNA damaging potential of G4 stabilising molecules. Moreover, profiling G4 distribution in other paired cell lines that have developed resistance to these other compounds would help uncover if the genomic G4 adaptation is primarily in cisplatin resistant cell lines, or more widely

attributed to acquired HGSOC drug resistance. Previous studies have focused on the DNA damaging capacity of G4 stabilising in this context, and so it would be interesting to see whether G4 targeting is also having an epigenetic regulatory impact to increase the sensitivity of resistant cells. Subsequently, mapping of G4s in a paired cell line with acquired resistance to alternative HGSOC treatments, such as PARP inhibitors, would be a beneficial expansion of this work. The transcriptome can also be profiled after G4 perturbation, as well as investigating if combining G4 perturbation with PARP inhibitors in a BRCA-proficient cells is enough to re-sensitise resistant cells to PARP inhibitors.

Finally, the G4 distribution and transcriptional changes were investigated in an alternative HGSOC patient derived cell line pair; PEA1 and PEA2. This was conducted to assess whether the trends identified in the PEO pair were unique to that cell line or could be more widely attributed to other HGSOC cell lines. This would strengthen the idea that G4s are an epigenetic adaption adopted by ovarian cancer cells in the acquisition of drug resistance and corroborate the suggestion that G4s could be a viable target in the treatment of drug-resistant HGSOC. Gratifyingly, the gain of a G4 in the drug-resistant PEA2 cells, relative to the drug-sensitive PEA1 cells, resulted in similar trends to those observed in the PEO cell line. G4s gained in non-promoter regions were associated with a statistically significant increase in gene expression relative to a gain in chromatin accessibility. Moreover, the gain in G4 in PEA2 cells was also linked to significantly upregulated gene expression of pathways associated with drug resistance, some of which are conserved with PEO4 pathways, such as mTOR. However, there were still only two cell line pairs investigated in this study. To strengthen the conclusions made in this report, additional cell line pairs should be profiled for the G4 distribution and transcriptional profiles ascertained to assess whether the observations made in PEO1/4 and PEA1/2 are valid for additional HGSOC cell line pairs of different genetic backgrounds. Whilst a promising preliminary study, this work can be further improved through expanding the profiling to additional pairs.

The conservation of gained G4s between the drug-resistant HGSOC cell lines, PEO4 and PEA2, was then investigated. Interestingly, few gained G4s were conserved between the

two cell lines, much below the overall conservation of all the G4s. This suggests that it might not be the formation of G4s in exact loci that is conducive to transcriptional changes in drug-resistant HGSOC cells, but more the accumulation of the G4s enriched in pathways that are associated with acquired drug resistance. Alternatively, it could be that only a gain in a few G4s in specific locations are required for driving drug resistance, and that G4 formation is a gene specific effect. This may well be the case, as it has been previously reported that only a few key genes require different methylation states to promote drug resistance, and so it is not illogical to consider that G4s may act in the same way.(58) Moreover, of those few gained G4s that are conserved between PEO4 and PEA2, they are enriched in signalling pathways previously associated with acquired drug resistance in HGSOC. Of note, mTOR signalling is enriched. This pathway is consistently enriched when considering upregulated gene expression associated with G4 formation and has been implicated with cisplatin resistance in PEO4 cells previously.(50) Additionally, targeting of mTOR signalling within these cells increased sensitivity of PEO4 cells to cisplatin. This highlights how G4s are present in key regulatory pathways that have been previously exploited to re-sensitise resistant cells to cisplatin. Hence, it could be considered that the conserved G4s between PEA2 and PEO4 that are forming in this pathway could be a gene specific driver to contribute to drug resistance.

Future work could look to specifically perturb G4-related mTOR signalling, to see if this is sufficient to alter sensitivity to chemotherapy in drug-resistant HGSOC cells, seeing as it is a pathway that persists in both PEO4 and PEA2 cells. To investigate this, conserved G4s within the mTOR signalling pathway could be preferentially targeted using an approach recently developed in the Di Antonio group, called ATENA.(279) This approach uses a G4 stabilising ligand attached to a catalytically inactive Cas9 protein that can be utilised to selectively target G4s in living cells. As such, the conserved G4s in the mTOR signalling pathway could be preferentially targeted, and the sensitivity of these perturbed cells to cisplatin could be determined. This would help uncover whether G4 formation in the mTOR signalling pathway of cisplatin resistant HGSOC cells is a strong contributor to transcriptional upregulation, that contributes to cisplatin resistance. Previous studies have highlighted that the inhibition of mTOR signalling can promote cisplatin-induced apoptosis, hence increasing the sensitivity of resistant cells to cisplatin.(49) Thus, G4s

could provide an alternative target in the regulation of mTOR signalling in HGSOC cells with acquired drug resistance. The specific targeting of this pathway via G4 stabilisation would also provide insight into if the perturbation of just this pathway alone is sufficient to re-sensitise cells to chemotherapy, or if it's the broad targeting of G4s in drug-resistant HGSOC cells that is required to exert an effect. This will help identify whether G4 formation in the mTOR pathway is a gene specific driver of acquired drug resistance in HGSOC cells.

As this study was performed in immortalised patient-derived cell lines, future work should look to see if these changes in G4 landscape are recapitulated in an environment more reminiscent of a tumour. This could involve the profiling of G4 landscapes from primary tumour samples, to assess whether the locations of the G4s gained in the drug resistant PEO4 and PEA2 cells are present in primary tumour cells. Ideally, the profiling of G4 maps could be conducted in primary tumour cells taken before and after the acquisition of drug resistance to provide a deeper insight into how the G4 landscape evolves during the development of drug resistance. This would overcome the limitations associated with using immortalised cell line models and provide stronger evidence for the involvement of G4s in acquired drug resistance. This would help uncover whether the altered G4 landscape, which contributes to drug resistance, is a just feature of HGSOC immortalised cell line models, or if it is also occurring in patient derived tumours. In turn, this would provide a more robust rationale for targeting G4s in the treatment of drug-resistant HGSOC, if the alteration of G4 landscape mirrors what has been identified in PEO4 and PEA2 cells, and is enriched in drug resistance associated pathways. Further expansion of this could then be to use patient-derived xenografts which retain key tumour features such as heterogeneity and the tumour microenvironment, which is implicated in drug resistance development.(280). These models could be used to test whether the re-sensitisation of HGSOC to chemotherapy can be recapitulated by using G4 stabilisers and in combination with chemotherapy, to assess in an environment that retains key biological features of the original tumour. This would expand on the work presented here, that shows transcriptional downregulation of pathways associated with acquired drug resistance coincides with partial re-sensitisation of cisplatin resistant cells when treated with G4 stabilising ligands. If effective, this would provide evidence that G4

targeting should be strongly considered as a therapeutic target to increase the sensitivity of resistant HGSOC to chemotherapy.

Previous work investigating G4 distribution in response to drug treatment and therapy resistance focused on how the G4 landscape is altered using temozolomide in gliomas.(181) Here, the authors found how temozolomide can cause guanine mutations which disrupt G4 structures, suggesting that G4 landscape can be altered due to drug treatment. This may explain why there are differing numbers of global G4 in PEO4 cells relative to PEO1, as they have received an extra round of chemotherapy treatment. This study also proposed that gliomas that are resistant to temozolomide are preferentially growth inhibited upon G4 stabilisation, due to the proposed exploitable vulnerability to guanine-rich DNA that temozolomide has. Indeed, this is complimentary to the re-sensitisation reported in this thesis that shows preferential re-sensitisation of the drug-resistant PEO4 cell line as opposed to the PEO1 cell line to cisplatin. It could be that there is an interplay between the mechanism of epigenetic regulation proposed in this work that is causing the shift in sensitivity, and the exploitation of vulnerability that is exposed after chemotherapy treatment. The temozolomide study did not explore the transcriptional impacts of an altered chromatin landscape, and so the epigenetic adaption cannot be compared beyond the acknowledgment that drug treatment and subsequent resistance results in an altered G4 landscape within the cells. However, it supports the idea presented in this thesis that the G4 landscape does alter as cancer cells evolve in the acquisition of drug resistance.

This PhD project aimed to further the understanding of G4s epigenetic regulatory capacity in the context of cancer evolution. This investigation has demonstrated that the G4 landscape can alter as HGSOC cells acquire drug resistance, and that the formation of unique G4s contributes to upregulation of gene expression of pathways associated with drug resistance. G4 formation plays an active role in this transcriptional regulation, and perturbing the G4 can partially re-sensitise resistant cells to cisplatin. This work requires validation to explore the full extent of G4s regulatory capacity, and to investigate whether this association persists in patient samples. However, it offers a unique perspective into

our understanding of G4s as epigenetic regulators in cancer development and offers a promising target in the treatment of drug-resistant HGSOC.

7. Materials and methods

7.1. Cell Culture

7.1.1. General

U2OS, PEO1, PEO4, PEA1, and PEA2 cells (obtained directly from ECACC) were grown in RPMI 1640 media with Glutamax (Gibco) and supplemented with 10% FBS (Gibco). Growth media was also supplemented with 2 mM Sodium Pyruvate (Gibco), for the PEO1 and PEO4 cell lines. Cell cultures were grown under standard growth conditions at 37°C and 5% CO₂ and routinely checked for mycoplasma. As cells were obtained directly from ECACC, authentication had been certified by STR-PCR.

7.1.2. MTS Cytotoxicity assay

Sensitivity of PEO1 and PEO4 cells to Cisplatin, PDS and PhenDC3, after 24 and 72 hours was measured using an MTS cytotoxicity assay. Cells were plated in a 96-well plate at varying densities depending on cell line. For 24-hour assays, PEO1 cells were plated at 6,000 cells/well, PEO4 at 9,000 cells/well. Cells were allowed to adhere for 24 hours. The following day, fresh growth medium with varying concentrations of the drugs of interest (**Table 7.1**) were added to the cells in a 1:2 serial dilution, leaving one well drug free. After either 24 or 72 hours, media was removed from the wells and replaced with fresh media (100 µL) and MTS reagent (20 µL) (CellTiter 96 AQueous One Solution Cell Proliferation Assay, Promega) and left to incubate under normal growth conditions for 2 hours. The absorbance of the plate was then measured at 490 nm using a CLARIOstar plus microplate reader. The percentage viability was then calculated by working out the relative absorbance of the drugged wells, to that of the control well (no drug) after background plate absorbance subtraction. Curves were then plotted using GraphPad Prism, using the Nonlinear regression (curve fit), inhibitor vs response – variable slope 4 parameter function to generate GI₅₀ values.

Drug name	Starting concentration 24-hour assay	Starting concentration 72-hour assay
Cisplatin	N/A	200 μ M
PDS	20 μ M	150 μ M
PhenDC3	50 μ M	200 μ M

Table 7.1 Cytotoxicity drug and concentrations.

Table of drugs used in MTS cytotoxicity assays and the highest concentration used in the 1:2 serial dilution.

7.2. BG4 Expression and Validation

7.2.1. BG4 expression

TES buffer: 50 mM Tris-HCl, 1 mM EDTA, 20% sucrose, pH 8

Equilibration buffer: 20 mM imidazole in PBS, pH 8

Wash buffer: 10 mM NaCl, 10 mM imidazole in PBS, pH 8

Elution buffer: 250 mM imidazole in PBS, pH 8

Intracellular salt buffer: 25 mM HEPES pH 7.6, 110 mM KCl, 10.5 mM NaCl, 1mM MgCl₂ in Milli-Q water

BL21 DE3 cells (50 μ L, New England Biolabs) were thawed on ice prior to being transformed with the pSANG10-3FBG4 plasmid (2 μ L, 36 ng/ μ L).(96) Transformation was achieved via heat shock, in which the plasmid and cells were first mixed gently by tapping and incubated on ice for 30 minutes, before being heated to 42°C for 10 seconds and left on ice for 5 minutes. Transformed cells were then mixed with SOC outgrowth media (950 μ L, New England Biolabs) for 1 hour at 37°C, 300 RPM. Following incubation, cells were expanded into YT media (50 mL), containing kanamycin (50 mg/ml, Gibco, and 2% glucose), and grown overnight at 37°C, 220 RPM. Of the small culture grown

overnight, 25 mL were expanded into the Invitrogen Magic autoinduction media (1 L) and further incubated for 6 hours at 37°C, 220 RPM, and then for 24 hours at 18°C, 220 RPM.

To generate cell lysate that the BG4 protein could be purified from, the cell suspension was first centrifuged for 30 minutes at 4000 g, 4°C. The supernatant was discarded, and the pelleted cells were dissolved in TES buffer (80 mL). The cell mixture was then treated with protease inhibitor tablets (Pierce protease inhibitor mini, 4 tablets) and left to stir at 4°C until the tablets had dissolved. Cells were then lysed by adding diluted TES buffer (1:5, 120 mL), MgCl₂ (2 mM), and benzonase (5 µL, Millipore), and stirred for 30 minutes at 4°C. Cellular lysate was then centrifuged at for 20 minutes at 8000 g, 4°C, with the BG4 protein being in the soluble fraction.

7.2.2. BG4 purification and quantification

The BG4 protein was purified from the soluble fraction of the cellular lysate using HisPur cobalt resin (2.5 mL, Thermo Scientific). The resin was first washed with Milli-Q water (12.5 mL) and equilibrated with equilibration buffer (5 mL). The washed resin was then continually mixed with the soluble fraction of the cellular lysate to facilitate protein binding to the resin for 30 minutes. The suspension was run through a column, and the resin was washed twice with wash buffer (7.5 mL/wash). BG4 was eluted from the resin following incubation of the resin with elution buffer (7.5 mL) for 15 minutes. The eluted fraction was concentrated using a 10 kDa centrifugal filter (Amicon) and spun down at 4000 g, 4°C until 1 mL remained. The remaining protein elution then underwent buffer exchange with intracellular salt buffer. This was achieved by the addition of intracellular salt buffer (8 mL) to the concentrated elution and subsequent centrifugation until a volume of 1 mL was achieved again. The concentrated BG4 protein was then aliquoted, flash frozen in liquid nitrogen and stored at -80°C.

7.2.3. BG4 quantification

Blocking buffer: 40% ethanol and 10% acetic acid in distilled water

Protein concentration and purity was quantified via SDS-PAGE gel by comparing an aliquot of the final protein stock to a protein ladder (2 μ L, Precision Plus all blue protein ladder), and a 1:2 serial dilution of BSA standards (ThermoFisher Scientific) from 0.3 mg/mL- 0.0375 mg/mL. The BG4 and BSA samples (12 μ L) were mixed with loading dye with SDS (6 μ L, ThermoFisher Scientific) and heated to 75°C for 10 minutes. The ladder and samples were run on a 12% SDS-PAGE gel at 180 V for 45 minutes. The gel was submerged in blocking buffer for 15 minutes and then washed 3 times with distilled water to remove any residual blocking buffer. The gel was then stained overnight with Coomassie blue dye (Bio-Rad), with gentle agitation at 100 RPM. Gel images were taken using an ImageQuant LAS 4000 gel imager, and relative intensity of bands was determined in ImageJ to calculate protein purity and concentration.

7.2.4. BG4 validation using Enzyme-Linked Immunosorbent Assay (ELISA)

Annealing buffer: 10 mM Tris-HCl, 100 mM KCl, pH 7.4

ELISA buffer: 50 mM H₂KPO₄, 100 mM KCl, pH 7.4

Blocking buffer: ELISA buffer with 3% BSA

ELISA buffer with TWEEN: ELISA buffer with 0.1% TWEEN-20

Selectivity of the BG4 protein was determined by an ELISA assay, using a protocol adapted from (96). The experiment was conducted with each oligonucleotide in triplicate, with an additional triplicate of no oligonucleotide wells to account for background absorbance generated by the plate.

Biotinylated oligonucleotides either with the ability to form a G4 (*c-MYC*, IDT) or mutated to no longer form a G4 (*c-MYC mutant*, IDT) had been previously validated by Dr Jenna Robinson, a previous PhD student from the Di Antonio group, using circular dichroism.

These oligonucleotides (500 nM) were first annealed in annealing buffer by heating to 95°C and slowly cooled to 21°C. The annealed oligonucleotides (50 nM, 100 µL/well) were bound to streptavidin-coated 96-well plates (Pierce) through their 5'biotinylated end for 1 hour, shaking at 450 RPM. The wells were washed three times with ELISA buffer (200 µL/well), shaking for 1 minute at 450 RPM for each wash. The washed wells were incubated for 1-hour with blocking buffer (100 µL/well). The blocking buffer was discarded from the wells, prior to the addition of fresh blocking buffer (100 µL/well) to all wells but the first column. To the first column, the highest concentration of BG4 was added (100 µL, 200 nM in blocking buffer), and a 1:2 serial dilution was performed across the remaining columns apart from the last that acts as a no protein control. The plate with BG4 was incubated for 1 hour, shaking at 450 RPM. Post incubation, the plate was washed 3 times with ELISA buffer with TWEEN (200 µL/well), shaking for 1 minute at 450 RPM per wash. Wells were then incubated with an HRP-conjugated anti-FLAG mouse antibody (100 µL/well, 1:10,000 dilution, mixed with blocking buffer, Abcam) for 1 hour and washed using ELISA buffer with TWEEN as described previously. The wells were air dried, ensuring the removal of any bubbles, and the HRP substrate TMB was added to the wells (100 µL/well). After at least 5 minutes, the reaction was quenched with H₂SO₄ (2 M, 50 µL/well), and the absorbance was immediately read.

The absorbance of the plate was measured at 450 nm using a CLARIOstar plus microplate reader. Binding curves were fitted using GraphPad Prism using the non-linear regression (curve fit) binding – saturation one-site total equation to generate K_D values and plots.

7.3. Chromatin Immunoprecipitation and Quantitative Polymerase Chain Reaction (ChIP-qPCR)

Chromatin immunoprecipitation was conducted following an optimised and adapted version of the protocol established by Hänsel-Hertsch *et al.*(99) For each cell line and condition, 3 biological replicates were performed, each containing 1 input sample and 2 technical replicates of the G4 immunoprecipitation.

7.3.1. Chromatin fixation and sonication

Cells were harvested at the optimised cell count per cell line; 4 million U2OS, 6 million for PEO1 cells, 8 million for PEO4 cells, and 4 million for both PEA1 and PEA2 cells. Cells were cross-linked in a suspension (15 mL) with 1% formaldehyde (growth medium mixed with Pierce™ 16% Formaldehyde (w/v), Methanol-free) for 10 minutes at 50 RPM. The cross-linking reaction was quenched using 2 M glycine (1 mL), and the cross-linked cell pellet was washed 3 times using ice cold PBS (10 mL, Sigma-Aldrich) and centrifugation for 5 minutes at 4°C and 2000 RPM. Washed cell pellets were then flash frozen in liquid nitrogen and stored at -80°C before progressing to sonication.

Prior to sonication, cross-linked cell pellets were thawed on ice and resuspended in ice-cold hypotonic buffer (250 µL, Chromatrap) and incubated on ice for 10 minutes. Samples were then centrifuged at 500 g, 4 °C for 5 minutes, before the supernatant was discarded and the remaining pellet was resuspended in lysis buffer (100 µL, Chromatrap). The suspension was incubated on ice for a further 10 minutes, prior to sonication using the Diagenode Pico Bioruptor sonicator. The samples were sonicated for 10 or 15 cycles for 30 seconds on/ 60 seconds off at 4°C. Sonicated chromatin samples were then aliquoted, flash frozen in liquid nitrogen and stored at -80°C.

7.3.2. Chromatin quality assessment

To assess the quality of chromatin sonication and viability for use in the G4 ChIP experiment, the sample concentration and fragment size distributions were evaluated. To establish concentration of the chromatin, the Qubit 4 Fluorometer using the Qubit dsDNA Broad range kit, was used following manufacturer's instructions. The optimal concentration of chromatin was determined to be in the range of 100-300 ng/µL. To validate the success of chromatin sonication to a fragment size of 100-600bp, two methods were implemented across the duration of this project. Originally, an agarose gel was used to separate DNA fragments via electrophoresis, before the distribution was later assessed using an Agilent TapeStation 4200 using a similar electrophoresis principle. For both methods, an initial protein digestion, de-crosslinking and purification step was required to produce purified DNA fragments to allow the assessment of their

size distribution. Sonicated chromatin (5 μ L) was mixed with TE buffer (70 μ L) and RNase A (1 μ L, Ambion™, affinity purified, 1 mg/mL) and incubated for 20 minutes at 37°C, shaking at 1400 RPM. Proteinase K (1 μ L, Invitrogen, 20 mg/mL) was added to the mixture and incubated at 65°C for a further 2 hours before the DNA was purified using Qiagens MinElute Reaction Cleanup Kit, following manufacturer's instructions.

For the agarose gel method of fragment distribution assessment, purified sonicated DNA (15 μ L) was mixed with purple gel loading dye with no SDS (3 μ L, NEB), and loaded into a 2% agarose gel (TopVision agarose and 1x TBE) with SYBR safe dye (ThermoFisher Scientific) along with a DNA ladder (1.5 kb, NEB). The gel was run for 1 hour at 120V. Gels were then imaged on Syngene G box gel imager. For the TapeStation method of fragment distribution analysis, the Agilent D5000 screen tape and reagents were used, following manufacturer's instructions and using the purified DNA. For both methods, successful sonication was deemed to be when the majority of DNA fragments were within the 100-600bp range.

7.3.3. BG4 immunoprecipitation

Intracellular salt buffer: 25 mM HEPES pH 7.6, 110 mM KCl, 10.5 mM NaCl, 1mM MgCl₂ in Milli-Q water

Blocking buffer: Intracellular salt buffer with 1% BSA dissolved, filter sterilised using 0.22 μ m syringe filter prior to use.

Wash buffer: 100 mM KCl, 0.1% TWEEN-20 and 10 mM Tris, pH 7.4, in Milli-Q water, filter sterilised using 0.22 μ m syringe filter prior to use.

For the beginning of the immunoprecipitation reaction, chromatin aliquots were thawed on ice prior to being supplemented with Triton X-100 (Sigma-Aldrich) at a final concentration of 1%. After 10 minutes of incubation at room temperature, chromatin (2.5 μ L) was added to blocking buffer (45 μ L) and RNase A (1 μ L, Ambion™ RNase A, affinity purified, 1 mg/mL) and incubated for 20 minutes at 37°C shaking at 1400 RPM. Post RNA digestion, the input samples were placed on ice, and the corresponding ChIP samples were mixed with BG4 (400 ng) and incubated at 16°C for 1 hour, 1400 RPM. Concurrently,

Anti-FLAG M2 magnetic beads (5 µL/sample, Millipore) were washed three times in blocking buffer (50 µL/sample) using the DynaMag-2 magnetic stand before being resuspended in blocking buffer (50 µL/sample). The washed bead suspension (50 µL/sample) was added to the ChIP samples, following completion of BG4 incubation, and incubated again at 16°C for 1 hour, 1400 RPM. ChIP samples were then washed three times at 4°C using wash buffer (200 µL/sample), and a further two washes were completed with a 10-minute incubation at 37°C, 1400 RPM.

Similarly to the chromatin quality assessment method, a protein digestion, de-crosslinking and purification process was completed to produce purified DNA. The washed beads, with the immunoprecipitated fragments still bound, were mixed with TE buffer (70 µL) and Proteinase K (1µL, Invitrogen, 20 mg/mL). The input samples, which had been on ice throughout the duration of the immunoprecipitation, were mixed with TE buffer (21.5 µL) and Proteinase K (1µL, Invitrogen, 20 mg/mL) to bring the total volume to be equivalent to that of the ChIP samples. Both input and ChIP samples were then incubated for 1 hour at 37°C, shaking at 1400 RPM, and then at 65°C for a further 2 hours before the DNA was purified using Qiagen's MinElute Reaction Cleanup Kit, following manufacturer's instructions. DNA was eluted in 30 µL of Elution buffer, before being diluted to a total volume of 45 µL.

7.3.4. Quantitative PCR and enrichment analysis

qPCR analysis was conducted to ensure enrichment of regions validated to contain G4 structures, over non-G4 regions as an indicator of successful immunoprecipitation. The primers used targeted 4 G4 regions; *MAZ*, *RPA3*, *KIF14* and *SPRED2*, and 2 non-G4 regions; *TMCC1* and *IL36G*, with the primer set sequences provided in **Table 7.2**.

Region name	Forward primer sequence (5'-3')	Reverse primer sequence (5'-3')
<i>MAZ</i>	ACT CAG CGC AGG ATT GTA AAT A	CCT CAT GCT TCG GCT TCC
<i>RPA3</i>	CGG AAG TTG ACA GAT ACA GGG	GAT CGC AGA AAG GTA GTC TCA G
<i>KIF14</i>	CGG TAG CCG TCT CTG AAT G	CTT TAG CAG AAC CCG AGG AG
<i>SPRED2</i>	AAC AGG AGG AGG AAG TAG GG	TTT CGG TCG CAA GTA GGA AG
<i>TMCC1</i>	GTG GTA CAC TGC CTA CAG TAT T	GTA TAA CGC CTG GGC TAT GT
<i>IL36G</i>	GCC CAC CTC TTT ACT TCC TTA	AAC ACT CTT TCA GCT CCA TCC

Table 7.2 Primer sequences used in ChIP-qPCR.

Primer sets for G4 and non-G4 regions, used in qPCR analysis of G4 ChIP. Table adapted from (99).

Purified DNA (2 µL) was mixed with Fast SYBR Master Mix (10 µL) and forward and reverse primers of the respective loci (4 µL of each, 200 nM, table x). An Agilent Technologies Stratagene Mx3005P Real-time PCR machine was used to run the following thermal programme:

Cycle 1: 95 °C for 20 seconds

Cycle 2: 95 °C for 3 seconds

Cycle 3: 60 °C for 30 seconds

Cycles 2-3 were then repeated 40 times

The number of cycles required for input and ChIP samples to reach a given fluorescence signal threshold calculated internally within the qPCR machine was termed to be the C_t value of the sample. To calculate the fraction of recovery of the immunoprecipitated samples (% Input) for each region, the C_t values of the ChIP samples were first normalised to the corresponding input samples using the following equation:

$$\% \text{ Input} = 2^{C_t(\text{input}) - C_t(\text{ChIP})} \times 100$$

The signal-to-noise-ratio (fold enrichment) was then calculated from the average of the % Input values for G4 regions vs non-G4 regions using the following equation:

$$\text{G4 enrichment} = \% \text{ Input}(\text{G4 regions}) / \% \text{ Input}(\text{non - G4 regions})$$

Successful immunoprecipitation was considered to be when the % Input recovery of G4-regions was $\geq 5\%$, and their relative enrichment over non-G4 regions was ≥ 5 -fold. Sequencing of ChIP and input samples was performed by the advanced sequencing facility at the Francis Crick Institute on a NovaSeq S2 for 200 cycles. Libraries were constructed by the facility using the NEBNext Ultra II DNA Library Prep Kit for Illumina. Final libraries were sequenced using 100 bp paired-end sequencing at 30 million reads/sample.

7.4. RNA-sequencing

7.4.1. RNA extraction

RNA was extracted from PEO1 and PEO4 cells (3.5×10^6 cells per biological replicate) using Qiagen's RNeasy extraction kit, following the manufacturer's instructions, either after no drug treatment, 2 or 6 hours of PDS or PhenDC3 treatment. Extraction was

performed in biological triplicate for each condition. In short, cells were lysed using the RNeasy kit lysis buffer, and the lysate was homogenized using the QIAshredder column. Cell lysate then underwent ethanol precipitation, before adding to RNeasy spin columns to remove impurities and for on column DNase I digestion using the RNase-Free DNase Set (Qiagen). The RNA was then purified by washing three times using a series of RNeasy wash buffers, before being eluted in RNase free water (50 µL).

7.4.2. RNA quality assessment and sequencing

The quality of the RNA was then assessed on an Agilent TapeStation 4200 using the RNA ScreenTape, following manufacturer's instructions. RNA samples that had a RINe score of 7 or higher were considered for sequencing, to ensure for minimal degradation. RNA sequencing was performed by Novogene, which is an external sequencing provider, who used polyA enrichment for extraction of mRNA, before performing 150 bp paired-ended NGS at 25 million reads/sample. Samples were sequenced on a NovaSeq S2 for 200 cycles. Raw fastq files for each sample were received ready for bioinformatic processing.

7.5. ATAC-seq

7.5.1. Profiling chromatin accessibility

Lysis buffer: 10 mM Tris-HCl, pH 7.4, 10 mM NaCl, 3 mM MgCl₂, 0.1% IGEPAL CA-630

Transposition reaction mix (per one reaction): 25 µL TD (2x reaction buffer), 2.5 µL TDE1 (Nextera Tn5 transposase), 22.5 µL Nuclease free H₂O. All reagents from Illumina Tagment DNA Enzyme and Buffer Small Kit.

Transposed DNA PCR mix: 10 µL Transposed DNA, 10 µL Nuclease free H₂O, 2.5 µL Custom Nextera PCR Primer 1, 2.5 µL Custom Nextera PCR Primer 2 (containing the barcode), 25 µL NEBNext High-Fidelity 2x PCR Master Mix

Chromatin accessibility was profiled using an established protocol from Buenrostro *et al.*(62) Experiments were conducted in biological triplicate. Briefly, 50,000 cells per

biological repeat were harvested and lysed with lysis buffer to produce a crude nuclei preparation. Nuclei were then resuspended in transposition reaction mix and incubated at 37°C for 30 minutes. After transposition, transposed DNA was purified and eluted in Elution buffer (10 µL) using the Qiagen MinElute PCR Purification Kit, following manufacturer's instructions. Purified DNA then underwent PCR amplification using customised barcoded primers (**Table 7.3**) following the thermal cycle detailed below:

1x cycle at 72°C for 5 minutes, 98 °C for 30 seconds.

8-10x cycles of 98°C for 10 seconds, 63°C for 30 seconds, 72°C for 1 minute.

Amplified library was then purified using the Qiagen MinElute PCR Purification Kit and eluted in 20 µL of elution buffer.

Primer name	Primer sequence (5'-3')
Ad1_noMX (Primer 1)	AATGATACGGCGACCAACGAGATCTACACTCGTCGGCAGCGTCAGATGTG
Ad2.1 (Primer 2)	CAAGCAGAAGACGGCATAACGAGATT TCGCCTTAGTCTCGTGGGCTCGGAGATGT
Ad2.2 (Primer 2)	CAAGCAGAAGACGGCATAACGAGAT CTAGTACGGTCTCGTGGGCTCGGAGATGT
Ad2.3 (Primer 2)	CAAGCAGAAGACGGCATAACGAGATT TCTGCCTGTCTCGTGGGCTCGGAGATGT
Ad2.4 (Primer 2)	CAAGCAGAAGACGGCATAACGAGAT GCTCAGGAGTCTCGTGGGCTCGGAGATGT
Ad2.5 (Primer 2)	CAAGCAGAAGACGGCATAACGAGAT AGGAGTCCGTCTCGTGGGCTCGGAGATGT
Ad2.6 (Primer 2)	CAAGCAGAAGACGGCATAACGAGAT CATGCCTAGTCTCGTGGGCTCGGAGATGT
Ad2.7 (Primer 2)	CAAGCAGAAGACGGCATAACGAGAT GTAGAGAGGTCTCGTGGGCTCGGAGATGT
Ad2.8 (Primer 2)	CAAGCAGAAGACGGCATAACGAGAT CCTCTCTGGTCTCGTGGGCTCGGAGATGT
Ad2.9 (Primer 2)	CAAGCAGAAGACGGCATAACGAGAT AGCGTAGCGTCTCGTGGGCTCGGAGATGT
Ad2.10 (Primer 2)	CAAGCAGAAGACGGCATAACGAGAT CAGCCTCGGTCTCGTGGGCTCGGAGATGT
Ad2.11 (Primer 2)	CAAGCAGAAGACGGCATAACGAGATT GCCTCTTGTCTCGTGGGCTCGGAGATGT
Ad2.12 (Primer 2)	CAAGCAGAAGACGGCATAACGAGATT CCTCTACGTCTCGTGGGCTCGGAGATGT
Ad2.13 (Primer 2)	CAAGCAGAAGACGGCATAACGAGATAT CACGACGTCTCGTGGGCTCGGAGATGT
Ad2.14 (Primer 2)	CAAGCAGAAGACGGCATAACGAGAT ACAGTGGTGTCTCGTGGGCTCGGAGATGT

Table 7.3 Primer sequences for ATAC-seq.
The barcode sequence is highlighted in bold.

7.5.2. Quality assessment and sequencing

To assess fragment size distribution and concentration, purified libraries were quantified using an Agilent 4200 TapeStation, using the Agilent High Sensitivity D1000 screen tape and reagents, following manufacturer's instructions. A successful transposition reaction was considered to show a distribution that shows a banding, bumping pattern and was then selected for sequencing.

Sequencing of ATAC libraries was performed by the advanced sequencing facility at the Francis Crick Institute. The libraries were pools by the facility, before being sequenced using 100 bp paired-end NGS at 50 million reads/sample on a NovaSeq S2 for 200 cycles. Raw fastq files were received ready for bioinformatic processing.

7.6. Bioinformatics

All code used in the completion of this project and below sections can be found at https://github.com/GemFlint/PhD_thesis.

7.6.1. ChIP-seq

Technical replicates were first merged and raw read quality assessed using *fastqc*.(202) Sequencing adaptors were removed using *fastp* (203) and trimmed reads were reassessed using *fastqc* to ensure successful adaptor trimming.(202) The hg38 human genome was downloaded and indexed, before trimmed reads were aligned using *Bowtie2*.(205) Aligned reads were filtered using *samtools* to remove any unmapped reads and those with secondary alignments (281) and deduplicated using the *Picard MarkDuplicates* tool (282) Blacklisted regions (207) were downloaded and removed using *BEDTools intersect*.(283) Peaks were called with *MACS2* peak calling software, using the default settings.(208) Peaks present in 2 out of 3 of the biological replicates

within each cell line were identified using *BEDTools* intersect -u function, and those distinct when comparing across different cell lines using the -v function.(283) Intersected peak files were annotated with their genomic location using *Homer annotatePeaks* and associated with the gene that has the closest transcriptional start site to the G4 peak.(284) Peaks and alignment tracks were visualised on IGV.(194)

Once peaks had been called and annotated, additional downstream processing was conducted for quality control and further investigative avenues. Sample clustering was performed using *DiffBind*.(285) Motif enrichment was conducted online using MEME-suite using the MEME function.(209) SuperG4s were identified using Rank Ordering of Super-enhancers (ROSE).(286)(287) Pathway enrichments were analysed using both ShinyGO (211) and GSEA.(288)

7.6.2.ATAC-seq

The merging of technical replicates, read quality assessment, adaptor trimming, alignment and aligned reads filtering and sorting was completed following the same bioinformatic pipeline as for the ChIP-seq data. An additional read filtering step was conducted to remove any mitochondrial reads using *samtools*.(281) Peaks were called this time using *MACS2* with *-nolambda* settings.(208) Peaks were intersected to identify overlapping and distinct peaks using the same process as in ChIP-seq; *bedtools intersect*.(283) ShinyGO was used for pathway enrichment analysis of ATAC-seq peaks.(211)

7.6.3.RNA-seq

Adaptor sequences were trimmed using *fastp*, similarly to initial ChIP-seq and ATAC-seq read processing.(203) The hg38 genome was built using *STAR*, with the corresponding overhang appropriate for the read length; read length minus 1.(218) Trimmed RNA reads were aligned to the built genome using *STAR* (218) and the aligned files were indexed using *samtools*.(281) Differential gene expression was assessed using DESeq2.(220)

Sample clustering, PCA plots, and pathway enrichment analysis was conducted using iDEP.(219)

7.6.4. Data integration

To investigate the gene expression change associated with gain in G4 ChIP or ATAC peaks, RNA-seq differential expression data was intersected with annotated peaks from either ChIP or ATAC data sets in Excel. Briefly, the list of gene names associated with peaks was copied into an Excel sheet, alongside gene names and Log₂FC values calculated during DESeq2.(220) The function =VLOOKUP(column containing peak gene names, columns containing RNA gene names and Log₂FC values, 2, FALSE) was used to output a list of Log₂FC values associated with the ChIP or ATAC peaks. These values were then plotted in GraphPad Prism to assess mean Log₂FC differential expression value associated with gain or losses in G4 ChIP or ATAC peaks.

7.7. Statistical tests

For RNA-seq data, statistically significant differentially expressed genes between cell lines were deemed to be when the Log₂(fold-change) of gene expression is >1, and the false discovery rate (FDR) was <0.05. Log₂(fold-change) >1 means that there is a 2-fold increase in gene expression in one cell line over the other, and an FDR of 0.05 means that only 5% of these genes are expected to be a false positive.

For comparison between differentially expressed genes associated with a gain in G4 or gain in accessible sites, a Mann-Whitney U-test was used, which is a non-parametric statistical test used to compare two distinct groups to see if there is a significant difference between them. This test was used to compare gain in G4 vs gain in accessible site, loss of the regions, gene downregulation after 2 and 6 hours of G4 stabilisation, and the transcriptional upregulation associated with a G4 gain in PEO4 or PEA2.

Bibliography

1. Shnaider P, Malyants I, Ivanova O, Gordeeva V, Svirina E, Zakharzhevskaya N, et al. Establishment of Novel High-Grade Serous Ovarian Carcinoma Cell Line OVAR79. *Int J Mol Sci*. 2024 Dec 10;25(24):13236.
2. Nougaret S, Lakhman Y, Gönen M, Goldman DA, Miccò M, D'Anastasi M, et al. High-Grade Serous Ovarian Cancer: Associations between *BRCA* Mutation Status, CT Imaging Phenotypes, and Clinical Outcomes. *Radiology*. 2017 Nov;285(2):472–81.
3. Ciucci A, Zannoni GF, Buttarelli M, Martinelli E, Mascilini F, Petrillo M, et al. Ovarian low and high grade serous carcinomas: hidden divergent features in the tumor microenvironment. *Oncotarget*. 2016 Oct 12;7(42):68033–43.
4. Chandra A, Pius C, Nabeel M, Nair M, Vishwanatha JK, Ahmad S, et al. Ovarian cancer: Current status and strategies for improving therapeutic outcomes. *Cancer Med*. 2019 Nov 27;8(16):7018–31.
5. Kurman RJ, Shih IM. Pathogenesis of Ovarian Cancer. *Int J of Gynecol Pathol*. 2008 Apr;27(2):151-60.
6. Wang H, Guo M, Wei H, Chen Y. Targeting p53 pathways: mechanisms, structures and advances in therapy. *Signal Transduct Target Ther*. 2023 Mar 1;8(1):92.
7. Testa U, Petrucci E, Pasquini L, Castelli G, Pelosi E. Ovarian Cancers: Genetic Abnormalities, Tumor Heterogeneity and Progression, Clonal Evolution and Cancer Stem Cells. *Medicines*. 2018 Feb 1;5(1):16.
8. Prakash R, Zhang Y, Feng W, Jasin M. Homologous Recombination and Human Health: The Roles of BRCA1, BRCA2, and Associated Proteins. *Cold Spring Harb Perspect Biol*. 2015 Apr 1;7(4):a016600.
9. Lontos M, Zografos E, Zoumpourlis P, Andrikopoulou A, Svarna A, Fiste O, et al. BRCA1/2 Mutation Types Do Not Affect Prognosis in Ovarian Cancer Patients. *Current Oncology*. 2021 Nov 3;28(6):4446–56.

10. Chen X, Peng F, Ji Y, Xiang H, Wang X, Liu T, et al. Brca2 deficiency drives gastrointestinal tumor formation and is selectively inhibited by mitomycin C. *Cell Death Dis.* 2020 Sep 26;11(9):812.
11. Tsonis O, Gkrozou F, Vlachos K, Paschopoulos M, Mitsis MC, Zakynthinakis-Kyriakou N, et al. Upfront debulking surgery for high-grade serous ovarian carcinoma: current evidence. *Ann Transl Med.* 2020 Dec;8(24):1707–1707.
12. Coleridge SL, Bryant A, Kehoe S, Morrison J. Neoadjuvant chemotherapy before surgery versus surgery followed by chemotherapy for initial treatment in advanced ovarian epithelial cancer. *Cochrane Database of Syst Rev.* 2021 Jul 30;2021(7).
13. Dasari S, Bernard Tchounwou P. Cisplatin in cancer therapy: Molecular mechanisms of action. *Eur J Pharmacol.* 2014 Oct;740:364–78.
14. Zhang C, Xu C, Gao X, Yao Q. Platinum-based drugs for cancer therapy and anti-tumor strategies. *Theranostics.* 2022;12(5):2115–32.
15. Alemasova EE, Lavrik OI. Poly(ADP-ribosyl)ation by PARP1: Reaction mechanism and regulatory proteins. *Nucleic Acids Res.* 2019 May 7;47(8):3811-3827.
16. Ray Chaudhuri A, Nussenzweig A. The multifaceted roles of PARP1 in DNA repair and chromatin remodelling. *Nat Rev Mol Cell Biol.* 017 Oct;18(10):610-621.
17. Murai J, Huang SYN, Das BB, Renaud A, Zhang Y, Doroshow JH, et al. Trapping of PARP1 and PARP2 by clinical PARP inhibitors. *Cancer Res.* 2012 Nov 1;72(21):5588-99.
18. Fu X, Li P, Zhou Q, He R, Wang G, Zhu S, et al. Mechanism of PARP inhibitor resistance and potential overcoming strategies. *Genes Dis.* 2024 Jan;11(1):306–20.
19. Lord CJ, Ashworth A. PARP inhibitors: Synthetic lethality in the clinic. *Science.* 2017 Mar 17;355(6330):1152-1158.
20. Helleday T. The underlying mechanism for the PARP and BRCA synthetic lethality: Clearing up the misunderstandings. *Mol Oncol.* 2011 Aug 22;5(4):387–93.

21. Norouzi-Barough L, Sarookhani MR, Sharifi M, Moghbelinejad S, Jangjoo S, Salehi R. Molecular mechanisms of drug resistance in ovarian cancer. *J Cell Physiol.* 2018 Jun 4;233(6):4546–62.
22. Wang L, Wang X, Zhu X, Zhong L, Jiang Q, Wang Y, et al. Drug resistance in ovarian cancer: from mechanism to clinical trial. *Mol Cancer.* 2024 Mar 28;23(1):66.
23. Silva R, Glennon K, Metoudi M, Moran B, Salta S, Slattery K, et al. Unveiling the epigenomic mechanisms of acquired platinum-resistance in high-grade serous ovarian cancer. *Int J Cancer.* 2023 Jul 20;153(1):120–32.
24. Alatise KL, Gardner S, Alexander-Bryant A. Mechanisms of Drug Resistance in Ovarian Cancer and Associated Gene Targets. *Cancers.* 2022 Dec 18;14(24):6246.
25. Wang X, Zhang H, Chen X. Drug resistance and combating drug resistance in cancer. *Cancer Drug Resist.* 2019 Jun 19;2(2):141–160.
26. Alhmoud JF, Woolley JF, Al Moustafa AE, Malki MI. DNA Damage/Repair Management in Cancers. *Cancers.* 2020 Apr 23;12(4):1050.
27. Burdett NL, Willis MO, Alsop K, Hunt AL, Pandey A, Hamilton PT, et al. Multiomic analysis of homologous recombination-deficient end-stage high-grade serous ovarian cancer. *Nat Genet.* 2023 Mar 27;55(3):437–50.
28. Sakai W, Swisher EM, Karlan BY, Agarwal MK, Higgins J, Friedman C, et al. Secondary mutations as a mechanism of cisplatin resistance in BRCA2-mutated cancers. *Nature.* 2008 Feb 10;451(7182):1116–20.
29. Swisher EM, Sakai W, Karlan BY, Wurz K, Urban N, Taniguchi T. Secondary *BRCA1* Mutations in *BRCA1* -Mutated Ovarian Carcinomas with Platinum Resistance. *Cancer Res.* 2008 Apr 15;68(8):2581–6.
30. Wen Y, Hou Y, Yi X, Sun S, Guo J, He X, et al. EZH2 activates CHK1 signaling to promote ovarian cancer chemoresistance by maintaining the properties of cancer stem cells. *Theranostics.* 2021;11(4):1795–813.

31. Huang Y, Hong W, Wei X. The molecular mechanisms and therapeutic strategies of EMT in tumor progression and metastasis. *J Hematol Oncol*. 2022 Sep 8;15(1):129.
32. Chen X, Zhang J, Zhang Z, Li H, Cheng W, Liu J. Cancer stem cells, epithelial-mesenchymal transition, and drug resistance in high-grade ovarian serous carcinoma. *Hum Pathol*. 2013 Nov;44(11):2373–84.
33. Brown R, Curry E, Magnani L, Wilhelm-Benartzi CS, Borley J. Poised epigenetic states and acquired drug resistance in cancer. *Nat Rev Cancer*. 2014 Nov 25;14(11):747–53.
34. Baribeau S, Chaudhry P, Parent S, Asselin É. Resveratrol Inhibits Cisplatin-Induced Epithelial-to-Mesenchymal Transition in Ovarian Cancer Cell Lines. *PLoS One*. 2014 Jan 22;9(1):e86987.
35. Haslehurst AM, Koti M, Dharsee M, Nuin P, Evans K, Geraci J, et al. EMT transcription factors snail and slug directly contribute to cisplatin resistance in ovarian cancer. *BMC Cancer*. 2012 Dec 19;12(1):91.
36. Marchini S, Fruscio R, Clivio L, Beltrame L, Porcu L, Nerini IF, et al. Resistance to platinum-based chemotherapy is associated with epithelial to mesenchymal transition in epithelial ovarian cancer. *Eur J Cancer*. 2013 Jan;49(2):520–30.
37. Wu Y, Zhou BP. TNF- α /NF- κ B/Snail pathway in cancer cell migration and invasion. *Br J Cancer*. 2010 Feb 19;102(4):639–44.
38. Teeuwssen M, Fodde R. Wnt Signaling in Ovarian Cancer Stemness, EMT, and Therapy Resistance. *J Clin Med*. 2019 Oct 11;8(10):1658.
39. Luo W, Shi Q, Han M, Zhang Z, Reiter RJ, Ashrafizadeh M, et al. TGF- β -driven EMT in cancer progression and drug resistance. *Cytokine Growth Factor Rev*. 2025 May;
40. Su H, Lai H, Lin Y, Liu C, Chen C, Chou Y, et al. Epigenetic silencing of SFRP5 is related to malignant phenotype and chemoresistance of ovarian cancer through Wnt signaling pathway. *Int J Cancer*. 2010 Aug 24;127(3):555–67.

41. Kim BG, Malek E, Choi SH, Ignatz-Hoover JJ, Driscoll JJ. Novel therapies emerging in oncology to target the TGF- β pathway. *J Hematol Oncol*. 2021 Apr 6;14(1):55.
42. Wang J, Wu GS. Role of Autophagy in Cisplatin Resistance in Ovarian Cancer Cells. *J Biol Chem*. 2014 Jun;289(24):17163–73.
43. Chang H, Zou Z. Targeting autophagy to overcome drug resistance: further developments. *J Hematol Oncol*. 2020 Dec 25;13(1):159.
44. Patel NH, Xu J, Saleh T, Wu Y, Lima S, Gewirtz DA. Influence of nonprotective autophagy and the autophagic switch on sensitivity to cisplatin in non-small cell lung cancer cells. *Biochem Pharmacol*. 2020 May;175:113896.
45. Hewitt G, Korolchuk VI. Repair, Reuse, Recycle: The Expanding Role of Autophagy in Genome Maintenance. *Trends Cell Biol*. 2017 May;27(5):340–51.
46. Pan Z, Zhang H, Dokudovskaya S. The Role of mTORC1 Pathway and Autophagy in Resistance to Platinum-Based Chemotherapeutics. *Int J Mol Sci*. 2023 Jun 26;24(13):10651.
47. Rinne N, Christie EL, Ardasheva A, Kwok CH, Demchenko N, Low C, et al. Targeting the PI3K/AKT/mTOR pathway in epithelial ovarian cancer, therapeutic treatment options for platinum-resistant ovarian cancer. *Cancer Drug Resist*. 2021 Apr 14;4(3):573-595.
48. Zhang X, Wang X, Song X, Liu C, Shi Y, Wang Y, et al. Programmed cell death 4 enhances chemosensitivity of ovarian cancer cells by activating death receptor pathway *in vitro* and *in vivo*. *Cancer Sci*. 2010 Oct;101(10):2163–70.
49. Wangpaichitr M, Wu C, You M, Kuo MT, Feun L, Lampidis T, et al. Inhibition of mTOR restores cisplatin sensitivity through down-regulation of growth and anti-apoptotic proteins. *Eur J Pharmacol*. 2008 Sep;591(1–3):124–7.
50. Taylor SJ, Hollis RL, Gourley C, Herrington CS, Langdon SP, Arends MJ. FANCD2 expression affects platinum response and further characteristics of high grade serous ovarian cancer in cells with different genetic backgrounds. *Exp Mol Pathol*. 2024 Aug;138:104916.

51. Hanahan D, Weinberg RA. Hallmarks of Cancer: The Next Generation. *Cell*. 2011 Mar;144(5):646–74.
52. Dupont C, Armant D, Brenner C. Epigenetics: Definition, Mechanisms and Clinical Perspective. *Semin Reprod Med*. 2009 Sep 26;27(05):351–7.
53. Fardi M, Solali S, Farshdousti Hagh M. Epigenetic mechanisms as a new approach in cancer treatment: An updated review. *Genes Dis*. 2018 Dec;5(4):304–11.
54. Moore LD, Le T, Fan G. DNA Methylation and Its Basic Function. *Neuropsychopharmacology*. 2013 Jan 11;38(1):23–38.
55. Jang H, Shin W, Lee J, Do J. CpG and Non-CpG Methylation in Epigenetic Gene Regulation and Brain Function. *Genes*. 2017 May 23;8(6):148.
56. Baldwin RL, Nemeth E, Tran H, Shvartsman H, Cass I, Narod S, et al. BRCA1 promoter region hypermethylation in ovarian carcinoma: a population-based study. *Cancer Res*. 2000 Oct 1;60(19):5329–33.
57. Teodoridis JM, Hall J, Marsh S, Kannall HD, Smyth C, Curto J, et al. CpG Island Methylation of DNA Damage Response Genes in Advanced Ovarian Cancer. *Cancer Res*. 2005 Oct 1;65(19):8961–7.
58. Strathdee G, MacKean MJ, Illand M, Brown R. A role for methylation of the hMLH1 promoter in loss of hMLH1 expression and drug resistance in ovarian cancer. *Oncogene*. 1999 Apr 8;18(14):2335–41.
59. Tian H, Yan L, Xiao-fei L, Hai-yan S, Juan C, Shan K. Hypermethylation of mismatch repair gene hMSH2 associates with platinum-resistant disease in epithelial ovarian cancer. *Clin Epigenetics*. 2019 Dec 30;11(1):153.
60. Bannister AJ, Kouzarides T. Regulation of chromatin by histone modifications. *Cell Res*. 2011 Mar 15;21(3):381–95.
61. Eberhart A, Becker PB. Histone acetylation: a switch between repressive and permissive chromatin. *EMBO Rep*. 2002 Mar;3(3):224–9.

62. Buenrostro JD, Wu B, Chang HY, Greenleaf WJ. ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide. *Curr Protoc Mol Biol*. 2015 Jan 5;109(1).
63. Grandi FC, Modi H, Kampman L, Corces MR. Chromatin accessibility profiling by ATAC-seq. *Nat Protoc*. 2022 Jun 27;17(6):1518–52.
64. Kim MG, Pak JH, Choi WH, Park JY, Nam JH, Kim JH. The relationship between cisplatin resistance and histone deacetylase isoform overexpression in epithelial ovarian cancer cell lines. *J Gynecol Oncol*. 2012;23(3):182.
65. Yano M, Miyazawa M, Ogane N, Ogasawara A, Hasegawa K, Narahara H, et al. Up-regulation of HDAC6 Results in Poor Prognosis and Chemoresistance in Patients With Advanced Ovarian High-grade Serous Carcinoma. *Anticancer Res*. 2021 Mar 3;41(3):1647–54.
66. Gallinari P, Marco S Di, Jones P, Pallaoro M, Steinkühler C. HDACs, histone deacetylation and gene transcription: from molecular biology to cancer therapeutics. *Cell Res*. 2007 Mar 27;17(3):195–211.
67. Watson ZL, Yamamoto TM, McMellen A, Kim H, Hughes CJ, Wheeler LJ, et al. Histone methyltransferases EHMT1 and EHMT2 (GLP/G9A) maintain PARP inhibitor resistance in high-grade serous ovarian carcinoma. *Clin Epigenetics*. 2019 Nov 27;11(1):165.
68. Snowden AW, Gregory PD, Case CC, Pabo CO. Gene-Specific Targeting of H3K9 Methylation Is Sufficient for Initiating Repression In Vivo. *Curr Biol*. 2002 Dec;12(24):2159–66.
69. Ara AI, Xia M, Ramani K, Mato JM, Lu SC. S-adenosylmethionine inhibits lipopolysaccharide-induced gene expression via modulation of histone methylation. *Hepatology*. 2008 May;47(5):1655–66.
70. Borley J, Brown R. Epigenetic mechanisms and therapeutic targets of chemotherapy resistance in epithelial ovarian cancer. *Ann Med*. 2015 Jul 4;47(5):359–69.

71. Chapman-Rothe N, Curry E, Zeller C, Liber D, Stronach E, Gabra H, et al. Chromatin H3K27me3/H3K4me3 histone marks define gene sets in high-grade serous ovarian cancer that distinguish malignant, tumour-sustaining and chemo-resistant ovarian tumour cells. *Oncogene*. 2013 Sep 19;32(38):4586–92.
72. Panigrahi A, O'Malley BW. Mechanisms of enhancer action: the known and the unknown. *Genome Biol*. 2021 Dec 15;22(1):108.
73. Borsari B, Villegas-Mirón P, Pérez-Lluch S, Turpin I, Laayouni H, Segarra-Casas A, et al. Enhancers with tissue-specific activity are enriched in intronic regions. *Genome Res*. 2021 Aug;31(8):1325–36.
74. Thurman RE, Rynes E, Humbert R, Vierstra J, Maurano MT, Haugen E, et al. The accessible chromatin landscape of the human genome. *Nature*. 2012 Sep 5;489(7414):75–82.
75. Whyte WA, Orlando DA, Hnisz D, Abraham BJ, Lin CY, Kagey MH, et al. Master Transcription Factors and Mediator Establish Super-Enhancers at Key Cell Identity Genes. *Cell*. 2013 Apr;153(2):307–19.
76. Pott S, Lieb JD. What are super-enhancers? *Nat Genet*. 2015 Jan 30;47(1):8–12.
77. Creyghton MP, Cheng AW, Welstead GG, Kooistra T, Carey BW, Steine EJ, et al. Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proc Natl Acad Sci U S A*. 2010 Dec 14;107(50):21931–6.
78. Shang S, Yang J, Jazaeri AA, Duval AJ, Tufan T, Lopes Fischer N, et al. Chemotherapy-Induced Distal Enhancers Drive Transcriptional Programs to Maintain the Chemoresistant State in Ovarian Cancer. *Cancer Res*. 2019 Sep 15;79(18):4599–611.
79. Liu S, Zou Q, Chen JP, Yao X, Guan P, Liang W, et al. Targeting enhancer reprogramming to mitigate MEK inhibitor resistance in preclinical models of advanced ovarian cancer. *J Clin Inv*. 2021 Oct 15;131(20).

80. Gallon J, Loomis E, Curry E, Martin N, Brody L, Garner I, et al. Chromatin accessibility changes at intergenic regions are associated with ovarian cancer drug resistance. *Clin Epigenetics*. 2021 Dec 5;13(1):122.
81. Croft W, Pounds R, Jeevan D, Singh K, Balega J, Sundar S, et al. The chromatin landscape of high-grade serous ovarian cancer metastasis identifies regulatory drivers in post-chemotherapy residual tumour cells. *Commun Biol*. 2024 Sep 28;7(1):1211.
82. Wang X, Yan J, Shen B, Wei G. Integrated Chromatin Accessibility and Transcriptome Landscapes of Doxorubicin-Resistant Breast Cancer Cells. *Front Cell Dev Biol*. 2021 Jul 30;9.
83. Mukherjee AK, Sharma S, Chowdhury S. Non-duplex G-Quadruplex Structures Emerge as Mediators of Epigenetic Modifications. *Trends Genetics*. 2019 Feb;35(2):129–44.
84. Spiegel J, Adhikari S, Balasubramanian S. The Structure and Function of DNA G-Quadruplexes. *Trends Chem*. 2020 Feb;2(2):123–36.
85. Gellert M, Lipsett MN, Davies DR. HELIX FORMATION BY GUANYLIC ACID. *Proc Natl Academy Sci U S A*. 1962 Dec;48(12):2013–8.
86. Bhattacharyya D, Mirihana Arachchilage G, Basu S. Metal Cations in G-Quadruplex Folding and Stability. *Front Chem*. 2016 Sep 9;4.
87. Spiegel J, Cuesta SM, Adhikari S, Hänsel-Hertsch R, Tannahill D, Balasubramanian S. G-quadruplexes are transcription factor binding hubs in human chromatin. *Genome Biol*. 2021 Dec 23;22(1):117.
88. Robinson J, Flint G, Garner I, Galli S, Maher TE, Kuimova MK, et al. G-quadruplex structures regulate long-range transcriptional reprogramming to promote drug resistance in ovarian cancer cells. *Genome Biol*. 2025 Jul 12;26(1):183.
89. Parkinson GN, Lee MPH, Neidle S. Crystal structure of parallel quadruplexes from human telomeric DNA. *Nature*. 2002 Jun 26;417(6891):876–80.

90. Puig Lombardi E, Londoño-Vallejo A. A guide to computational methods for G-quadruplex prediction. *Nucleic Acids Res.* 2020 Jan 10;48(1):1–15.
91. Huppert JL, Balasubramanian S. Prevalence of quadruplexes in the human genome. *Nucleic Acids Res.* 2005;33(9):2908–16.
92. Mukundan VT, Phan AT. Bulges in G-Quadruplexes: Broadening the Definition of G-Quadruplex-Forming Sequences. *J Am Chem Soc.* 2013 Apr 3;135(13):5017–28.
93. Bedrat A, Lacroix L, Mergny JL. Re-evaluation of G-quadruplex propensity with G4Hunter. *Nucleic Acids Res.* 2016 Feb 29;44(4):1746–59.
94. Chambers VS, Marsico G, Boutell JM, Di Antonio M, Smith GP, Balasubramanian S. High-throughput sequencing of DNA G-quadruplex structures in the human genome. *Nat Biotechnol.* 2015 Aug;33(8):877-81.
95. Marsico G, Chambers VS, Sahakyan AB, McCauley P, Boutell JM, Antonio M Di, et al. Whole genome experimental maps of DNA G-quadruplexes in multiple species. *Nucleic Acids Res.* 2019 May 7;47(8):3862–74.
96. Biffi G, Tannahill D, McCafferty J, Balasubramanian S. Quantitative visualization of DNA G-quadruplex structures in human cells. *Nat Chem.* 2013 Mar;5(3):182-6.
97. Hänsel-Hertsch R, Beraldi D, Lensing S V, Marsico G, Zyner K, Parry A, et al. G-quadruplex structures mark human regulatory chromatin. *Nat Genet.* 2016 Oct 12;48(10):1267–72.
98. Park PJ. ChIP-seq: advantages and challenges of a maturing technology. *Nat Rev Genet.* 2009 Oct 8;10(10):669–80.
99. Hänsel-Hertsch R, Spiegel J, Marsico G, Tannahill D, Balasubramanian S. Genome-wide mapping of endogenous G-quadruplex DNA structures by chromatin immunoprecipitation and high-throughput sequencing. *Nat Protoc.* 2018 Mar;13(3):551-564.

100. Mao SQ, Ghanbarian AT, Spiegel J, Martínez Cuesta S, Beraldi D, Di Antonio M, et al. DNA G-quadruplex structures mold the DNA methylome. *Nat Struct Mol Biol.* 2018 Oct 1;25(10):951–7.
101. Huppert JL, Balasubramanian S. G-quadruplexes in promoters throughout the human genome. *Nucleic Acids Res.* 2007;35(2):406–13.
102. Balasubramanian S, Hurley LH, Neidle S. Targeting G-quadruplexes in gene promoters: a novel anticancer strategy? *Nat Rev Drug Discov.* 2011 Apr;10(4):261–75.
103. Xu H, Handoko L, Wei X, Ye C, Sheng J, Wei CL, et al. A signal–noise model for significance analysis of ChIP-seq with negative control. *Bioinformatics.* 2010 May 1;26(9):1199–204.
104. Lyu J, Shao R, Kwong Yung PY, Elsässer SJ. Genome-wide mapping of G-quadruplex structures with CUT&Tag. *Nucleic Acids Res.* 2022;50(3).
105. Hui WWI, Simeone A, Zyner KG, Tannahill D, Balasubramanian S. Single-cell mapping of DNA G-quadruplex structures in human cancer cells. *Sci Rep.* 2021 Dec 8;11(1):23641.
106. Galli S, Flint G, Růžicková L, Di Antonio M. Genome-wide mapping of G-quadruplex DNA: a step-by-step guide to select the most effective method. *RSC Chem Biol.* 2024;5(5):426–38.
107. Thompson MD, Byrd AK. Untargeted CUT&Tag reads are enriched at accessible chromatin and restrict identification of potential G4-forming sequences in G4-targeted CUT&Tag experiments. *Nucleic Acids Res.* 2025 Jul 19;53(14).
108. Yu Z, Spiegel J, Melidis L, Hui WWI, Zhang X, Radzevičius A, et al. Chem-map profiles drug binding to chromatin in cells. *Nat Biotechnol.* 2023 Sep 23;41(9):1265–71.
109. Esnault C, Magat T, Zine El Aabidine A, Garcia-Oliver E, Cucchiaroni A, Bouchouika S, et al. G4access identifies G-quadruplexes and their associations with open chromatin and imprinting control regions. *Nat Genet.* 2023 Aug 3;55(8):1359–69.

110. Colando S, Schulz D, Hardin J. Selecting ChIP-seq normalization methods from the perspective of their technical conditions. *Brief Bioinform.* 2025 Jul 2;26(4).
111. Bryan TM. G-Quadruplexes at Telomeres: Friend or Foe? *Molecules.* 2020 Aug 13;25(16):3686.
112. Rhodes D, Lipps HJ. G-quadruplexes and their regulatory roles in biology. *Nucleic Acids Res.* 2015 Oct 15;43(18):8627–37.
113. Robinson J, Raguseo F, Nuccio SP, Liano D, Di Antonio M. DNA G-quadruplex structures: more than simple roadblocks to transcription? *Nucleic Acids Res.* 2021 Sep 7;49(15):8419–31.
114. Raguseo F, Wang Y, Li J, Petrić Howe M, Balendra R, Huyghebaert A, et al. The ALS/FTD-related C9orf72 hexanucleotide repeat expansion forms RNA condensates through multimolecular G-quadruplexes. *Nat Commun.* 2023 Dec 13;14(1):8272.
115. Wellinger RJ, Sen D. The DNA structures at the ends of eukaryotic chromosomes. *Eur J Cancer.* 1997 Apr;33(5):735–49.
116. Tran PLT, Mergny JL, Alberti P. Stability of telomeric G-quadruplexes. *Nucleic Acids Res.* 2011 Apr;39(8):3282–94.
117. De Cian A, Cristofari G, Reichenbach P, De Lemos E, Monchaud D, Teulade-Fichou MP, et al. Reevaluation of telomerase inhibition by quadruplex ligands and their mechanisms of action. *Proc Natl Academy Sci U S A.* 2007 Oct 30;104(44):17347–52.
118. Moye AL, Porter KC, Cohen SB, Phan T, Zyner KG, Sasaki N, et al. Telomeric G-quadruplexes are a substrate and site of localization for human telomerase. *Nat Commun.* 2015 Jul 9;6(1):7643.
119. Rizzo A, Salvati E, Porru M, D'Angelo C, Stevens MF, D'Incalci M, et al. Stabilization of quadruplex DNA perturbs telomere replication leading to the activation of an ATR-dependent ATM signaling pathway. *Nucleic Acids Res.* 2009 Sep;37(16):5353–64.

120. Gauthier LR, Granotier C, Hoffschir F, Etienne O, Ayoub A, Desmaze C, et al. Rad51 and DNA-PKcs are involved in the generation of specific telomere aberrations induced by the quadruplex ligand 360A that impair mitotic cell progression and lead to cell death. *Cellular and Mol Life Sci.* 2012 Feb 20;69(4):629–40.
121. Pfeiffer V, Lingner J. Replication of Telomeres and the Regulation of Telomerase. *Cold Spring Harb Perspect Biol.* 2013 May 1;5(5):a010405–a010405.
122. Desmaze C, Soria JC, Freulet-Marrière MA, Mathieu N, Sabatier L. Telomere-driven genomic instability in cancer cells. *Cancer Lett.* 2003 May;194(2):173–82.
123. Lerner LK, Sale JE. Replication of G Quadruplex DNA. *Genes.* 2019 Jan 29;10(2):95.
124. Valton AL, Prioleau MN. G-Quadruplexes in DNA Replication: A Problem or a Necessity? *Trends in Genetics.* 2016 Nov;32(11):697–706.
125. Sun H, Karow JK, Hickson ID, Maizels N. The Bloom's Syndrome Helicase Unwinds G4 DNA. *J Biol Chem.* 1998 Oct;273(42):27587–92.
126. Fry M, Loeb LA. Human Werner Syndrome DNA Helicase Unwinds Tetrahelical Structures of the Fragile X Syndrome Repeat Sequence d(CGG). *J Biol Chem.* 1999 Apr;274(18):12797–802.
127. Wu Y, Shin-ya K, Brosh RM. FANCDJ Helicase Defective in Fanconia Anemia and Breast Cancer Unwinds G-Quadruplex DNA To Defend Genomic Stability. *Mol Cell Biol.* 2008 Jun 1;28(12):4116–28.
128. Sarkies P, Reams C, Simpson LJ, Sale JE. Epigenetic Instability due to Defective Replication of Structured DNA. *Mol Cell.* 2010 Dec;40(5):703–13.
129. Prorok P, Artufel M, Aze A, Coulombe P, Peiffer I, Lacroix L, et al. Involvement of G-quadruplex regions in mammalian replication origin activity. *Nat Commun.* 2019 Jul 22;10(1):3274.
130. Wang B, Zhang L, Dai T, Qin Z, Lu H, Zhang L, et al. Liquid–liquid phase separation in human health and diseases. *Signal Transduct Target Ther.* 2021 Aug 2;6(1):290.

131. Lafontaine DLJ, Riback JA, Bascetin R, Brangwynne CP. The nucleolus as a multiphase liquid condensate. *Nat Rev Mol Cell Biol.* 2021 Mar 1;22(3):165–82.
132. Boija A, Klein IA, Sabari BR, Dall’Agnese A, Coffey EL, Zamudio A V., et al. Transcription Factors Activate Genes through the Phase-Separation Capacity of Their Activation Domains. *Cell.* 2018 Dec;175(7):1842-1855.e16.
133. Wang W, Li D, Xu Q, Cheng J, Yu Z, Li G, et al. G-quadruplexes promote the motility in MAZ phase-separated condensates to activate CCND1 expression and contribute to hepatocarcinogenesis. *Nat Commun.* 2024 Feb 5;15(1):1045.
134. Siddiqui-Jain A, Grand CL, Bearss DJ, Hurley LH. Direct evidence for a G-quadruplex in a promoter region and its targeting with a small molecule to repress c-MYC transcription. *Proc Natl Acad Sci U S A.* 2002 Sep 3;99(18):11593–8.
135. Lago S, Nadai M, Cernilogar FM, Kazerani M, Domíniguez Moreno H, Schotta G, et al. Promoter G-quadruplexes and transcription factors cooperate to shape the cell type-specific transcriptome. *Nat Commun.* 2021 Jun 23;12(1):3885.
136. Esain-Garcia I, Kirchner A, Melidis L, Tavares R de CA, Dhir S, Simeone A, et al. G-quadruplex DNA structure is a positive regulator of MYC transcription. *Proc Natl Acad Sci U S A.* 2024 Feb 13;121(7).
137. Andersson R, Sandelin A. Determinants of enhancer and promoter activities of regulatory elements. *Nat Rev Genet.* 2020 Feb 11;21(2):71–87.
138. Kyrchanova O V., Bylino O V., Georgiev PG. Mechanisms of enhancer-promoter communication and chromosomal architecture in mammals and Drosophila. *Front Genet.* 2022 Dec 1;13.
139. Karpinska MA, Oudelaar AM. The role of loop extrusion in enhancer-mediated gene activation. *Curr Opin Genet Dev.* 2023 Apr;79:102022.
140. Ren G, Jin W, Cui K, Rodrigez J, Hu G, Zhang Z, et al. CTCF-Mediated Enhancer-Promoter Interaction Is a Critical Regulator of Cell-to-Cell Variation of Gene Expression. *Mol Cell.* 2017 Sep;67(6):1049-1058.e6.

141. Hou Y, Li F, Zhang R, Li S, Liu H, Qin ZS, et al. Integrative characterization of G-Quadruplexes in the three-dimensional chromatin structure. *Epigenetics*. 2019 Sep 2;14(9):894–911.
142. Wulfridge P, Yan Q, Rell N, Doherty J, Jacobson S, Offley S, et al. G-quadruplexes associated with R-loops promote CTCF binding. *Mol Cell*. 2023 Sep;83(17):3064-3079.e5.
143. Santos-Pereira JM, Aguilera A. R loops: new modulators of genome dynamics and function. *Nat Rev Genet*. 2015 Oct 15;16(10):583–97.
144. Roy D, Lieber MR. G Clustering Is Important for the Initiation of Transcription-Induced R-Loops In Vitro, whereas High G Density without Clustering Is Sufficient Thereafter. *Mol Cell Biol*. 2009 Jun 1;29(11):3124–33.
145. Meng Y, Zou L. Building an integrated view of R-loops, transcription, and chromatin. *DNA Repair*. 2025 May;149:103832.
146. Lee CY, McNerney C, Ma K, Zhao W, Wang A, Myong S. R-loop induced G-quadruplex in non-template promotes transcription by successive R-loop formation. *Nat Commun*. 2020 Jul 7;11(1):3392.
147. Tan J, Wang X, Phoon L, Yang H, Lan L. Resolution of ROS-induced G-quadruplexes and R-loops at transcriptionally active sites is dependent on BLM helicase. *FEBS Lett*. 2020 May 9;594(9):1359–67.
148. De Magis A, Manzo SG, Russo M, Marinello J, Morigi R, Sordet O, et al. DNA damage and genome instability by G-quadruplex ligands are mediated by R loops in human cancer cells. *Proc Natl Academy Sci U S A*. 2019 Jan 15;116(3):816–25.
149. Xu H, Di Antonio M, McKinney S, Mathew V, Ho B, O’Neil NJ, et al. CX-5461 is a DNA G-quadruplex stabilizer with selective lethality in BRCA1/2 deficient tumours. *Nat Commun*. 2017 Feb 17;8:14432.
150. Li L, Williams P, Ren W, Wang MY, Gao Z, Miao W, et al. YY1 interacts with guanine quadruplexes to regulate DNA looping and gene expression. *Nat Chem Biol*. 2021 Feb 16;17(2):161–8.

151. Roy SS, Bagri S, Sengupta A, Then CR, Kumar R, Sridharan S, et al. Artificially inserted G-quadruplex DNA secondary structures induce long-distance chromatin activation. *eLife*. 2024 Jun 13:RP96216
152. Liano D, Chowdhury S, Di Antonio M. Cockayne Syndrome B Protein Selectively Resolves and Interact with Intermolecular DNA G-Quadruplex Structures. *J Am Chem Soc*. 2021 Dec 15;143(49):20988–1002.
153. Zhang X, Spiegel J, Martínez Cuesta S, Adhikari S, Balasubramanian S. Chemical profiling of DNA G-quadruplex-interacting proteins in live cells. *Nat Chem*. 2021 Jul 28;13(7):626–33.
154. Tsai RX, Fang KC, Yang PC, Hsieh YH, Chiang IT, Chen Y, et al. TERRA regulates DNA G-quadruplex formation and ATRX recruitment to chromatin. *Nucleic Acids Res*. 2022 Nov 28;50(21):12217–34.
155. Esnault C, Zine El Aabidine A, Robert MC, Cucchiaroni A, Magat T, Pigeot A, et al. G-quadruplexes are promoter elements controlling nucleosome exclusion and RNA polymerase II pausing. *Nat Genet*. 2025 Aug 22;57(8):1981–93.
156. Ahmed AA, Neidle S. A G-Quadruplex-Binding Small Molecule and the HDAC Inhibitor SAHA (Vorinostat) Act Synergistically in Gemcitabine-Sensitive and Resistant Pancreatic Cancer Cells. *Molecules*. 2020 Nov 19;25(22):5407.
157. Niu K, Xiang L, Zhang X, Li X, Yao T, Li J, et al. DNA 5mC methylation inhibits the formation of G-quadruplex structures in the genome. *Genome Biol*. 2025 Jul 11;26(1):202.
158. Matsumoto S, Tateishi-Karimata H, Sugimoto N. DNA methylation is regulated by both the stability and topology of G-quadruplex. *Chem Comm*. 2022;58(89):12459–62.
159. Kosiol N, Juranek S, Brossart P, Heine A, Paeschke K. G-quadruplexes: a promising target for cancer therapy. *Mol Cancer*. 2021 Feb 25;20(1):40.

160. Biffi G, Tannahill D, Miller J, Howat WJ, Balasubramanian S. Elevated Levels of G-Quadruplex Formation in Human Stomach and Liver Cancer Tissues. *PLoS One*. 2014 Jul 17;9(7):e102711.
161. Tseng TY, Chen WW, Chu IT, Wang CL, Chang CC, Lin MC, et al. The G-quadruplex fluorescent probe 3,6-bis(1-methyl-2-vinyl-pyridinium) carbazole diiodide as a biosensor for human cancers. *Sci Rep*. 2018 Oct 31;8(1):16082.
162. Eddy J, Maizels N. Gene function correlates with potential for G4 DNA formation in the human genome. *Nucleic Acids Res*. 2006 Sep 1;34(14):3887–96.
163. Cogoi S, Xodo LE. G-quadruplex formation within the promoter of the KRAS proto-oncogene and its effect on transcription. *Nucleic Acids Res*. 2006 May 10;34(9):2536–49.
164. Kim NW, Piatyszek MA, Prowse KR, Harley CB, West MD, Ho PLC, et al. Specific Association of Human Telomerase Activity with Immortal Cells and Cancer. *Science*. 1994 Dec 23;266(5193):2011–5.
165. Tauchi T, Shin-ya K, Sashida G, Sumi M, Nakajima A, Shimamoto T, et al. Activity of a novel G-quadruplex-interactive telomerase inhibitor, telomestatin (SOT-095), against human leukemia cells: involvement of ATM-dependent DNA damage response pathways. *Oncogene*. 2003 Aug 14;22(34):5338–47.
166. Phatak P, Cookson JC, Dai F, Smith V, Gartenhaus RB, Stevens MFG, et al. Telomere uncapping by the G-quadruplex ligand RHPS4 inhibits clonogenic tumour cell growth in vitro and in vivo consistent with a cancer stem cell targeting mechanism. *Br J Cancer*. 2007 Apr 3;96(8):1223–33.
167. Rodriguez R, Miller KM, Forment J V, Bradshaw CR, Nikan M, Britton S, et al. Small-molecule-induced DNA damage identifies alternative DNA structures in human genes. *Nat Chem Biol*. 2012 Mar 5;8(3):301–10.
168. Zimmer J, Tacconi EMC, Folio C, Badie S, Porru M, Klare K, et al. Targeting BRCA1 and BRCA2 Deficiencies with G-Quadruplex-Interacting Compounds. *Mol Cell*. 2016 Feb 4;61(3):449–60.

169. Groelly FJ, Porru M, Zimmer J, Benainous H, Visser Y De, Kosova AA, et al. Anti-tumoural activity of the G-quadruplex ligand pyridostatin against BRCA1/2-deficient tumours. *EMBO Mol Med*. 2022 Mar 7;14(3):e14501.
170. Li HX, He YM, Fei J, Guo M, Zeng C, Yan PJ, et al. The G-quadruplex ligand CX-5461: an innovative candidate for disease treatment. *J Transl Med*. 2025 Apr 18;23(1):457.
171. Sanij E, Hannan KM, Xuan J, Yan S, Ahern JE, Trigos AS, et al. CX-5461 activates the DNA damage response and demonstrates therapeutic efficacy in high-grade serous ovarian cancer. *Nat Comm*. 2020 May 26;11(1):2641.
172. Li MZ, Meng T, Song SS, Bao X Bin, Ma LP, Zhang N, et al. Discovery of MTR-106 as a highly potent G-quadruplex stabilizer for treating BRCA-deficient cancers. *Invest New Drugs*. 2021 Oct;39(5):1213-1221.
173. Nalabothula N, Al-jumaily T, Eteleeb AM, Flight RM, Xiaorong S, Moseley H, et al. Genome-Wide Profiling of PARP1 Reveals an Interplay with Gene Regulatory Regions and DNA Methylation. *PLoS One*. 2015 Aug 25;10(8):e0135410.
174. Krishnakumar R, Gamble MJ, Frizzell KM, Berrocal JG, Kininis M, Kraus WL. Reciprocal Binding of PARP-1 and Histone H1 at Promoters Specifies Transcriptional Outcomes. *Science*. 2008 Feb 8;319(5864):819–21.
175. Edwards AD, Marecki JC, Byrd AK, Gao J, Raney KD. G-Quadruplex loops regulate PARP-1 enzymatic activation. *Nucleic Acids Res*. 2021 Jan 11;49(1):416-431.
176. Salvati E, Botta L, Amato J, Di Leva FS, Zizza P, Gioiello A, et al. Lead Discovery of Dual G-Quadruplex Stabilizers and Poly(ADP-ribose) Polymerases (PARPs) Inhibitors: A New Avenue in Anticancer Treatment. *J Med Chem*. 2017 May 11;60(9):3626-3635.
177. Sengar A, Vandana JJ, Chambers VS, Di Antonio M, Winnerdy FR, Balasubramanian S, et al. Structure of a (3+1) hybrid G-quadruplex in the PARP1 promoter. *Nucleic Acids Res*. 2019 Feb 20;47(3):1564-1572.

178. Li H, Liu ZY, Wu N, Chen YC, Cheng Q, Wang J. PARP inhibitor resistance: the underlying mechanisms and clinical implications. *Mol Cancer*. 2020 Dec 20;19(1):107.
179. Dallavalle S, Princiotta S, Mattio LM, Artali R, Musso L, Aviñó A, et al. Investigation of the complexes formed between PARP1 inhibitors and PARP1 G-quadruplex at the gene promoter region. *Int J Mol Sci*. 2021 Aug 14;22(16):8737.
180. Richl T, Kuper J, Kisker C. G-quadruplex-mediated genomic instability drives SNVs in cancer. *Nucleic Acids Res*. 2024 Mar 21;52(5):2198–211.
181. Tiek DM, Erdogan B, Razaghi R, Jin L, Sadowski N, Alamillo-Ferrer C, et al. Temozolomide-induced guanine mutations create exploitable vulnerabilities of guanine-rich DNA and RNA regions in drug-resistant gliomas. *Sci Adv*. 2022 Jun 24;8(25).
182. Rodriguez R, Müller S, Yeoman JA, Trentesaux C, Riou JF, Balasubramanian S. A novel small molecule that alters shelterin integrity and triggers a DNA-damage response at telomeres. *J Am Chem Soc*. 2008 Nov 26;130(47):15758–9.
183. Marchetti C, Zyner KG, Ohnmacht SA, Robson M, Haider SM, Morton JP, et al. Targeting Multiple Effector Pathways in Pancreatic Ductal Adenocarcinoma with a G-Quadruplex-Binding Small Molecule. *J Med Chem*. 2018 Mar 22;61(6):2500–17.
184. Hänsel-Hertsch R, Simeone A, Shea A, Hui WWI, Zyner KG, Marsico G, et al. Landscape of G-quadruplex DNA structural regions in breast cancer. *Nat Genet*. 2020 Sep 3;52(9):878–83.
185. Lund RJ, Huhtinen K, Salmi J, Rantala J, Nguyen E V., Moulder R, et al. DNA methylation and Transcriptome Changes Associated with Cisplatin Resistance in Ovarian Cancer. *Sci Rep*. 2017 May 4;7(1):1469.
186. Ma Q, Yang F, Mackintosh C, Jayani RS, Oh S, Jin C, et al. Super-Enhancer Redistribution as a Mechanism of Broad Gene Dysregulation in Repeatedly Drug-Treated Cancer Cells. *Cell Rep*. 2020 Apr;31(3):107532.

187. Wolf CR, Hayward IP, Lawrie SS, Buckton K, McIntyre MA, Adams DJ, et al. Cellular heterogeneity and drug resistance in two ovarian adenocarcinoma cell lines derived from a single patient. *Int J Cancer*. 1987 Jun 15;39(6):695–702.
188. Wang QE, Milum K, Han C, Huang YW, Wani G, Thomale J, et al. Differential contributory roles of nucleotide excision and homologous recombination repair for enhancing cisplatin sensitivity in human ovarian cancer cells. *Mol Cancer*. 2011;10(1):24.
189. Stukova M, Hall MD, Tsotsoros SD, Madigan JP, Farrell NP, Gottesman MM. Reduced accumulation of platinum drugs is not observed in drug-resistant ovarian cancer cell lines derived from cisplatin-treated patients. *J Inorg Biochem*. 2015 Aug;149:45–8.
190. Subeha MR, Goyeneche AA, Bustamante P, Lisio MA, Burnier J V., Telleria CM. Nelfinavir Induces Cytotoxicity towards High-Grade Serous Ovarian Cancer Cells, Involving Induction of the Unfolded Protein Response, Modulation of Protein Synthesis, DNA Damage, Lysosomal Impairment, and Potentiation of Toxicity Caused by Proteasome Inhibition. *Cancers*. 2021 Dec 26;14(1):99.
191. Corr BR, Finlay-Schultz J, Rosen RB, Qamar L, Post MD, Behbakht K, et al. Cytokeratin 5–Positive Cells Represent a Therapy Resistant subpopulation in Epithelial Ovarian Cancer. *Int J Gynecol Cancer*. 2015 Nov;25(9):1565–73.
192. Sakai W, Swisher EM, Jacquemont C, Chandramohan KV, Couch FJ, Langdon SP, et al. Functional Restoration of BRCA2 Protein by Secondary *BRCA2* Mutations in *BRCA2* -Mutated Ovarian Carcinoma. *Cancer Res*. 2009 Aug 15;69(16):6381–6.
193. Yoshida K, Miki Y. Role of BRCA1 and BRCA2 as regulators of DNA repair, transcription, and cell cycle in response to DNA damage. *Cancer Sci*. 2004 Nov 19;95(11):866–71.
194. Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, et al. Integrative genomics viewer. *Nat Biotechnol*. 2011 Jan 1;29(1):24–6.

195. Cooke SL, Ng CKY, Melnyk N, Garcia MJ, Hardcastle T, Temple J, et al. Genomic analysis of genetic heterogeneity and evolution in high-grade serous ovarian carcinoma. *Oncogene*. 2010 Sep 2;29(35):4905–13.
196. Rocha CRR, Silva MM, Quinet A, Cabral-Neto JB, Menck CFM. DNA repair pathways and cisplatin resistance: an intimate relationship. *Clinics*. 2018;73:e478s.
197. Guillemette S, Serra RW, Peng M, Hayes JA, Konstantinopoulos PA, Green MR, et al. Resistance to therapy in *BRCA2* mutant cells due to loss of the nucleosome remodeling factor CHD4. *Genes Dev*. 2015 Mar 1;29(5):489–94.
198. Chen X, Li Y, Xiong K, Wagner TE. A self-initiating eukaryotic transient gene expression system based on cotransfection of bacteriophage T7 RNA polymerase and DNA vectors containing a T7 autogene. *Nucleic Acids Res*. 1994;22(11):2114–20.
199. Schneider CA, Rasband WS, Eliceiri KW. NIH Image to ImageJ: 25 years of image analysis. *Nat Methods*. 2012 Jul 28;9(7):671–5.
200. Javadekar SM, Nilavar NM, Paranjape A, Das K, Raghavan SC. Characterization of G-quadruplex antibody reveals differential specificity for G4 DNA forms. *DNA Res*. 2020 Dec 3;27(5).
201. Manley LJ, Ma D, Levine SS. Monitoring Error Rates In Illumina Sequencing. *J Biomol Tech*. 2016 Dec;27(4):125–8.
202. Andrews S. FastQC: a quality control tool for high throughput sequence data. *Babraham Bioinformatics*. 2010. Available from: <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>
203. Chen S, Zhou Y, Chen Y, Gu J. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics*. 2018 Sep 1;34(17):i884–90.
204. Aganezov S, Yan SM, Soto DC, Kirsche M, Zarate S, Avdeyev P, et al. A complete reference genome improves analysis of human genetic variation. *Science*. 2022 Apr;376(6588).

205. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods*. 2012 Apr 4;9(4):357–9.
206. Bailey T, Krajewski P, Ladunga I, Lefebvre C, Li Q, Liu T, et al. Practical Guidelines for the Comprehensive Analysis of ChIP-seq Data. *PLoS Comput Biol*. 2013 Nov 14;9(11):e1003326.
207. Amemiya HM, Kundaje A, Boyle AP. The ENCODE Blacklist: Identification of Problematic Regions of the Genome. *Sci Rep*. 2019 Jun 27;9(1):9354.
208. Feng J, Liu T, Qin B, Zhang Y, Liu XS. Identifying ChIP-seq enrichment using MACS. *Nat Protoc*. 2012 Sep 30;7(9):1728–40.
209. Bailey TL, Johnson J, Grant CE, Noble WS. The MEME Suite. *Nucleic Acids Res*. 2015 Jul 1;43(W1):W39–49.
210. Yang C, Bolotin E, Jiang T, Sladek FM, Martinez E. Prevalence of the initiator over the TATA box in human and yeast genes and identification of DNA motifs enriched in human TATA-less core promoters. *Gene*. 2007 Mar;389(1):52–65.
211. Ge SX, Jung D, Yao R. ShinyGO: a graphical gene-set enrichment tool for animals and plants. *Bioinformatics*. 2020 Apr 15;36(8):2628–9.
212. David-West G, Ernlund A, Gadi A, Schneider RJ. mTORC1/2 inhibition re-sensitizes platinum-resistant ovarian cancer by disrupting selective translation of DNA damage and survival mRNAs. *Oncotarget*. 2018 Sep 4;9(69):33064–76.
213. Akrida I, Bravou V, Papadaki H. The deadly cross-talk between Hippo pathway and epithelial–mesenchymal transition (EMT) in cancer. *Mol Biol Rep*. 2022 Oct 23;49(10):10065–76.
214. Ma YT, Li C, Shen Y, You WH, Han MX, Mu YF, et al. Mechanisms of the JNK/p38 MAPK signaling pathway in drug resistance in ovarian cancer. *Front Oncol*. 2025 Apr 24;15.

215. Colic E, Patel PU, Kent OA. Aberrant MAPK Signaling Offers Therapeutic Potential for Treatment of Ovarian Carcinoma. *Onco Targets Ther.* 2022 Nov;Volume 15:1331–46.
216. Chen Y, Simeone A, Melidis L, Cuesta SM, Tannahill D, Balasubramanian S. An Upstream G-Quadruplex DNA Structure Can Stimulate Gene Transcription. *ACS Chem Biol.* 2024 Mar 15;19(3):736–42.
217. Stadhouders R, van den Heuvel A, Kolovos P, Jorna R, Leslie K, Grosveld F, et al. Transcription regulation by distal enhancers. *Transcription.* 2012 Jul 18;3(4):181–6.
218. Dobin A, Gingeras TR. Mapping RNA-seq Reads with STAR. *Curr Protoc Bioinformatics.* 2015 Sep 3;51(1).
219. Ge SX, Son EW, Yao R. iDEP: an integrated web application for differential expression and pathway analysis of RNA-Seq data. *BMC Bioinformatics.* 2018 Dec 19;19(1):534.
220. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 2014 Dec 5;15(12):550.
221. Liu M, Zhou X, Sun L, Tan S, Liu T, Xiao W, et al. Chronic stress induces platinum and Niraparib resistance in mouse models of ovarian cancer. *Exp Cell Res.* 2022 Jan;410(2):112935.
222. Pinato DJ, Graham J, Gabra H, Sharma R. Evolving concepts in the management of drug resistant ovarian cancer: Dose dense chemotherapy and the reversal of clinical platinum resistance. *Cancer Treat Rev.* 2013 Apr;39(2):153–60.
223. Fang F, Cardenas H, Huang H, Jiang G, Perkins SM, Zhang C, et al. Genomic and Epigenomic Signatures in Ovarian Cancer Associated with Resensitization to Platinum Drugs. *Cancer Res.* 2018 Feb 1;78(3):631–44.
224. Lin Y, Zhou X, Ni Y, Zhao X, Liang X. Metabolic reprogramming of the tumor immune microenvironment in ovarian cancer: A novel orientation for immunotherapy. *Front Immunol.* 2022 Oct 14;13.

225. Martincuks A, Song J, Kohut A, Zhang C, Li YJ, Zhao Q, et al. PARP Inhibition Activates STAT3 in Both Tumor and Immune Cells Underlying Therapy Resistance and Immunosuppression In Ovarian Cancer. *Front Oncol*. 2021 Dec 7;11.
226. Pokhriyal R, Hariprasad R, Kumar L, Hariprasad G. Chemotherapy Resistance in Advanced Ovarian Cancer Patients. *Biomark Cancer*. 2019 Jan 5;11.
227. Hu X, li J, Fu M, Zhao X, Wang W. The JAK/STAT signaling pathway: from bench to clinic. *Signal Transduct Target Ther*. 2021 Nov 26;6(1):402.
228. Al-Bataineh R, Al Sharie AH, Altal OF. The emerging role of JAK inhibitors in ovarian cancer: new kids on the block? *Ann Med Surg*. 2024 Feb;86(2):631–4.
229. Buenrostro JD, Wu B, Chang HY, Greenleaf WJ. ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide. *Curr Protoc Mol Biol*. 2015 Jan 5;109(1).
230. Szerlong HJ, Hansen JC. Nucleosome distribution and linker DNA: connecting nuclear function to dynamic chromatin structure. *Biochem and Cell Biol*. 2011 Feb;89(1):24–34.
231. Cherchame E, Vromman M, Guégan J, Coutelier M. ATAC-seq methods for consensus peak generation to investigate differences in chromatin accessibility per sample group v1. 2024. Available from: <https://www.protocols.io/view/atac-seq-methods-for-consensus-peak-generation-to-daea2bae.html>
232. Tikhonova P, Pavlova I, Isaakova E, Tsvetkov V, Bogomazova A, Vedekhina T, et al. DNA G-Quadruplexes Contribute to CTCF Recruitment. *Int J Mol Sci*. 2021 Jun 30;22(13):7090.
233. Belton JM, McCord RP, Gibcus JH, Naumova N, Zhan Y, Dekker J. Hi-C: A comprehensive technique to capture the conformation of genomes. *Methods*. 2012 Nov;58(3):268–76.
234. Chen X, Xu H, Shu X, Song CX. Mapping epigenetic modifications by sequencing technologies. *Cell Death Differ*. 2025 Jan 1;32(1):56–65.

235. Kanehisa M. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.* 2000 Jan 1;28(1):27–30.
236. Ediriweera MK, Tennekoon KH, Samarakoon SR. Role of the PI3K/AKT/mTOR signaling pathway in ovarian cancer: Biological and therapeutic significance. *Semin Cancer Biol.* 2019 Dec;59:147–60.
237. Son B, Lee W, Kim H, Shin H, Park HH. Targeted therapy of cancer stem cells: inhibition of mTOR in pre-clinical and clinical research. *Cell Death Dis.* 2024 Sep 30;15(9):696.
238. Ma H, Tian T, Cui Z. Targeting ovarian cancer stem cells: a new way out. *Stem Cell Res Ther.* 2023 Feb 14;14(1):28.
239. Easwaran H, Tsai HC, Baylin SB. Cancer Epigenetics: Tumor Heterogeneity, Plasticity of Stem-like States, and Drug Resistance. *Mol Cell.* 2014 Jun;54(5):716–27.
240. Das PK, Islam F, Lam AK. The Roles of Cancer Stem Cells and Therapy Resistance in Colorectal Carcinoma. *Cells.* 2020 Jun 3;9(6):1392.
241. Pillich RT, Chen J, Churas C, Fong D, Gyori BM, Ideker T, et al. NDEx IQuery: a multi-method network gene set analysis leveraging the Network Data Exchange. *Bioinformatics.* 2023 Mar 1;39(3).
242. Nagaprashantha LD, Singhal J, Li H, Warden C, Liu X, Horne D, et al. 2'-Hydroxyflavanone effectively targets RLIP76-mediated drug transport and regulates critical signaling networks in breast cancer. *Oncotarget.* 2022 Jan 1;13(1):331–331.
243. Roberts-Thomson SJ, Chalmers SB, Monteith GR. The Calcium-Signaling Toolkit in Cancer: Remodeling and Targeting. *Cold Spring Harb Perspect Biol.* 2019 Aug;11(8):a035204.
244. Bonnefond ML, Lambert B, Giffard F, Abeilard E, Brotin E, Louis MH, et al. Calcium signals inhibition sensitizes ovarian carcinoma cells to anti-Bcl-xL strategies through Mcl-1 down-regulation. *Apoptosis.* 2015 Apr 28;20(4):535–50.

245. Lee H, Kim JW, Kim DK, Choi DK, Lee S, Yu JH, et al. Calcium Channels as Novel Therapeutic Targets for Ovarian Cancer Stem Cells. *Int J Mol Sci*. 2020 Mar 27;21(7):2327.
246. Deng F, Fu M, Zhao C, Lei J, Xu T, Ji B, et al. Calcium signals and potential therapy targets in ovarian cancer. *Int J Oncol*. 2023 Sep 12;63(5):125.
247. Karp PD, Midford PE, Caspi R, Khodursky A. Pathway size matters: the influence of pathway granularity on over-representation (enrichment analysis) statistics. *BMC Genomics*. 2021 Dec 16;22(1):191.
248. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Academy Sci U S A*. 2005 Oct 25;102(43):15545–50.
249. Harrington BS, Annunziata CM. NF- κ B Signaling in Ovarian Cancer. *Cancers*. 2019 Aug 15;11(8):1182.
250. Marolt N, Pavlič R, Kreft T, Gjorgoska M, Rižner TL. Targeting estrogen metabolism in high-grade serous ovarian cancer shows promise to overcome platinum resistance. *Biomedicine & Pharmacotherapy*. 2024 Aug;177:117069.
251. Zheng X, Andruska N, Lambrecht MJ, He S, Parissenti A, Hergenrother PJ, et al. Targeting multidrug-resistant ovarian cancer through estrogen receptor α dependent ATP depletion caused by hyperactivation of the unfolded protein response. *Oncotarget*. 2018 Mar 13;9(19):14741–53.
252. Chun J. Isoalantolactone Suppresses Glycolysis and Resensitizes Cisplatin-Based Chemotherapy in Cisplatin-Resistant Ovarian Cancer Cells. *Int J Mol Sci*. 2023 Aug 3;24(15):12397.
253. Chakraborty PK, Mustafi SB, Xiong X, Dwivedi SKD, Nesin V, Saha S, et al. MICU1 drives glycolysis and chemoresistance in ovarian cancer. *Nat Commun*. 2017 May 22;8(1):14634.

254. Niu Z, He J, Wang S, Xue B, Zhang H, Hou R, et al. Targeting Glycolysis for Treatment of Breast Cancer Resistance: Current Progress and Future Prospects. *Int J Biol Sci*. 2025 Mar 24;21(6):2589–605.
255. Jurkovicova D, Neophytou CM, Gašparović AČ, Gonçalves AC. DNA Damage Response in Cancer Therapy and Resistance: Challenges and Opportunities. *Int J Mol Sci*. 2022 Nov 24;23(23):14672.
256. Hientz K, Mohr A, Bhakta-Guha D, Efferth T. The role of p53 in cancer drug resistance and targeted chemotherapy. *Oncotarget*. 2017 Jan 31;8(5):8921–46.
257. J Devor E, E Thomas A, M Schickling B. BRCA2 Status Alters the Effect of the P53 Reactivator HO-3867 in Ovarian Cancer Cells. *J Cancer Sci Clin Ther*. 2023;07(02).
258. Alatise KL, Gardner S, Alexander-Bryant A. Mechanisms of Drug Resistance in Ovarian Cancer and Associated Gene Targets. *Cancers*. 2022 Dec 18;14(24):6246.
259. Tian T, Chen YQ, Wang SR, Zhou X. G-Quadruplex: A Regulator of Gene Expression and Its Chemical Targeting. *Chem*. 2018 Jun;4(6):1314–44.
260. De Magis A, Manzo SG, Russo M, Marinello J, Morigi R, Sordet O, et al. DNA damage and genome instability by G-quadruplex ligands are mediated by R loops in human cancer cells. *Proc Natl Academy Sci U S A*. 2019 Jan 15;116(3):816–25.
261. Tyner JW, Haderk F, Kumaraswamy A, Baughn LB, Van Ness B, Liu S, et al. Understanding Drug Sensitivity and Tackling Resistance in Cancer. *Cancer Res*. 2022 Apr 15;82(8):1448–60.
262. Brasseur K, Gévry N, Asselin E. Chemoresistance and targeted therapies in ovarian and endometrial cancers. *Oncotarget*. 2017;8(3).
263. Kilanowska A, Ziółkowska A, Stasiak P, Gibas-Dorna M. cAMP-Dependent Signaling and Ovarian Cancer. *Cells*. 2022 Nov 29;11(23):3835.
264. Mungenast F, Thalhammer T. Estrogen Biosynthesis and Action in Ovarian Cancer. *Front Endocrinol*. 2014 Nov 12;5.

265. Gao Y, Qiao X, Liu Z, Zhang W. The role of E2F2 in cancer progression and its value as a therapeutic target. *Front Immunol*. 2024 May 14;15.
266. Salanci Š, Vilková M, Martinez L, Mirossay L, Michalková R, Mojžiš J. The Induction of G2/M Phase Cell Cycle Arrest and Apoptosis by the Chalcone Derivative 1C in Sensitive and Resistant Ovarian Cancer Cells Is Associated with ROS Generation. *Int J Mol Sci*. 2024 Jul 9;25(14):7541.
267. Reyes-González JM, Armaiz-Peña GN, Mangala LS, Valiyeva F, Ivan C, Pradeep S, et al. Targeting c-MYC in Platinum-Resistant Ovarian Cancer. *Mol Cancer Ther*. 2015 Oct 1;14(10):2260–9.
268. Dhanasekaran R, Deutzmann A, Mahauad-Fernandez WD, Hansen AS, Gouw AM, Felsher DW. The MYC oncogene — the grand orchestrator of cancer growth and immune evasion. *Nat Rev Clin Oncol*. 2022 Jan 10;19(1):23–36.
269. Liu T, Li Y, Wang X, Yang X, Fu Y, Zheng Y, et al. The role of interferons in ovarian cancer progression: Hinderer or promoter? *Front Immunol*. 2022 Dec 21;13.
270. Selvendiran K, Bratasz A, Kuppusamy ML, Tazi MF, Rivera BK, Kuppusamy P. Hypoxia induces chemoresistance in ovarian cancer cells by activation of signal transducer and activator of transcription. *Int J Cancer*. 2009 Nov 21;125(9):2198–204.
271. He M, Wu H, Jiang Q, Liu Y, Han L, Yan Y, et al. Hypoxia-inducible factor-2 α directly promotes *BCRP* expression and mediates the resistance of ovarian cancer stem cells to adriamycin. *Mol Oncol*. 2019 Feb 14;13(2):403–21.
272. Langdon SP, Lawrie SS, Hay FG, Hawkes MM, McDonald A, Hayward IP, et al. Characterization and properties of nine human ovarian adenocarcinoma cell lines. *Cancer Res*. 1988 Nov 1;48(21):6166–72.
273. Cooke SL, Ng CKY, Melnyk N, Garcia MJ, Hardcastle T, Temple J, et al. Genomic analysis of genetic heterogeneity and evolution in high-grade serous ovarian carcinoma. *Oncogene*. 2010 Sep 2;29(35):4905–13.

274. Huang TT, Lampert EJ, Coots C, Lee JM. Targeting the PI3K pathway and DNA damage response as a therapeutic strategy in ovarian cancer. *Cancer Treat Rev.* 2020 Jun;86:102021.
275. Tsubakihara Y, Moustakas A. Epithelial-Mesenchymal Transition and Metastasis under the Control of Transforming Growth Factor β . *Int J Mol Sci.* 2018 Nov 20;19(11):3672.
276. Zou Z, Tao T, Li H, Zhu X. mTOR signaling pathway and mTOR inhibitors in cancer: progress and challenges. *Cell Biosci.* 2020 Dec 10;10(1):31.
277. Musa F, Alard A, David-West G, Curtin JP, Blank S V., Schneider RJ. Dual mTORC1/2 Inhibition as a Novel Strategy for the Resensitization and Treatment of Platinum-Resistant Ovarian Cancer. *Mol Cancer Ther.* 2016 Jul 1;15(7):1557–67.
278. Gremke N, Polo P, Dort A, Schneikert J, Elmshäuser S, Brehm C, et al. mTOR-mediated cancer drug resistance suppresses autophagy and generates a druggable metabolic vulnerability. *Nat Commun.* 2020 Sep 17;11(1):4684.
279. Nuccio SP, Cadoni E, Nikoloudaki R, Galli S, Ler AJ, Sanchez-Cabanillas C, et al. Chemically modified CRISPR-Cas9 enables targeting of individual G-quadruplex and i-motif structures, revealing ligand-dependent transcriptional perturbation. *bioRxiv [preprint].* 2025 Jul 22:2024.10.14.618195.
280. Salemme V, Centonze G, Avalor L, Natalini D, Piccolantonio A, Arina P, et al. The role of tumor microenvironment in drug resistance: emerging technologies to unravel breast cancer heterogeneity. *Front Oncol.* 2023 May 17;13.
281. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics.* 2009 Aug 15;25(16):2078–9.
282. Broad Institute. Picard toolkit: MarkDuplicates. Broad Institute; 2010. Available from: <https://broadinstitute.github.io/picard/>
283. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics.* 2010 Mar 15;26(6):841–2.

284. Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, et al. Simple Combinations of Lineage-Determining Transcription Factors Prime cis-Regulatory Elements Required for Macrophage and B Cell Identities. *Mol Cell*. 2010 May;38(4):576–89.
285. Stark R, Brown G. DiffBind: differential binding analysis of ChIP-Seq peak data. 2011. Available from: <https://bioconductor.org/packages/DiffBind>
286. Whyte WA, Orlando DA, Hnisz D, Abraham BJ, Lin CY, Kagey MH, et al. Master Transcription Factors and Mediator Establish Super-Enhancers at Key Cell Identity Genes. *Cell*. 2013 Apr;153(2):307–19.
287. Lovén J, Hoke HA, Lin CY, Lau A, Orlando DA, Vakoc CR, et al. Selective Inhibition of Tumor Oncogenes by Disruption of Super-Enhancers. *Cell*. 2013 Apr;153(2):320–34.
288. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Academy Sci U S A*. 2005 Oct 25;102(43):15545–50.