Investigating radioresistance: A systems biology and functional genomics approach to identify mechanisms of resistance to radiation and develop a predictive classifier

A thesis submitted for the degree Doctor of Philosophy

by

Philip Charlton

BSc MA MBBS MRCP FRCR

St Cross College

Hilary 2025

Publications:

Charlton PV, O'Reilly D, Philippou Y, Rao SR, Lamb ADG, Mills IG, Higgins GS, Hamdy FC, Verrill C, Buffa FM, Bryant RJ. Molecular analysis of archival diagnostic prostate cancer biopsies identifies genomic similarities in cases with progression post-radiotherapy, and those with de novo metastatic disease. Prostate. 2024 Jul;84(10):977-990. doi: 10.1002/pros.24715. Epub 2024 Apr 23. PMID: 38654435; PMCID: PMC11253896

Presentations:

Charlton P, O'Reilly D, Philippou Y, Rao S, Lamb A, Higgins G, Hamdy F, Verrill C, Bryant R, Buffa F, PO-1160: A pilot dual-platform transcriptomic analysis of diagnostic prostate biopsies & radical RT response, Radiotherapy and Oncology, 2020-11-01 | Conference poster, DOI: 10.1016/S0167-8140(21)01178-6

Poster Presentation ESTRO Nov 2020

For Lucy, Daisy and Beatrix

Acknowledgments:

I would like to thank my Supervisors, Francesca Buffa and Geoff Higgins for their guidance and patience, Timothy Maughan for his support and Eric O'Neill for his encouragement. I would also like to thank Clare Verrill and Richard Bryant for their invaluable contributions.

I would like to thank my father Tom and my late mother Cyril for their enduring support. I wish that they could have both been here to see the completion of my research.

I would like to acknowledge the constant support and devotion from my wife Lucy, without her effort and sacrifice this work would not have been possible.

THESIS CONTENTS:

CHAPTER 1

1.1 Introduction

CHAPTER 2

2.1 Methods

CHAPTER 3

3.1 Radioresistant Cell Line Generation and Characterisation

CHAPTER 4

4.1 Radioresistant Cell Line Gene expression and Methylation Analysis

CHAPTER 5

5.1 Prostate cancer radical radiotherapy clinical dataset

CHAPTER 6

6.1 Discussion

List of Figures

List of Tables:

List of Abbreviations:

ALDH1: Aldehyde dehydrogenase 1

ANOVA: analysis of variance

ATP: Adenosine triphosphate

C12FDG: 5-Dodecanoylaminofluorescein Di-β-D-Galactopyranoside

CCP: Cell Cycle Progression

CCK-8: Cell Counting Kit-8

CO2: Carbon Dioxide

cfDNA: cell free DNA

CMS: consensus molecular subtypes

DDR: DNA Damage Response

DNA: deoxyribonucleic acid

DSBs: Double Strand Breaks

EGFR: epidermal growth factor receptor

FDR: false discovery rate

FFPE: Formalin Fixed Paraffin Embedded

GARD: Genome adjusted radiation dose

GEO: Gene Expression Omnibus

GFP: green fluorescent protein

GO: gene ontology

Gy: Gray

H3K27me3: Histone H3 on lysine 27 Tri-Methylation

H&N: Head and Neck

HR: Homologous Recombination

IHC: immunohistochemistry

IR: Ionising Radiation

KEGG: Kyoto Encyclopedia of Genes and Genomes

LFC: Log Fold Change

MeV: Mega Electron Volt

MHC: Major Histo-Compatability Complex

mRNA: messenger Ribonucleic acid

NICE: National Institute of Clinical Excellence

NSCLC: non-small cell lung cancer

NHEJ: Non-Homologous End Joining

ORA: Over-Representation Analysis

OS: overall survival

padj: adjusted p value

PBS: Phosphate Buffer Solution

PC: principal component

PC1: principal component 1

PCA: principal component analysis

PCR: Polymerase chain reaction

PROMPT: Prostate cancer: Mechanisms of Progression and Treatment

PSA: Prostate Specific Antigen

3’ RNA Seq: 3’ RNA Sequencing

RIN: RNA Integrity Number

RSI: RadioSensitivity Index

SCC: Squamous Cell Carcinoma

SCORT: Stratification in Colorectal cancer

SF2: surviving fraction at 2Gy

STR: Short Tandem Repeat

TCGA: The Cancer Genome Atlas

TKI: tyrosine kinase inhibitors

VEGF: Vascular endothelial growth factor

X-gal: 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside

Abstract

Radiotherapy forms an integral part of modern cancer management, but despite advances in molecular medicine that have revolutionised the field of medical oncology enabling predictive and prognostic molecular markers and targeted therapies, these have not translated to the field radiation therapy. The work in this thesis uses the existing literature to inform the design of cell-line experiments to explore molecular factors associated with radioresistance, applies these findings to clinical cohorts, and investigates the predictive potential of pre-treatment diagnostic biopsies in prostate cancer clinical samples from patients treated with radiotherapy.

Gene expression and cell characterisation identifies that generated radioresistant cell lines exhibit cellular and gene expression changes consistent with senescence and altered methylation. Ecological modelling using experimentally derived growth and radiosensitivity parameters suggests a mechanism whereby a subpopulation of slower growing radioresistant cells through competition, can become the dominant population under the pressure of clinically relevant fractionated radiation schedules.

Genes that were significantly increased in expression in generated radioresistant lung cancer cell lines were applied to a clinical dataset and demonstrated to be prognostic in a radiotherapy treated cohort of patients with lung cancer, but not in patients who had not been treated with radiotherapy.

In a small cohort who underwent multi-omic sequencing, coherent gene expression and methylation patterns were demonstrated in pre-treatment biopsies of patients progressing following radical radiotherapy, similar to patients who presented with metastatic disease and distinct from patients who had stable disease. The significantly differentially expressed genes demonstrated the potential to predict for treatment failure in a large external dataset of radiotherapy treated patients.

CHAPTER 1

1.1 Introduction

1.11 Radiotherapy

Radiotherapy is the clinical use of ionising radiation to treat diseases, most commonly malignancy. Radiotherapy forms an integral part of the treatment of cancer contributing to the cure of 40% of cancer patients and as the sole curative treatment modality in 16% of patients[1,2].

Clinically Ionising radiation is most commonly delivered by Linear Accelerators as high energy photons in the Mega Electron Volt (MeV) range where the Compton effect dominates. This ionising radiation induces deoxyribonucleic acid (DNA) damage predominantly via the indirect effect, whereby free radicals generated near DNA strands cause damage, or to a lesser extent by the direct effect where the DNA damage occurs directly [3]. DNA double strand breaks which cannot be overcome by the classical repair pathways (Homologous Recombination) HR and Non-Homologous End Joining (NHEJ) result in mitotic catastrophe and tumour cell death. One Gray (Gy), the unit of absorbed dose of ionizing radiation, defined as the absorption of one Joule per kilogram, generates approximately 40 Double Strand Breaks (DSBs) per cell with a typical radical radiotherapy total treatment dose delivering upwards of 60Gy at 2Gy per fraction. The rapid advances in the delivery of radiation (image guided radiotherapy, volumetric arc therapy, respiratory motion gating etc) mean that tumours can be ever more precisely targeted and normal tissues spared[4]. Advances in the characterisation of cancers at the mutational, transcriptomic and proteomic levels have revolutionised the field of oncology providing prognostic and predictive information to guide clinical management and stratification for targeted therapies, but as yet this has not been translated to the field of radiation oncology.

1.12 Radiotherapy and key radiobiological concepts

Key Radiobiological Principles which have been described from the 1970s include Repair, Redistribution, Repopulation, Reoxygenation and Radiosensitivity[5]. Repair relates to the ability of the cell to repair damage from the radiation treatment. Reassortment outlines that different stages of the cell cycle exhibit varying degrees of sensitivity to radiation, most sensitive during G2-M phase and least sensitive in late S phase, and that relative proportions of cells within each phase of the cell cycle can be redistributed after irradiation. Repopulation outlines a clinical effect whereby tumour cell re-growth affects efficacy of treatment, of clinical significance to overcome when total treatment times are lengthened e.g. due to delays in therapy. Reoxygenation describes the effect whereby tumour hypoxia, which renders cells relatively radioresistant as oxygen enhances the biological effects of ionising radiation, is reduced in the surviving population of tumour cells during treatment. Radiosensitivity relates to the intrinsic radiosensitivity of the tumour cell, often seen in clinical practice with a large disparity between the radio sensitivities of different tumour types of Lymphoma relatively radiosensitive and Melanoma relatively resistant[5].

There is a growing understanding of the complex nature of cellular interactions *in vivo*, that there are likely to be additional mechanisms involved in the clinical response to radiation including Cellular metabolism[6], Stemness[7], Cellular Senescence[8], Immune Activation[9].

Previous studies establish the paradigm for utilising a pan-cancer gene expression-based predictor of cancer cell intrinsic radiosensitivity in cell lines[10]. This has been applied to clinical datasets however has not yet proved sufficiently robust to enter clinical application. The clinical benefit of a predictor of radioresistance is that it could allow management stratification e.g. treatment intensification e.g. increased dose or addition of radiosensitisers, or increased post treatment observation. The advantages of transcriptional assessments are that they can allow simultaneous measures of global gene expression, are quantitative and can be refined to develop rapid polymerase chain reaction based methods for translation to the diagnostic setting e.g. oncotype Dx[11].

The potential for a pan-cancer transcriptomic predictor of radiosensitivity has been established most notably in the RadioSensitivity Index (RSI) which utilised linear regression to identify baseline (pre-radiation) expression of genes correlating to radiosensitivity, as measured by surviving fraction at 2Gy (SF2)[10]. This was originally performed across 48 cancer cell lines using micro-array data for 6817 genes. The study identified a hub of 10 key genes, the expression of which when ranked and applied to a formula provide a surrogate for SF2 and has demonstrated clinical applicability to cohorts of radiotherapy treated patients. The RSI has also been adapted to provide a Genome Adjusted Radiation dose in retrospective clinical cohorts on the assumption that a baseline assessment of tumour radiosensitivity can be used to modulate the required dose[12]. In addition to gene expression a further study also identifies copy number alterations and several key mutations as underpinning the genetic basis for differing sensitivity to radiation[14]. This study performed high throughput screens on over 500 cell lines, utilising an ATP dependent luminescence technique as a surrogate for survival. Another known contributor to radioresistance is the effect of ploidy with a direct correlation demonstrated between ploidy and mean number of lethal chromosomal aberrations occurring[14]. It is important to note that these studies did not directly address all the clinical aspects of radiobiology known to affect the response to radiotherapy including DNA repair, cell cycle distribution, effect of hypoxia and reoxygenation and repopulation[6].

1.13 Predictive genomic biomarkers in oncology

Significant technical improvements have been made to improve the delivery of radiation, enabling the ever more precise delivery of high doses and sparing of normal tissues. The rise of systems biology and bioinformatics. originally defined as the information in biological systems, permeates the field of modern oncology and underpins much of the information integral to decision making in the oncology clinic. Bioinformatics enables an unprecedented degree of molecular characterisation of individual cancers, covering at once genomic, transcriptomic, and proteomic levels. To guide clinical management, such information can be

utilised to refine an increasingly individualised and accurate understanding of each patient's diagnosis, prognosis and prediction of response to therapy.

The quest for personalised medicine in the post-Human Genome Project era led to the elucidation of numerous "oncogene addiction" pathways whereby single mutations drive carcinogenesis[15]. This allowed the development of agents to specifically target such pathways, giving rise to a new arsenal of targeted therapies. Robust bioinformatic analysis facilitates the development of such drugs and provides a framework to stratify patients who may derive clinical benefit from targeted agents.

Molecular classification due to wider assessments of gene expression have enabled the identification of biologically distinct molecular subgroups within histological tumour types. Improved understanding of underlying biological mechanisms of pathogenesis provides both prognostic and predictive information to guide management in the clinic.

1.14 Targeted Therapy in Oncology

The canonical small molecule targeted agent imatinib was developed to treat Chronic Myeloid Leukaemia. Investigations of the Philadelphia Chromosome (ch9:22 translocation) found in most cases and subsequent BCR-ABL fusion protein identified a constitutively active tyrosine kinase causing constant proliferative signalling[16]. Imatinib was developed to inhibit this tyrosine kinase and was first-in-class of the now commonly used tyrosine kinase inhibitors (TKIs).

Gefitinib, the first TKI in solid tumours, targets the epidermal growth factor receptor (ErbB1), overexpressed in non-small cell lung cancer (NSCLC). Early trials in unstratified cohorts of previously chemotherapy-treated patients reported promising response rates however, further unselected trials showed limited success[17]. Subsequently specific activating epidermal growth factor receptor (EGFR) mutations were identified, e.g. exon 19 deletion and exon 21 substitution (L8R8) which comprise 85% of drug sensitive mutations[18] whilst

other mutations e.g. exon 20 insertions, confer resistance to TKI therapy, necessitating mutational characterisation prior to therapy initiation.

Resistance to EGFR TKI therapy develops due to further mutation in the EGFR binding pocket domain (T790M mutation~ 60%). Molecular modelling techniques led to the development of Osimertinib, a TKI preferentially targeting mutant EGFRs over wild type[19]. The collaborative development and bringing to market of which has been considered a paradigm for targeted drug development[19]. Due to their activity overlaps some TKIs are used outside their initial indication. e.g. imatinib in GastroIntestinal Stromal Tumours[20]. The potential for this type of targeted therapy design has not yet been translated to the field of radiation therapy.

Molecular stratification to guide targeted therapy is not restricted to small molecule agents. Cetuximab, a chimeric monoclonal antibody (mAb) targeting the EGFR receptor and used in the treatment of metastatic colorectal cancer, requires stratification of patients by molecular screening to exclude the presence of RAS and other mutations downstream in the EGFR signalling pathway which negate benefit from cetuximab therapy[21,22].

Not all mAbs require mutational selection, stratification for Traztuzumab a humanised monoclonal antibody (mAb) targeting the HER2 receptor (ErbB2) used in breast and oesophageal/gastric cancer, occurs via Immumo-Histo-Chemistry (IHC) assessment[23]. Computational insights into protein structure enabled antibody engineering of humanised[24] and later human mAbs against a range of oncological targets e.g. HER2, EGFR, PDL1.

Previous attempts to intensify radiotherapy treatment by targeting specific molecular pathways such as the EGFR pathway which are known to stimulate cell growth in response to radiation, have met with limited success in large scale Phase III randomised trials in Lung (RTOG0617), Head and Neck Cancers (RTOG1016) and Oesophageal (SCOPE1/RTOG 0436)[25]. The further elucidation of tumour specific radiosensitivity mechanisms has recently led to the development of novel radiotherapy-drug combinations targeting cancer specific

cell intrinsic repair mechanisms and hypoxia pathways e.g. Ataxia telangiectasia and Rad3 related (ATR) and PolQ[26] and modifying hypoxia with Atavaquone[28]. The ability to specifically target radioresistant cancer cell specific pathways therefore holds great potential but requires the elucidation of radioresistance mechanisms.

1.15 Prognostic Molecular Classification and Predictive Signatures in Oncology

Beyond single gene characterisation, wider consideration of multiple genes at the transcriptional level is clinically informative with potential to impact clinical management. Analysis of relative expression levels of multiple genes across multiple patient samples can identify transcriptionally similar groups or clusters. In the first instance to allow visualisation and interpretation of such large and complex datasets, heatmaps or similar high visual content plots are often generated, however behind and beyond such maps many analytical steps are carried out to assess the validity and robustness of the patient classification[28]. These additional steps can include quality control metrics of the gene signatures used to identify classifying groups e.g. co-correlation of genes, compactness, effect of data standardisation but ultimately must rely on a validation of prognostic utility in external datasets utilising the same original method to demonstrate robustness and clinical utility[29,30].

The molecular classification of breast cancer originally used complementary DNA microarrays to analyse gene expression of over eight thousand genes from fresh frozen breast cancer samples[31]. An unsupervised hierarchical clustering approach was performed based on the expression of five hundred "intrinsic" genes and identified the familiar molecular subtypes (Luminal A, Luminal B, HER2+ and Basal). These subtypes recapitulated in separate datasets, demonstrated significant prognostic implications in numerous large clinical cohorts[32]. However, a large methodological discussion started following publication of these studies on the robustness of patient classification[33,34,35].

Similar recent attempts to classify colorectal cancer resulted in six separately published systems, the consensus molecular subtypes (CMS) unify these classifications and identifies

4 subtypes: CMS1 inflammatory/immune enriched, CMS2 canonical, CMS3 metabolic and CMS4 mesenchymal[36]. Pooled expression data from all six studies generated a core consensus gene set identifying the four CMS subtypes and methods were developed to assign CMS groups to non-classified samples. Evidence from Phase III trials demonstrates independent prognostic significance of CMS and trials assessing molecular subtype directed therapy are ongoing[37]. Extensive discussion on the optimal context of application of the classifiers followed[38] and additional classifiers have been suggested e.g. colorectal cancer intrinsic subtypes (CRISs), to address the question of reliability of patient classification from a single biopsy sample[39].

Hypoxia signatures, provide an example of a phenotypic classifier applicable across multiple tumour types[40,41]. Tumour hypoxia has been demonstrated to be prognostic across multiple cancers and has been associated with genomic instability and specific driver mutations[42]. A compact hypoxia metagene has been derived and demonstrated to be highly prognostic in multiple cancer types[43].

Computational intensity and resource limitations prohibit derivation of many classifiers via the original methods for routine clinical application. Novel methods such as Nanostring, a fluorescence-based assay, offer improved turnaround and reduced computational intensity but remain research based[44]. Several examples of specifically developed transcriptomic predictive tests are in routine clinical use notably in the field of breast cancer.

Following the discovery of breast cancer molecular subtypes, several predictive gene expression-based tests were developed and commercialised. The UK National Institute of Clinical Excellence (NICE) recommends the use of EndoPredict, Oncotype DX and Prosigna tests to guide adjuvant therapy in ER +ve, HER2-ve, node negative disease for patients at intermediate risk of distant recurrence[45]. The goal of such predictive tests is that they are rapid, sensitive and specific. Oncotype Dx as an example uses reverse transcription Polymerase Chain Reaction (PCR) from Formalin Fixed Paraffin Embedded (FFPE) biopsies to measure RNA expression of 16 cancer related genes and 5 reference genes. All samples

are processed at a central laboratory and from the relative expression of these genes a recurrence score (0-100) is derived to assign a risk stratification (low/intermediate/high). Central processing ensures reproducibility however limits technical appreciation of sensitivity and specificity of the technique to published reports. The recent TAILORx trial validated prognostic use of the recurrence score to stratify patients for adjuvant hormone and chemotherapy[46]. An association with risk of local recurrence has been identified from retrospective cohorts and is being assessed in the TAILOR RT trial[39].

1.16 Current biomarkers for radiotherapy outcome

Despite the application of the Radiation Sensitivity Index, a 10 gene radiosensitivity classifier [48], to clinical cohorts generating a genomic-adjusted radiation dose[49] and validation of utility of hypoxia signatures in clinical cohorts (bladder and head and neck cancer)[50,51], there are no genomic predictive radiotherapy classifiers in routine clinical use. The phase 3 trial Adjuvant Radiotherapy Intensification Classifier (ARCTIC) trial reported a 28 gene classifier to be prognostic of loco-regional recurrence and predictive of radiotherapy benefit for early breast cancer in the adjuvant setting[52]. Retrospective and prospective cohorts of prostate cancer patients have also demonstrated the utility of genomic classifiers (PORTOS/DECIPHER) for adjuvant radiotherapy post-radical prostatectomy with ongoing studies[53-55].

Genome Adjusted Radiation Dose is an mRNA based, 10 gene, individualised RadioSensitivity Index (RSI) utilised in combination with the linear quadratic model. Scott demonstrates variation in radiation response not only between different tumour types but also heterogeneity within tumour types with clinical data from 8,271 patients across 20 tumour sites[48].

There has been some critique of the validity of this approach which relies on the assumption of an interaction between RSI and dose, Mistry identifying that the addition of an interaction

term between RSI and dose does not explain more of the outcome variance[56]. As yet there has been no prospective clinical validation of the utility of GARD. Further work by Berk modelling the effect of intratumoral heterogeneity of radiosensitivity on tumour response over the course of fractionated radiation therapy suggests that tumour cell radiosensitivity heterogeneity alters the tumour response and that tumour control will likely depend on the radiosensitivity of the most resistant clones[57]. The potential for tumour radiosensitivity to alter over the course of treatment is one potential barrier to the clinical utility of predictive gene signatures in radiotherapy.

1.17 Limitations & Future perspectives

Insights into intratumoral heterogeneity identify potential limitations of representative tumour sampling from single biopsy-based assays[58]. Circulating tumour or cell free DNA (cfDNA) can identify mutations in subpopulations and minimal residual disease[59,60] potentially overcoming sampling issues. However rigorous validation of emerging technologies is necessary prior to clinical implementation, illustrated by possible confounding sources of mutations in circulating cfDNA analysis from the haematopoietic system[61] and growing understanding of factors affecting cfDNA release kinetics[62]. Spatial transcriptomics and single cell sequencing can enable deconvolution of signal from tumour, stromal and immune compartments over bulk sequencing[63].

1.18 Literature Review of gene expression studies investigating radiosensitivity

A review of the existing literature of gene expression studies investigating radiosensitivity was conducted to identify the current knowledge base, review methods previously utilised and aid the development of a combined cell line experimental method. The search terms utilised in the literature review are detailed in Appendix (Table 1).

Review of the literature identified significant heterogeneity of experimental design and minimal overlap of identified genes when studies were grouped by tumour site. Studies were classified as either clinical studies - utilising tissue specimens or cell line *in vitro* studies.

75 studies were identified in total with 31 studies utilising clinical tissue samples and 44 studies *in vitro* using cell line samples. Significant inter-experimental variations were identified in radiation doses, method of transcriptional assessment and thresholds for analysis performed. Notably many studies include radiation doses that have little relevance to clinically used dose fractionation schedules. This was particularly apparent in the subset of studies which generates radioresistant lines. A framework was developed to categorise studies by methodological design and enable methodologically similar studies to be compared (Appendix Figure 1)

Within cell line studies significant variation in experimental design was identified however gene expression studies were broadly categorised into three distinct methodological groups.

1) Gene expression measured at baseline (pre-radiation) compared between cell lines with different radiosensitivity.
2) Gene expression measured post-radiation compared between cell lines with different radiosensitivity.
3) Gene expression measured at baseline compared between a generated isogenic radioresistant cell line and parental cell line with differing radiosensitivity, whereby radioresistant lines were established by repeat irradiation.

Table 1.1 Radiation Resistance Studies Summary

* This Table was compiled with C.R Mendez as part of his Radiation Biology MSc

Category	Reference	Author	Date	PMID	Sample Type	Tumour Type
Baseline Expression	64	Ahn et al.	2014	25202076	Cell line	Lung
Baseline Expression	65	Chen et al.	2020	33179733	Clinical	Breast
Baseline Expression	66	Li et al.	2020	31865249	Clinical	Glioblastoma
Baseline Expression	67	Bosma et al.	2020	31525407	Clinical	Breast
Baseline Expression	68	Liu et al.	2020	33014019	Clinical	Head and Neck
Baseline Expression	69	Ahmad et al.	2020	32266559	Clinical	Head and Neck
Baseline Expression	70	Knoops et al.	2007	17483295	Clinical	Haematological
Baseline Expression	71	Liu et al.	2014	24691555	Clinical	Gynaecological
Baseline Expression	72	Smirnov et al.	2012	21844125	Clinical	Haematological
Baseline Expression	73	Supiot et al.	2013	23745026	Clinical	Colo-Rectal
Baseline Expression	74	Tsujiguchi et al.	2016	26661850	Clinical	Haematological
Baseline Expression	75	Weidhaas et al.	2009	19509178	Clinical	Gynaecological
Baseline Expression	76	Zempolich et al.	2008	18299147	Clinical	Gynaecological
Baseline Expression	77	Zhou et al.	2017	28560382	Clinical	Glioblastoma
Baseline Expression	78	Chen et al.	2020	32020207	Cell Line	Lung
Baseline Expression	79	Kim et al.	2020	33012153	Cell Line	Head and Neck
Baseline Expression	80	Khan et al.	2021	33577642	Cell Line	Bladder
Baseline Expression	81	Liu et al.	2020	33038765	Cell Line	Bone
Baseline Expression	82	Ahmed et al.	2009	19451095	Cell Line	Colo-Rectal
Baseline Expression	83	Aryankalayil et al.	2014	25003313	Cell Line	Prostate
Baseline Expression	84	Bassi et al.	2008	19273547	Cell Line	Glioblastoma
Baseline Expression	85	Ding et al.	2009	15966761	Cell Line	Other
Baseline Expression	86	Kis et al.	2006	17069989	Cell Line	Other
Baseline Expression	87	Ma et al.	2013	24176853	Cell Line	Glioblastoma
Baseline Expression	88	Park et al.	2002	12466973	Cell Line	Haematological
Baseline Expression	89	Otomo et al.	2004	15133290	Cell line	Glioblastoma
Baseline Expression	90	Kim et al.	2012	22846430	Cell Line	Other
Baseline Expression	91	Speers et al.	2015	25904749	Cell Line	Breast
Baseline Expression	92	Spitzner et al.	2010	20970032	Cell Line	Colo-Rectal
Baseline Expression	93	Torres-Roca et al.	2005	16103067	Cell Line	Other
Baseline Expression	94	Yard et al.	2016	27109210	Cell Line	Other
Baseline Expression	95	Zhang et al.	2014	25032244	Cell Line	Other
Baseline Expression	96	Ishigami et al.	2007	17290400	Cell Line	Head and Neck
Baseline Expression	97	Chaudhry et al.	2003	12767531	Cell line	Other
Baseline Expression	98	Du et al.	2009	19212620	Cell line	Gynaecological
Expression Post-Irradiation	99	Zhao et al.	2019	31121035	Clinical	Glioblastoma
Expression Post-Irradiation	100	Liu et al.	2020	31698994	Clinical	Head and Neck
Expression Post-Irradiation	101	Sjöström et al.	2020	31618132	Clinical	Breast
Expression Post-Irradiation	102	Li et al.	2020	32198434	Clinical	Head and Neck
Expression Post-Irradiation	103	Zhang et al.	2019	31186755	Clinical	Head and Neck
Expression Post-Irradiation	104	Naghavi et al.	2020	31833149	Clinical	Head and Neck
Expression Post-Irradiation	105	Amundson et al.	2004	15374940	Clinical	Haematological
Expression Post-Irradiation	106	Bo et al.	2004	15226608	Clinical	Head and Neck
Expression Post-Irradiation	107	Fu et al.	2015	25940978	Clinical	Gynaecological
Expression Post-Irradiation	108	Klopp et al.	2008	18406887	Clinical	Gynaecological
Expression Post-Irradiation	109	Piening et al.	2009	19267539	Clinical	Haematological
Expression Post-Irradiation	110	Wang et al.	2006	16638715	Clinical	Haematological
Expression Post-Irradiation	111	Foy et al.	2017	28859688	Clinical	Head and Neck
Expression Post-Irradiation	112	Young et al.	2014	25369795	Clinical	Prostate
Expression Post-Irradiation	113	Thi et al.	2019	31387015	Cell line	Colo-Rectal
Expression Post-Irradiation	114	He et al.	2020	32329416	Cell line	Other
Expression Post-Irradiation	115	Akerman et al.	2005	15657912	Cell line	Haematological
Expression Post-Irradiation	116	Amundson et al.	2003	12692264	Cell line	Haematological
Expression Post-Irradiation	117	Klopp et al.	2008	18406887	Cell line	Gynaecological
Expression Post-Irradiation	118	Summerer et al.	2015	26328888	Cell line	Head and Neck
Expression Post-Irradiation	119	Eschrich et al.	2009	19735874	Cell line	Other
Expression Post-Irradiation	120	Huang et al.	2011	21344162	Cell line	Colo-Rectal
Expression Post-Irradiation	121	Kobunai et al.	2011	22199273	Cell line	Colo-Rectal
Generated Radioresistant	122	Lee et al.	2010	20510196	Cell line	Lung
Generated Radioresistant	123	Guo et al.	2005	15966762	Cell line	Lung
Generated Radioresistant	124	Achary et al.	2000	11173827	Clinical	Gynaecological
Generated Radioresistant	125	Ojima et al.	2007	17876542	Clinical	Colo-Rectal
Generated Radioresistant	126	Tewari et al.	2005	16109440	Clinical	Gynaecological
Generated Radioresistant	127	Ahn et al.	2014	25202076	Cell line	Lung
Generated Radioresistant	128	Chang et al.	2007	17671084	Cell line	Head and Neck
Generated Radioresistant	129	Fukuda et al.	2004	15365572	Cell line	Oesophageal
Generated Radioresistant	130	Khodarev et al.	2004	14755057	Cell line	Head and Neck
Generated Radioresistant	131	Li et al.	2014	24498188	Cell line	Head and Neck
Generated Radioresistant	132	Michna et al.	2016	27455841	Cell line	Head and Neck
Generated Radioresistant	133	Ogawa et al.	2008	18477249	Cell line	Pancreatic
Generated Radioresistant	134	H. Zhang et al.	2010	20868556	Cell line	Oesophageal
Generated Radioresistant	135	Zhou et al.	2010	20726715	Cell line	Head and Neck
Generated Radioresistant	136	Seifert et al.	2019	31682594	Cell line	Prostate
Generated Radioresistant	137	Todorovic et al.	2019	31775835	Cell line	Head and Neck
Generated Radioresistant	138	You et al.	2019	30642292	Cell line	Head and Neck

Table 1.2 Most Common differentially expressed genes

* This Table was compiled with C.R Mendez as part of his Radiation Biology MSc

HGNC Symbol	Ensembl ID	Description	Studies
GADD45A	ENSG00000116717	growth arrest and DNA damage inducible alpha	12
CDKN1A	ENSG00000124762	cyclin dependent kinase inhibitor 1A	9
MDM2	ENSG00000135679	MDM2 proto-oncogene	9
DDB2	ENSG00000134574	damage specific DNA binding protein 2	8
GDF15	ENSG00000130513	growth differentiation factor 15	8
PCNA	ENSG00000132646	proliferating cell nuclear antigen	8
STAT1	ENSG00000115415	signal transducer and activator of transcription 1	8
TRIM22	ENSG00000132274	tripartite motif containing 22	8
CASP1	ENSG00000137752	caspase 1	7
CCNB1	ENSG00000134057	cyclin B1	7
FAS	ENSG00000026103	Fas cell surface death receptor	7
GADD45B	ENSG00000099860	growth arrest and DNA damage inducible beta	7
HNRNPD	ENSG00000138668	heterogeneous nuclear ribonucleoprotein D	7
NQO1	ENSG00000181019	NAD(P)H quinone dehydrogenase 1	7
RAD51	ENSG00000051180	RAD51 recombinase	7
XPC	ENSG00000154767	XPC complex subunit, DNA damage recognition and repair factor	7

Examining the relative strengths of each method the first is most suited to develop a predictive classifier as it most accurately reflects the clinical situation where the only sample is a pre-treatment biopsy, however the comparison of radiosensitive vs radioresistant tumours in clinical samples is complex and likely required the utilisation of large-scale datasets for validation.

The second method identifies differences in gene expression in response to radiation potentially giving insight into mechanisms and pathways in response to radiation that may account for differential radiosensitivity e.g. upregulated DNA damage response. Given the underlying differences in both radiosensitivity and baseline (pre-radiation) gene expression

between cell lines, comparison of changes in gene expression post-radiation may be limited unless these factors are considered.

The third method potentially offers the clearest identification of genes involved in mechanisms of acquired radioresistance as it involves comparison of generated isogenic lines. Like the situation in clinical practice, this allows assessment from pre-radiation samples but the comparison between radioresistant lines derived from parental lines, potentially overcomes the intrinsic differences in gene expression between different cell lines, enabling a clearer elucidation of gene expression changes associated with acquired radioresistance. Of these studies there was still significant variation in the methods used to derive radioresistant cell-lines.

1.19 Generation of Radioresistant Cell Lines in the literature:

Ahn *et al* irradiated the lung cancer cell lines A549 and H1650 with an increasing dose of radiation (2-5Gy) every 2 weeks over 26 weeks and established their radioresistance through clonogenic assay. Microarray analysis was then used for identification of differentially expressed genes between parental and resistant lines[127].

Chang *et al* established two radioresistant subclone cell lines derived from two established Naso-Pharyngeal Cancer parental cell lines, by harvesting the fourth generation following exposure to 10Gy and comparing gene expression to baseline. Fukada established Radioresistant subclones from existing Oesophageal Squamous Cell Carcinoma cell lines through repeat irradiation with 60Gy (30 x 2Gy) fractions[128].

Lee *et al* established radioresistant cell lines through repeated irradiation with 2Gy (80Gy total) and identified differentially expressed genes to the parental line[122].

Li *et al* compared the differences of both miRNA and mRNA expression profiles in the generated radioresistant NPC CNE2-IR and radiosensitive NPC CNE2 cells which they generated through repeat irradiation with 11 Gy over 5 generations. They compared baseline gene expression between the parental and radioresistant lines[131].

Michna *et al* established radioresistant Head and Neck (H&N) Squamous Cell Carcinoma (SCC) Cell Lines through repeat 2 Gy irradiation (20Gy total) and performed differential gene expression following a further 8Gy irradiation[132].

Ogawa *et al* utilised pancreatic cancer cell lines (PK-1, PK-8, PK-9, T3M4, MiaPaCa2 and PANC-1) and treated with 10 Gy fractionated irradiation at approximately two-week intervals (total dose 150-180 Gy) to generate radioresistant sublines. Gene expression was performed between radioresistant and parental samples at baseline[133].

Ojima *et al* used Colorectal Cell lines (colo205, DLD-1, and LoVo) and radioresistant cell lines were established through repeat irradiation to a total of 100Gy. Gene expression was performed to compare resistant and baseline samples and 30 clinical specimens[125].

Zhang *et al* established radioresistant cell lines from Oesophageal Adenocarcinoma cell line Seg-1, using 2Gy irradiation up to 60Gy in total. Gene expression analysis was then performed at 0Hr, 8Hr, 24Hr post IR[134].

Zhou *et al* established radioresistance of Hep-2 H&N (Larynx) Cell Line through repeat irradiation with 6.37Gy until exponential growth to a total of 76.44 Gy. Gene expression was performed using microarrays to compare parental and resistant lines at baseline[135].

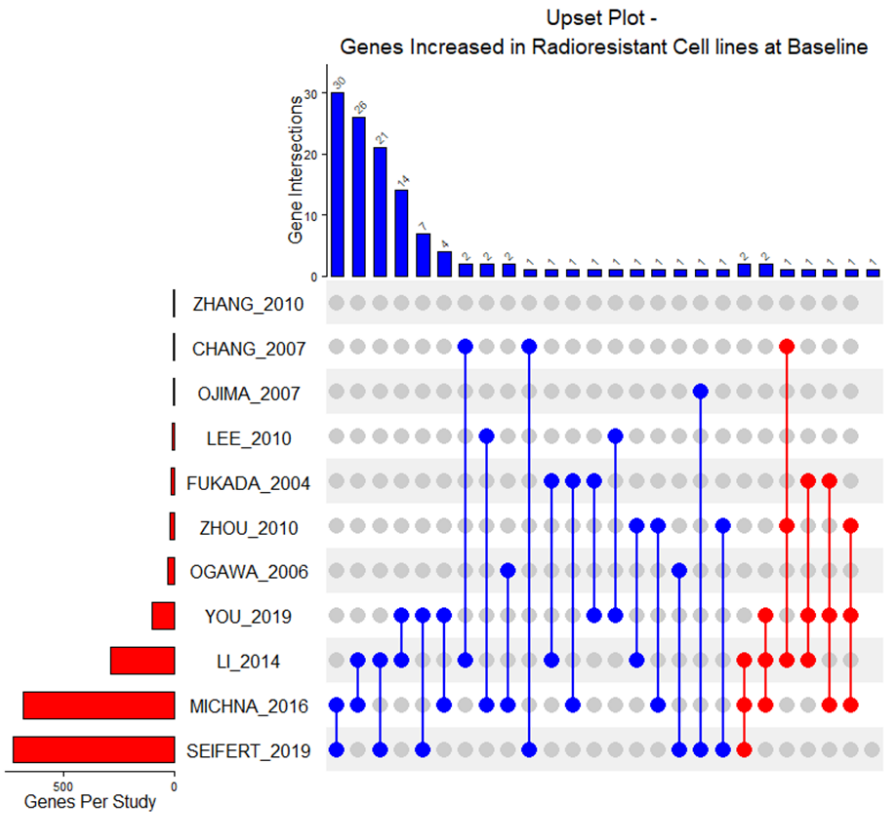
Siefert *et al* established radioresistant cell lines of the prostate cancer cell lines DU145 and LNCaP through repeat irradiation with 2Gy. Gene expression of parental and Radioresistant lines was compared at baseline using micro-arrays[136].

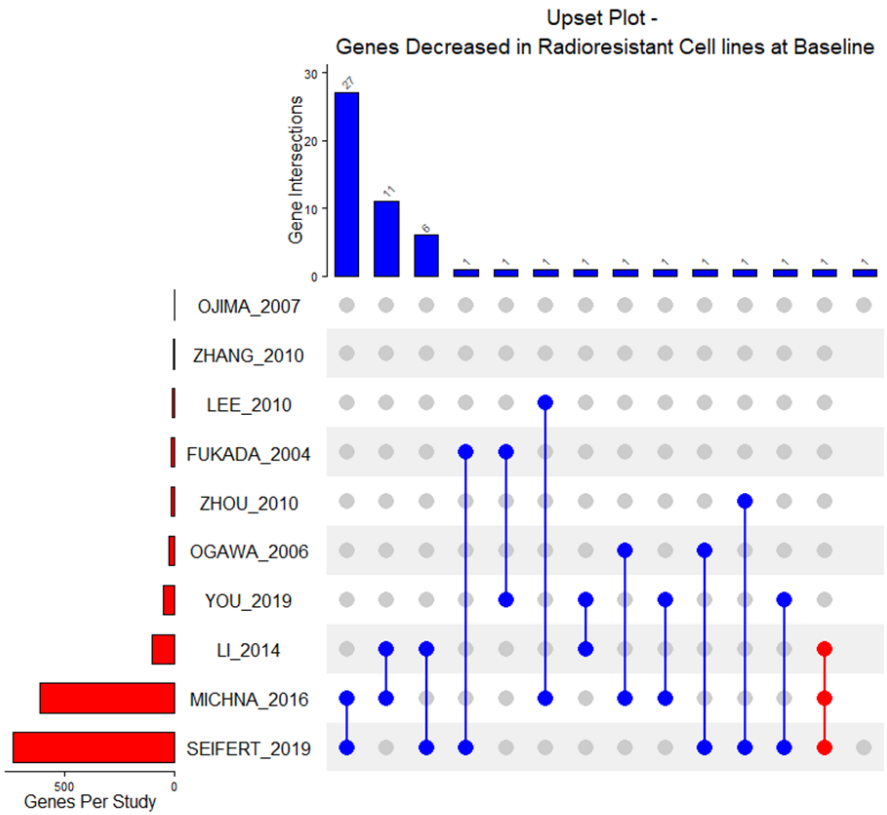
Todorovic *et al* established radioresistant cells derived from parental Head and Neck Cancer FaDu cell lines, by repeated exposure to ionizing radiation using 2Gy to a total dose of 30 Gy, 60 Gy, and 120 Gy. Mirco-array analysis was used to compare gene expression in parental and resistant lines at baseline and 5 hours post a further 5Gy irradiation[137].

You *et al* induced radioresistance in H&N cell lines OECM1, FaDu, Detroit using 2Gy irradiation over 3 months up to a total of 60Gy. Gene expression was performed using microarray to compare parental and resistant lines at baseline[138].

Three studies; Ojima[125], Achary[124] and Tewari[126] derived radioresistant cell lines from clinical tumour samples from Colo-rectal tumours (Ojima) and Gynaecological tumours (Achary and Tewari).

When considering the overlaps and concordance between generated radioresistant cell line studies Figure 1.1 demonstrates that 8 genes were identified in 3 or more studies as being increased in generated radioresistant lines compared to parental, and 117 genes from 2 studies. Only one gene was identified as being decreased in 3 studies of radioresistant cell lines with 51 genes decreased in 2 studies comparing radioresistant to parental cell lines. This degree of overlap between a smaller group of studies suggests that by considering experimental design and classifying studies by methodology a more coherent result is obtained, potentially due to minimisation of confounding factors caused by heterogeneity of study design.





Histology Key:

ZHANG, FUKADA –Oesophageal , **CHANG, LI,** Head & Neck Squamous Cell Carcinoma (NasoPharynx), **OJIMA** – Colorectal, **LEE,MICHNA**- Head & Neck Squamous Cell Carcinoma, **ZHOU, YOU** - Head & Neck Squamous Cell Carcinoma, **SIEFERT** - Prostate Adenocarcinoma

Figure 1.1: Overlaps in gene expression between radioresistant and parental cell lines

Upset Plot demonstrating overlap between gene expression studies with genes increased in radioresistant cells on the left and genes decreased on the right. The y axis barplot shows the number of genes per study. The x axis barplot shows the number of overlapping genes. The central panel demonstrates which studies contain overlapping genes overlapping 2 studies in blue and genes overlapping 3 studies in red.

Studies which used clinical tissue measure radiosensitivity through surrogate endpoints, such as tumour-shrinkage or overall survival[19,20]. Whilst these endpoints are useful in large scale studies to identify prognostic/predictive features the utility in smaller scale studies comparing one or two samples is likely to be limited, given significant confounding factors which are not listed within the studies, including treatment dose fractionation, development of distant metastatic disease vs local recurrence, which renders the utility of small scale signatures uncertain.

Numerous predictive or prognostic gene signatures are not successfully validated in external datasets, with minimal publication of negative validation results[30]. Several factors are likely to influence this, one key factor is likely to be the failure of experimental design to consider classical radiobiological factors[6]. Additional influencing factors are likely to include the tumour microenvironment and immune system, and intra-tumour heterogeneity[139].

There is a lack of clear protocols and methodology for the design of such experiments, and this is reflected in the significant heterogeneity of methods used in both biological aspects e.g. dose and fractionation or radiation used, but also in gene expression and statistical analysis thresholds such as log fold change and statistical cutoffs for significance.

Another key issue is the lack of statistical rigour often employed with such analyses, minimum sample size calculations in the development stage of studies are necessary for the development of studies with adequate power to translate findings to larger external datasets. Small sample sizes overestimate signature performance which may result in failure to validate.

Signature quality metrics are a key consideration for the development and application of signatures, including basic measures of expression of the signature genes and variability in datasets[29]. A signature developed in a specific dataset would lose significant prognostic power if the signature genes either are not expressed within an external dataset or have minimal variability. A similar and related problem relates to the lack of stability of gene

identifiers, and difficulties with the translation of results from one sequencing technique to another. Furthermore, there are specific differences between micro-array data, dependent on the array used and different RNA sequencing techniques, related to technical aspects of sequencing pipelines, including the exact tools and part of the sequence being measured, the human genome build that reads are aligned to and how reads are processed and counted[140]. These technical aspects should be accounted for when considering the result of comparisons between studies but given the complexity of nomenclature, data analysis techniques and the rate of technological change this poses a significant challenge, and a significant potential limitation of comparisons.

Significant heterogeneity exists in the literature in the generation of signatures associated with radiosensitivity and radiation response. Many factors known to be critical in the response to radiation differ between studies. e.g. cell lines vs *in vivo* samples, radiation dose delivered and fractionation, measurements of baseline gene expression or expression post-irradiation, limiting the ability to draw direct comparisons.

Key radiobiological factors such as dose/fraction size/frequency of exposure, particularly hypofractionation, can have significant effects on the clinical efficacy of radiotherapy and as such there are recommended clinical guidelines for different tumour types[133]. There is also emerging evidence of differential effects of altered dose/fractionation schedules *in vitro*[132].

As suggested by the gene changes and pathways outlined by different strategies, potentially the most clinically relevant changes may be those after exposure to radiation which may be seen in other systems e.g. Vascular endothelial growth factor (VEGF) where response after perturbation is key[141]. Comparisons between methodologically similar studies may result in greater homogeneity, and potentially a more unified approach to methodology may enable for more meaningful comparisons between different studies and signatures.

When radiation response is considered the comparison of cell lines vs clinical samples is a potentially useful one, with the implication that cell lines act as a model to explore intrinsic

cell response whereas clinical samples include effects of the tumour microenvironment immune response and hypoxia, which have previously been demonstrated to have significant effects on radiation response[9,43].

Radiation response gene signatures aim to identify cellular responses induced by Ionising Radiation (IR). For example, DNA Damage Response (DDR) pathways are known to activate in response to Ionising Radiation, but these pathways are less likely to be differentially expressed before IR is administered, and indeed not seen on pathway analysis (Appendix Figure 3-7).

Clinical utility may arise by assessing gene expression changes in a patient after the first few sessions of Radiotherapy to predict long-term response. Clinical studies investigating response to radiotherapy treatment to predict outcome are subject to multiple confounding factors which often arise because of clinical endpoints, additional therapy, and variation in patient characteristics, requiring large datasets for clinical validation is necessary for any signature to have utility. Radioresistant phenotypes could potentially result from multiple different gene expression patterns—posing a significant challenge for biomarker development. However, the advantage of RNA sequencing technologies is that it can allow multiple aspects to be assessed simultaneously e.g. hypoxia, immune component etc. External validation in clinical datasets and ideally prospective clinical validation are necessary before the utility of a gene signature can be determined to be suitable for clinical application. Within this thesis I aim to use a systems biology and functional genomics approach to identify mechanisms of resistance to radiation and develop a predictive classifier.

1.20 Aims of Thesis:

1) To generate radioresistant cell lines using clinically relevant dose and fractionation regimens and identify baseline characteristics of radioresistant cell lines compared to parental cell lines.

It is clear from the literature that the generation of radioresistant cell lines provides a suitable method to identify genomic changes between parental cell lines and cell lines where radioresistance has been established through exposure to radiation. Previous studies have utilised non-clinically relevant irradiation schedules and have failed to consider basic radiobiological factors. The results in support of this aim are outlined in Chapter 3 and provide the foundation for gene expression and methylation studies in Chapter 4. The baseline characteristics of generated radioresistance lines are further utilised with ecological modelling techniques to explore the role of tumour heterogeneity in the response to radiotherapy.

2) To identify molecular features associated with a radioresistant phenotype through gene expression and methylation analysis

The differential cellular mechanisms resulting in a radioresistant phenotype between the radioresistant and radiosensitive parental cell lines should be detectable at a transcriptional level as they are isogenic. The hypothesis at test is that the molecular features indicative of a radioresistant phenotype can be identified, the results of the work towards this aim are described in Chapter 4.

3) To identify the predictive value of features associated with radioresistant cell line phenotype in clinical cohorts.

The identified molecular features indicative of a radioresistant phenotype may provide clinical utility in predicting outcomes in clinical radiotherapy treated cohorts. The hypothesis at test is that the identified radioresistant features of a radioresistant phenotype are predictive of response to radiotherapy in clinical cohorts. The results of the work towards this aim are described in the latter part of Chapter 4.

4) To identify features of a radioresistant phenotype from a clinical cohort of patients with prostate cancer treated with radiotherapy and identify their predictive value in an external cohort of prostate cancer patients treated with radiotherapy.

By comparing sequencing data from patients who have progressed following radical radiotherapy treatment to those who have not, molecular features can be identified associated with progression following radiotherapy. The hypothesis at test is that the identified features of progression post-radiotherapy are predictive of progression post-radiotherapy in an external clinical cohort. The results of the work towards this aim are described in Chapter 5.

CHAPTER 2

2.1 Methods

2.11 In vitro experiments cell culture

Cell lines were obtained from FB lab (HT29) or GH lab (H460/A549), originally from American Type Culture Collection (ATCC) and European Collection of Authenticated Cell Cultures (ECACC). Authentication was performed using Short Tandem Repeat testing compared to ATCC standard by Northgene. Mycobacterium testing was performed on acquiring cell lines and repeated post long-term irradiation using the MycoAlert™ mycoplasma detection kit (Lonza).

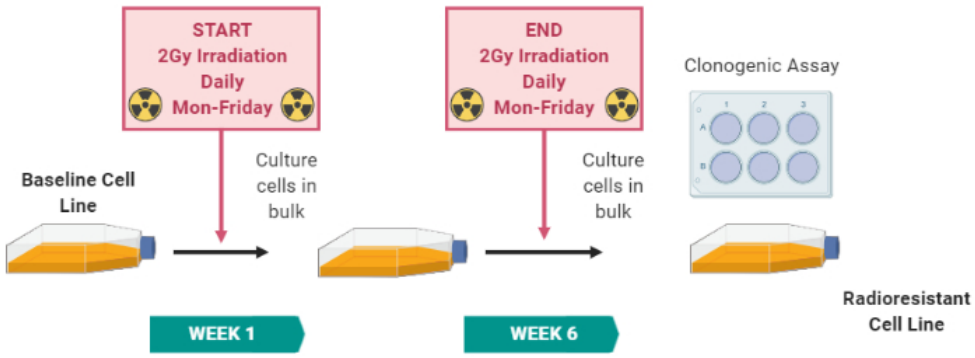
Cell culture was performed using standard aseptic techniques and appropriate media H460(RPMI-1640), A549 (F-12K Medium) and HT29(McCoy's 5a Medium Modified), with the addition of Fetal Bovine Serum 10% (Thermo Fisher, MA) and 100U/mL Penicillin, 100 mg/mL of Streptomycin (1%) and l-Glutamine (1%) (Thermo Fisher, MA). Cells were trypsinized (Trypsin 0.05% Thermo Fisher, MA) and split prior to achieving confluency, briefly adherent cells once approaching 80% confluence cells were washed with Phosphate Buffer Solution (PBS 1X, pH 7.4 Sigma Aldrich) incubated with trypsin at 37°C for 5 minutes. Cells were gently pipetted, transferred to a Falcon tube before centrifugation at 1000 rpm for 5 minutes. Supernatant was removed, cells resuspended in fresh culture medium, diluted and transferred to fresh tissue culture flasks. Media was changed approximately every 2 days Monday/Wednesday/Friday.

2.12 *In vitro* experiments generation of radioresistant cell lines

Cells from H460, A549, HT29 were cultured in 75 cm^2 culture flasks (Thermo Fisher T75) and irradiated daily using a caesium-137 irradiator (Gamma Service GSR D1). Medical Physics calculations were performed weekly to ensure that required dose was achieved, using the standard worksheet. Dose rate varied from 2Gy/95s -98s over the course of experiment.

Baseline cell lines were irradiated daily (five times per week) with 2 Gy, Monday to Friday for 6 weeks to mimic clinical radiotherapy regimens. Cells were trypsinized and split once approaching confluency. Media changes were performed 2-3x per week. During the following 6 weeks of irradiation cells were grown in culture until sufficient numbers recovered and clonogenic assay performed.

Figure 2.1 Schema of *in-vitro* generation of radioresistant cell lines



2.13 *In vitro* experiments clonogenic assays

Standard Clonogenic assays were performed to establish the radiosensitivity of baseline and irradiated cell lines [17]. Using the standard clonogenic method in 6 well plates, six replicates were performed at each dose (0, 2, 4 and 6 Gy), plating 200, 400, 800 and 2000 cells respectively with a minimum of 4 hours for attachment prior to irradiation. Cells were irradiated using a caesium irradiator. Cells were incubated in an incubator (ThermoFisher) at

38°C, 5% CO_2 for 10-14 days. Colony Counts were performed using automated colony-counter (Agilent Technologies) and software, only including colonies with over 50 cells. Clonogenic experiments were repeated 3 times for each cell line.

Statistical analysis of CFA results was performed in R 3.52 using the CFAssay package 3.10. Surviving Fraction was calculated using the formula colonies counted/cells seeded x [plating efficiency/100]. For each cell line survival curves were generated using Maximum Likelihood method and performing comparison tests between curves or between mean values of survival fractions. Statistical comparison of parental and radioresistant survival curves was performed using Analysis of Variance (ANOVA).

2.14 *In vitro* experiments cell line characterisation: senescence staining

Senescence staining was performed using a Senescence Detection (Abcam) kit per protocol. Briefly this involved irradiated cells being trypsinized and transferred into 6 well plates 24 hours prior to detection. Subconfluent cells were washed in Phosphate Buffer Solution (PBS) fixed then stained using 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal) and staining solution. Cells were incubated overnight at 37°C in an airtight pocket to prevent CO_2 ingress. Images were captured under magnification on an electronic microscope (Nikon XLE).

2.15 *In vitro* experiments cell line characterisation: senescence flow cytometry

Cells were plated and 4 hours later either irradiated or sham irradiated. They were incubated for 24 hours (37°C, 5% Carbon Dioxide (CO_2), 1hr Lysosomal acidification with (Bafilomycin 0.1uM Invitrogen) followed by 2 hours incubation with 5-Dodecanoylaminofluorescein Di-β-D-Galactopyranoside (C12FDG obtained from Invitrogen). A positive control was generated using Doxorubicin 0.1uM incubation for 48 hours (37°C, 5% Carbon Dioxide (CO_2)).

2.16 *In vitro* experiments cell line characterisation: cell cycle analysis

Cells were grown in relevant media, plated onto T25 flasks Thermo Fisher) and left for attachment for 4 hours and then either irradiated or sham irradiated with 2Gy in a caesium-

137 irradiator. Samples were incubated at 37°C at 5% CO_2 for 24 hours before being stained with Propidium Iodide (PI) (Thermo Fisher). Using a Flow Cytometer (CytoFLEX Beckman Coulter) cells were gated (Dotplot FSC/SSC),Single Cells gated (ECD-H, ECD-A) and Cell Cycle Histogram (ECD-A) analysis performed in FlowJo.

2.16 *In vitro* experiments cell line characterisation: ALDH1 flow cytometry

ALDEFLUOR Kit (Catalog no. 01700) was purchased from STEMCELL Technologies, Inc. (Vancouver, Canada). Cells were trypsinized cells at a density of 1×10^6 cells/mL and washed in phosphate-buffer, incubated for 45 min at 37°C in the dark after the addition of ALDEFLUOR reagent (5 µL/10^6 cells). As a negative control, 1.5 mM diethylaminobenzaldehyde (DEAB) was added to a final concentration of 7.5 µM. Cells were centrifuged, resuspended in ALDEFLUOR assay buffer, and analyzed by direct immunofluorescence flow cytometry. The percentage of ALDH (+) was determined by subtracting from the positive percentage of the substrate administered group.

2.17 Ecological Modeling

Ecological modeling equations were utilised:

Logistic Growth:

$$\frac{dx}{dt} = rx\left(1 - \frac{x}{K}\right).$$

Where x = size of the population at a given time, r = growth rate, K = carrying capacity.

Lotka–Volterra equations:

$$\frac{dx_1}{dt} = r_1 x_1 \left(1 - \left(\frac{x_1 + \alpha_{12} x_2}{K_1}\right)\right)$$
$$\frac{dx_2}{dt} = r_2 x_2 \left(1 - \left(\frac{x_2 + \alpha_{21} x_1}{K_2}\right)\right).$$

Where α12 represents the effect species 2 has on the population of species 1 and α21 represents the effect species 1 has on the population of species 2. Utilising the competitive version of the model, all interactions (competition) and therefore all α-values are positive. The code was adapted from previously published Ecologial Competition Model by K Shoemaker https://kevintshoemaker.github.io/NRES-470/LECTURE16.html#Competition

The following R code was adapted based on the above to include two additional terms RS1 and RS2 to simulate daily reduction by radiation - corresponding to SF2.

```
System <- data.frame(n1 = rep(InitN1,(Ndays+1)),n2 = InitN2)

doday <- function(prevday){
  n1 <- {prevday[1] + prevday[1] * Rmax1 * (1-((prevday[1]+Alpha*prevday[2])/(K1)))  }* RS1
  n2 <- {prevday[2] + prevday[2] * Rmax2 * (1-((prevday[2]+Beta*prevday[1])/(K2)))  }* RS2
  return(c(n1,n2))
}

## Do simulation
for(i in 1:(Ndays+1)){
  System[1+i,] <- doday(System[i,])
}
```

n = population size
K = carrying capacity As population size (N) approaches carrying capacity (K), the numerator (K-N) becomes smaller but the denominator (K) stays the same and the second term decreases. The addition of this term describes a rate of population growth that slows down as population size increases, until the population reaches its carrying capacity
R = intrinsic rate of increase [r] times the population size [N]) describes a population's growth without competition
Alpha/Beta = competition coefficient effect of one sub-population on the others growth
RS = Radiosensitivity % of cells surviving 2Gy radiation differs between RESISTANT and SENSITIVE, assumes that RT is given over 33 consecutive days

Initial Growth Parameters

Alpha <- 0.8
Beta <- 1.2
InitN1 <- 10000000
InitN2 <- 100000
K1 <- 5000000000
K2 <- 5000000000
Rmax1 <- 0.62
Rmax2 <- 0.51
Ndays <- 42
RS1 <- 1
RS2 <- 1

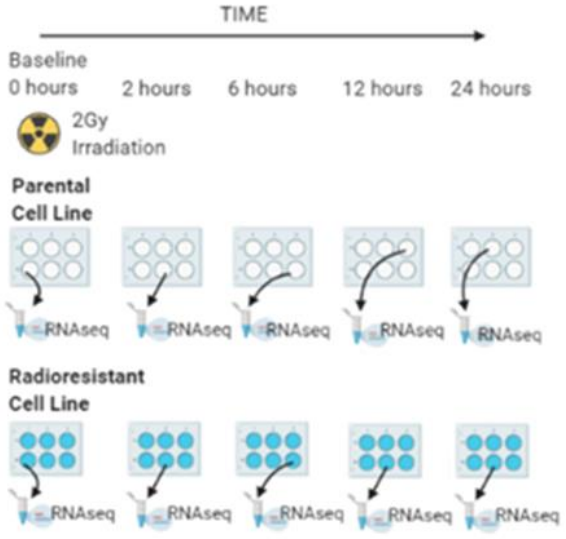
Radiation Treatment Parameters

Alpha <- 0.8
Beta <- 1.2
InitN1 <- System_time$n1[Ndays]
InitN2 <- System_time$n2[Ndays]]
K1 <- 5000000000
K2 <- 5000000000
Rmax1 <- 0.62
Rmax2 <- 0.51
Ndays <- 33
RS1 <- 0.52
RS2 <- 0.75

2.18 Cell Line Irradiation Experiments

Cells were plated into 6 well plates and grown overnight to sub-confluency. Cells were harvested at Time 0 (pre-radiation) and 2, 4, 6, 12, and 24 hours post irradiation with 2 Gy as above. The experiment was repeated for 3 biological replicates.

Figure 2.2. Schema of Cell Line RNA extraction



2.19 Cell line nucleic acid extraction

RNA extraction from Trizol (Invitrogen) was performed as per manufacturer's guidelines. Briefly this included addition of 0.75 mL TRIzol™ to each well and mechanical detachment of

cells using a cell scraper and homogenization by repeated pipetting. Followed by addition of 0.2 mL of chloroform, incubation for 2–3 minutes, centrifugation for 15 minutes at 12,000 × g at 4°C. The aqueous phase containing the RNA was transferred to a new tube by angling the tube at 45° and pipetting the solution out.

RNA was precipitated using 0.5ml isopropanol (Sigma-Aldrich), centrifuged at 12,000g at 4^{o}C and supernatant discarded. RNA was washed and resuspended in 1 ml 75% ethanol (Sigma-Aldrich), re-centrifuged and RNA solubilised in RNAse free water (Thermo Fisher). RNA integrity was confirmed using the absorbance technique at 260nm and 280nm to calculate concentration and A260/A280 ratio using a NanoDrop Spectrophotometer (Thermo Scientific).

DNA was extracted using the equivalent pathway for DNA - this involved isolation of DNA from the interphase and the lower phenol-chloroform phase. Precipitation of the DNA with 0.3ml 100% ethanol, washing of the DNA in 1 ml 0.1M sodium citrate in 10% ethanol, and solubilization of the DNA in 8nM NaOH (Gibco). DNA yield was quantified using the absorbance technique as above with a NanoDrop Spectrophotometer.

2.20 Cell line RNA sequencing alignment and quality assurance

3’ RNA sequencing (3’ RNA seq) was performed by the Wellcome Trust Centre for Human Genetics, using Lexogen 3’ RNA library preparation kit and Illumina Nextseq platform.

Prior to alignment sequences were trimmed and Fastq files were aligned to human genome Hg38 using STAR to generate gene counts. Counts were normalised prior to analysis of differential expression. Differential analyses were performed in R 3.52 using Deseq2, EDgeR and Limma packages and visualised with ClusterProfiler, pheatmap.

3’RNAseq Fastq files were concatenated per sample, and polyA and Illumina adapter trimming (AGATCGGAAGAGC) was performed using trimmomatic, prior to alignment with STAR to Hg38 (GRCh38.95.gtf), utilising the following parameters -m15, --nextseq-trim=20,-

a AGATCGGAAGAGC -a AAAAAAAAAAAAAAAAAAA -n2. Feature counts and Subread packages were used to generate counts and alternate transcripts.

3'RNAseq quality control (QC) was performed with FASTQC pre- and post-trimming.

Filtering of lowly expressed genes was performed, followed by differential expression analysis using DESEQ2, EDGER and limma. Over-representation analysis was performed using Clusterprofiler, enrichr and gprofiler. The GSVA package was used to perform ssGSEA.

2.21 Clinical Samples Identification of Patient Samples

Patient samples were retrospectively identified from the Prostate cancer Mechanisms of Progression and Treatment (PROMPT) cohort courtesy of Associate Professor Richard Bryant University of Oxford. Patient samples which had incomplete or missing clinical information on T/N/M, Gleason Score or relapse status were discounted and not included in the study.

2.22 Clinical samples Pathology assessment and macro-dissection

Original sample Formalin Fixed Paraffin Embedded (FFPE) sections and blocks were initially inspected by a Consultant Histopathologist (Thanks to Professor Clare Verrill Oxford University Hospitals) to determine whether they contained sufficient material. Sections were inspected and tumour containing regions marked up on a single slide which was digitally captured. Macro-dissection of tumour-containing regions was performed on 4 x 5µm unstained slides, using the digital reference slide, using sterile 10 blade scalpels, with multiple slides pooled to a single sample per patient.

2.23 Nucleic acid extraction, 3’ RNA sequencing, alignment and quality assurance

Samples were pooled per patient, and RNA extracted using the HighPure FFPE RNA (Roche) extraction kit as per the protocol. Briefly this involved deparaffinization with 800ul Xylene, and repeat centrifugation at 12,000g with 800ul 70% ethanol, followed by RNA extraction comprising addition of tissue lysis buffer, proteinase K digestion for 3 hours at 55^{o}C, sequential addition of binding buffers I & II and DNAse treatment, and repeat centrifugation through High Pure filter tubes followed by elution of RNA in elution buffer. RNA integrity was confirmed using the absorbance technique at 260nm and 280nm to calculate concentration and A260/A280 ratio using a NanoDrop Spectrophotometer.

8 patient samples contained sufficient material for 4 further slides to be utilised for DNA extraction using the FPPE high-pure DNA kit (Roche) per protocol which utilizes similar steps to above but with RNAse in place of DNAse.

3’ RNAseq sequencing library preparation was performed with QuantSeq 3'-mRNA-Seq Library Prep Kit (Lexogen), and libraries were sequenced on Illumina NextSeq flow cells (75 bp fragments) at the Wellcome Trust Centre for Human Genetics (University of Oxford).

For NanoString analysis, n=12 samples with optimal spectrophotometric, (determined by NanoDrop) characteristics were selected (comprising n=4 individuals from each group). Samples underwent nCounter® Human PanCancer Panel Gene Expression profiling according to the manufacturer’s instructions, and data was acquired using the nCounter® SPRINT profiler (Bryant laboratory University of Oxford).

For DNA methylation analysis, bisulfite conversion was performed using the EZ-96 DNA Bisulfite Zymo Research conversion protocol. The Illumina Infinium HD Methylation Assay protocol was followed, and samples were hybridised to Human MethylationEPIC beadchips (Illumina) by Genomics Birmingham (University of Birmingham).

RIN scores were not used to select sample suitability owing to the ability of Quantseq to utilise 75bp fragments. Samples were run on the Agilent 2100 bioanalyzer with RIN 2-3.

3'RNAseq quality control (QC) was performed with FASTQC pre- and post-trimming. One sample failed sequencing, with <100,000 counts, and was not used for further analysis. Although not strictly required for DESEQ2 analysis, counts were filtered to only include genes with >1 count in ≥5 samples (minimum group size).

For nanoString analysis, data were imported into nSolver^{TM} analysis software v2.5, and QC (Quality Control) was performed according to nanoString guidelines, with gene transcripts normalized to housekeeping genes. Differential expression analysis was conducted using the NSolver pipeline.

For DNA methylation QC was performed which included, filtering of poor performing probes, removal of cross-reactive probes, and removal of known SNPs at CpG sites), normalization was performed using the minifi package, prior to differential methylation analysis.

As a measure of the performance characteristics of our RNAseq dataset, we used the SigQC package to generate QC metrics for previously validated signatures (DECIPHER, Prolaris, Oncotype, and prostate hypoxia, each obtained from the original published literature). We compared the performance of these validated signatures in our dataset to two previously generated 'gold standard' datasets (the TCGA and Jain RRT). Sigcheck was used to assess the performance of identified gene sets compared to random genes and known signatures with 1000 iterations.

2.24 Additional Statistical Analysis

Sequencing and clinical data from the external dataset GSE116918 were downloaded from Gene Expression Omnibus (GEO). Probes were filtered to only include non-overlapping exonic probes, and multiple probes were merged to the mean value per gene.
In order to overcome technology platform differences as GSE116918 utilised a different sequencing technology (microarray) to those used in this study 3' RNA seq, mean values of fully exonic probes for the relevant genes were used from the GSE116918 dataset. nanoString probes are designed to span exons, such that only mature mRNAs are counted. The Quantseq method sequences mature transcripts with polyA tails, and this was used with the FeatureCounts and Subread package for quantification and identifications of alternate transcripts. Survminer and Survival packages were used for survival analysis with cox proportional hazards models utilizing known clinic-pathological features, Staging and Gleason score. KMunicate and survival packages were used for survival analysis with extended risk tables for time to event analysis.

2.25 Clinical Datasets Stratification in Colorectal Cancer (SCORT)– sequencing data and platform characteristics

Access to the dataset was kindly provided by Prof Timothy Maughan. Clinical datasets and sequencing data were downloaded from the SCORT bioportal utilizing mean per gene dataset. RNA sequencing was performed by the SCORT Consortium following deparaffinization and RNA extraction as described above, using hybridisation to XCel Micro-arrays. Normalisation was performed using (Robust MultiArray Algorithm (RMA) and batch-correction performed.

2.26 The Cancer Genome Atlas – sequencing data and platform characteristics

RNA was purified and distributed throughout the TCGA network and analysis platforms included exome and whole genome DNA sequencing, RNA sequencing. Clinical and sequencing data was obtained from the Genomic Data Commons Data Portal (https://portal.gdc.cancer.gov/)

2.27 GSE116918 - sequencing data and platform characteristics

Clinical and expression data was obtained from Gene Expression Omnibus (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE116918), and used tailored Affymetric microarray (Almac Diagnostics Prostate Disease Specific Array) [96].

CHAPTER 3

3.1 Radioresistant Cell Line Generation and Characterisation

3.11 Aims:

To generate radioresistant cell lines using clinically relevant dose and fractionation regimes and identify baseline characteristics of radioresistant cell lines compared to parental cell lines.

3.12 Cell Line Experiments

Parental cell lines were confirmed to be Mycoplasma free and were authenticated by STR testing. Mutational characteristics of the parental cell lines are demonstrated in Table 3.1, and represent commonly found mutations within colorectal and non-small cell lung cancer and have been demonstrated to be associated with radioresistance [13, 14, 15].

Table 3.1 Parental Cell Line Mutational Characteristics:

Cell Line	**Histology**	**Karyotype & modal chromosome number**	**Mutations**
H460	Non-Small Cell Lung Cancer (Large Cell)	Hypotriploid 57 (53-65)	P53 WT KRAS missense mutation; AA mutation p.Q61H PIK3CA missense mutation;** AA mutation p.E545K
A549	Non-Small Cell Lung Cancer (Adenocarcinoma)	Hypotriploid 66	P53 WT KRAS missense mutation;** AA mutation p.G12S KEAP1 missense mutation;** AA mutation p.G333C
HT29	Colorectal (Adenocarcinoma)	Hypertriploid 71	P53 missense mutation;** AA mutation p.R273H KRAS WT APC Frame Shift Insertion; AA mutation p.E1554Gfs*5 BRAF missense mutation; AA mutation p.V600E

**Data obtained from ATCC with additional mutational information from the COSMIC database.

3.12 *In vitro* cell line irradiation experiments

STR authenticated and Mycoplasma free parental cell lines were irradiated with clinically relevant dose fractionation, 60 Gy in 2 Gy per fraction, delivered monday to friday over 6 weeks, to generate radioresistant cell lines as per methods.

3.13 Morphological changes on irradiation

During the generation of radioresistant cell lines it was noted that a number of morphological changes occurred after irradiation. Cells became larger and stellate with a phenotype consistent with cellular senescence (appendix 2). Cells were taken from radioresistant lines generated using 2Gy/# schedule and were stained with Beta-galactosidase, demonstrating increased staining during irradiation consistent with cellular senescence.

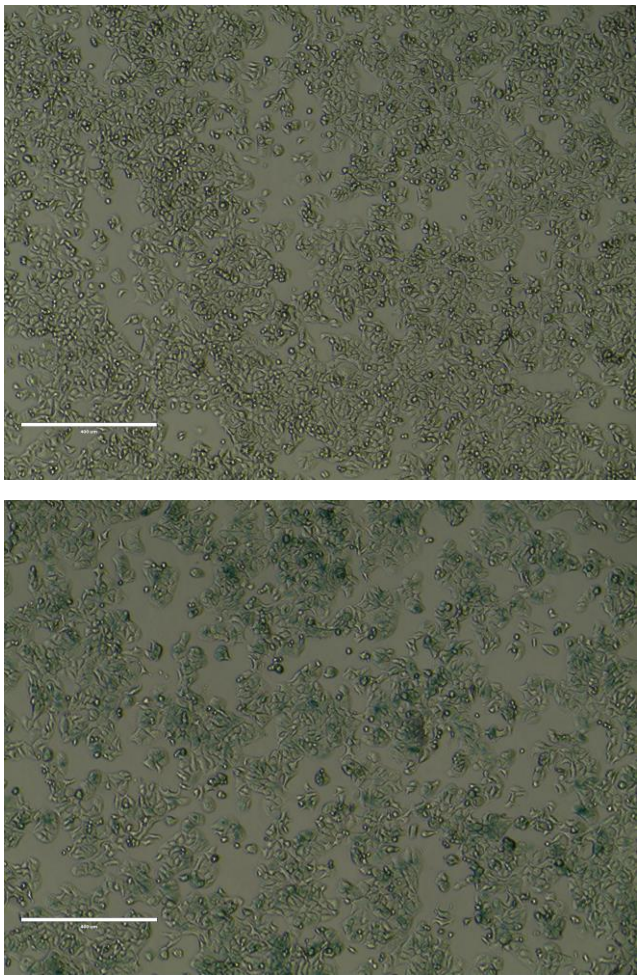


Figure 3.1: B-Galactosidase Staining in irradiated HT29 cell lines

Baseline/pre-radiation cells are shown in the top panel, Bottom panel shows cells 48 hours following 18Gy. Both panels are at 10x magnification (scale bar is 400μm)

3.14 A549 Radioresistant Cell Line

A549 cells survived fractionated irradiation and were allowed to recover for 2 weeks as per methods prior to clonogenic experiments. Clonogenic survival curves demonstrate that the surviving cell line is more radioresistant than the parental cell line with a higher mean SF2 of 0.75 compared to the parental line 0.52 . The difference in survival curves was confirmed to be statistically significant by analysis of variance (ANOVA) with a p value of 2.29×10^{-10}, (F Value = 38.69) comparing models with 2 separate curves vs a single curve.

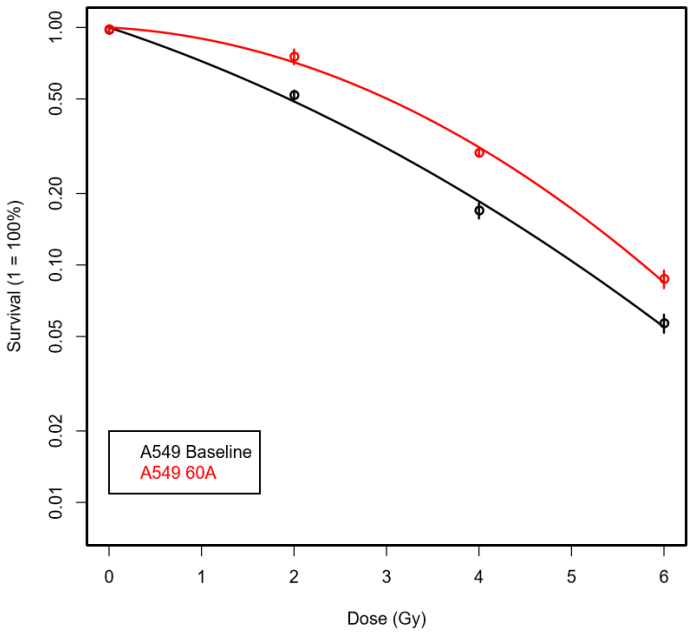


Figure 3.2 Cell Line Clonogenic Survival Curves, A549 Baseline and A549 60A
Clonogenic survival curves, mean survival curve of 3 experiments, each individual experiment comprises one 6 well plate per dose. Full survival curves in Appendix Figure 1.2. Analysis of variance (ANOVA) p value of 2.29×10^{-1}

Table 3.2 Mean Surviving Fraction per Dose A549 Cell Lines

CELL LINE	Mean Surviving Fraction 2Gy +/- SD	Mean Surviving Fraction 4Gy +/- SD	Mean Surviving Fraction 6Gy +/- SD
A549 BASE	**0.52** +/-0.023	**0.17** +/-0.012	**0.06** +/- 0.005
A549 60A	**0.75** +/- 0.054	**0.29** +/- 0.010	**0.083** +/- 0.008

3.15 H460 Radioresistant Cell Line

H460 cells surviving 6 weeks of fractionated irradiation (60Gy) were designated H460 60A. Fractionated irradiation for one of the cell lines was terminated early after 5 weeks (50Gy in 2 Gy per fraction), owing to reduced numbers of surviving cells, this cell line was designated H460 50B. Clonogenic survival curves demonstrate that the surviving cell lines are more radioresistant than the parental cell line with a higher mean SF2 of 0.48 (H460 60A) and 0.57 (H460 50B) compared to the parental line 0.47. The difference in survival curves between H460 60A and the parental line was confirmed to be statistically significant by analysis of variance (ANOVA) with a pvalue of 2.2 $x10^{-6}$, (F Value = 23) comparing models with 2 separate curves vs a single curve.

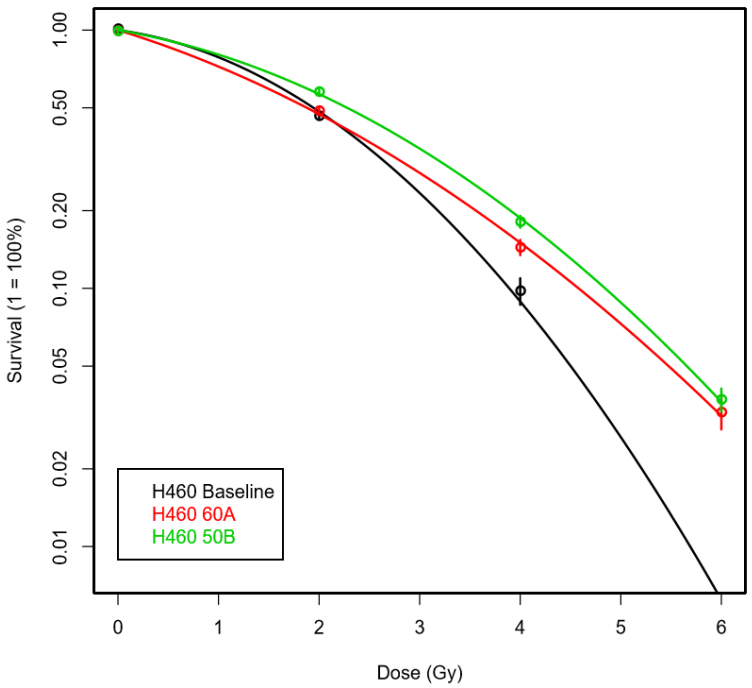


Figure 3.3 Cell Line Clonogenic Survival Curves H460 Baseline, H460 60A, and H460 50B

Clonogenic survival curves, mean survival curve of 3 experiments, each individual experiment comprises one 6 well plate per dose. Full survival curves in Appendix Figure 1.3. Analysis of Variance demonstrates p value = $5.4x10^{-10}$ pH460A vs Baseline p value = $2.2x10^{-6}$ H450B vs Baseline

Table 3.3 Mean Surviving Fraction per dose H460 Cell Lines

CELL LINE	Mean Surviving Fraction 2Gy +/- SD	Mean Surviving Fraction 4Gy +/- SD	Mean Surviving Fraction 6Gy +/- SD
H460 BASELINE	**0.47** +/-0.016	**0.10** +/-0.012	**0.005** +/- 0.0006
H460 60A	**0.48** +/- 0.015	**0.14** +/- 0.099	**0.033** +/- 0.0038
H460 50B	**0.57** +/- 0.017	**0.18** +/- 0.097	**0.037** +/- 0.0047

3.16 HT29 Radioresistant Cell Lines

HT29 cells survived fractionated irradiation, two cell lines were generated HT29 60A and HT29 60B. Clonogenic survival curves demonstrate that the surviving cell lines are more radioresistant than the parental cell line with a higher mean SF2 of 0.74 and 0.72 (HT29 60A and HT2960B respectively) compared to the parental line 0.62. The difference in survival curves was confirmed to be statistically significant by analysis of variance (ANOVA) comparing models with 2 separate curves vs a single curve.

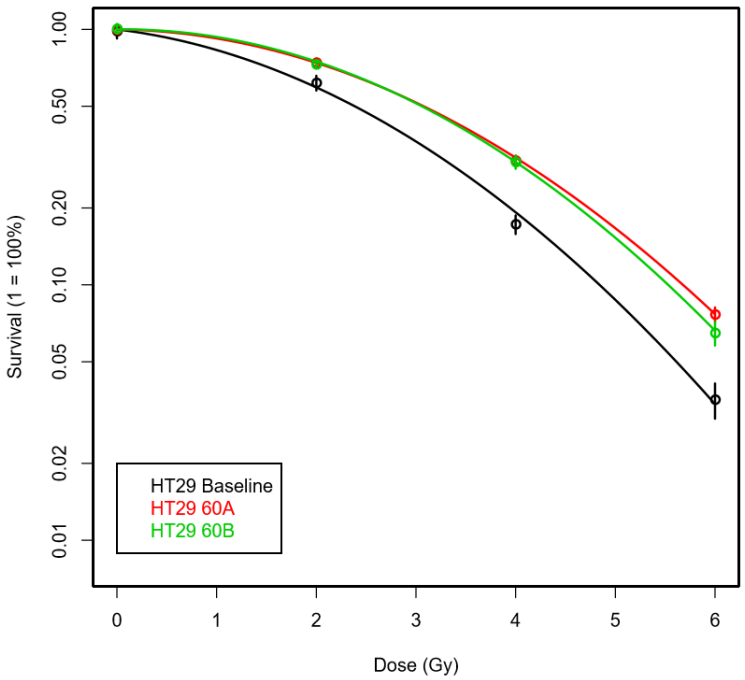


Figure 3.4 HT29 Survival Curves

Clonogenic survival curves, mean survival curve of 3 experiments, each individual experiment comprises one 6 well plate per dose. Full curves in Appendix Figure 1.3. Analysis of Variance demonstrates p value = 3.4x10-10 HT2960A vs Baseline p value = 3.7x10-6 HT2960B vs Baseline

Table 3.4 Mean Surviving Fraction per Dose HT29 Cell Lines

CELL LINE	Mean Surviving Fraction 2Gy +/- SD	Mean Surviving Fraction 4Gy +/- SD	Mean Surviving Fraction 6Gy +/- SD
HT29 BASE	**0.617** +/- 0.062	**0.172** +/- 0.04	**0.0356** +/- 0.006
HT29 60A	**0.741**+/- 0.022	**0.306** +/- 0.028	**0.076** +/- 0.005
HT29 60B	**0.72** +/- 0.025	**0.302**+/- 0.018	**0.06**+/-0.007

3.17 Review of Literature for published SF2

To establish that experimentally derived values of SF2 were consistent with those previously published in the literature a review by Matsui et al [16] of the previously published SF2s was undertaken for the parental cell lines. From the literature mean SF2s, ranges and standard deviations were calculated. For each of the parental cell lines experimental results for SF2 are within one standard deviation of the mean of previously published results see Table 3.5.

Table 3.5 Comparison of experimentally derived SF2s with the literature.

CELL LINE	SF2	Standard Dev	Range	No of Studies	Experimental Result
A549	0.66	+/- 0.14	0.25-0.84	192	**0.52** +/-0.023
H460	0.62	+/-0.16	0.24-0.94	95	**0.47** +/-0.016
HT29	0.70	+/- 0.12	0.42-0.94	67	**0.617** +/- 0.062

3.18 Characterisation of radioresistant cell lines cell cycle analysis H460

To assess whether parental and radioresistant cell lines had different proportions of cells within the cell cycle, cell cycle analysis was performed using flow-cytometry to identify relative proportions of each cell line and the mean of 3 biological replicates calculated. The proportions of cells in G1, G2 and S phase were not significantly different between H460 Baseline, H460 60A and H460 50B (Figure 3.5)

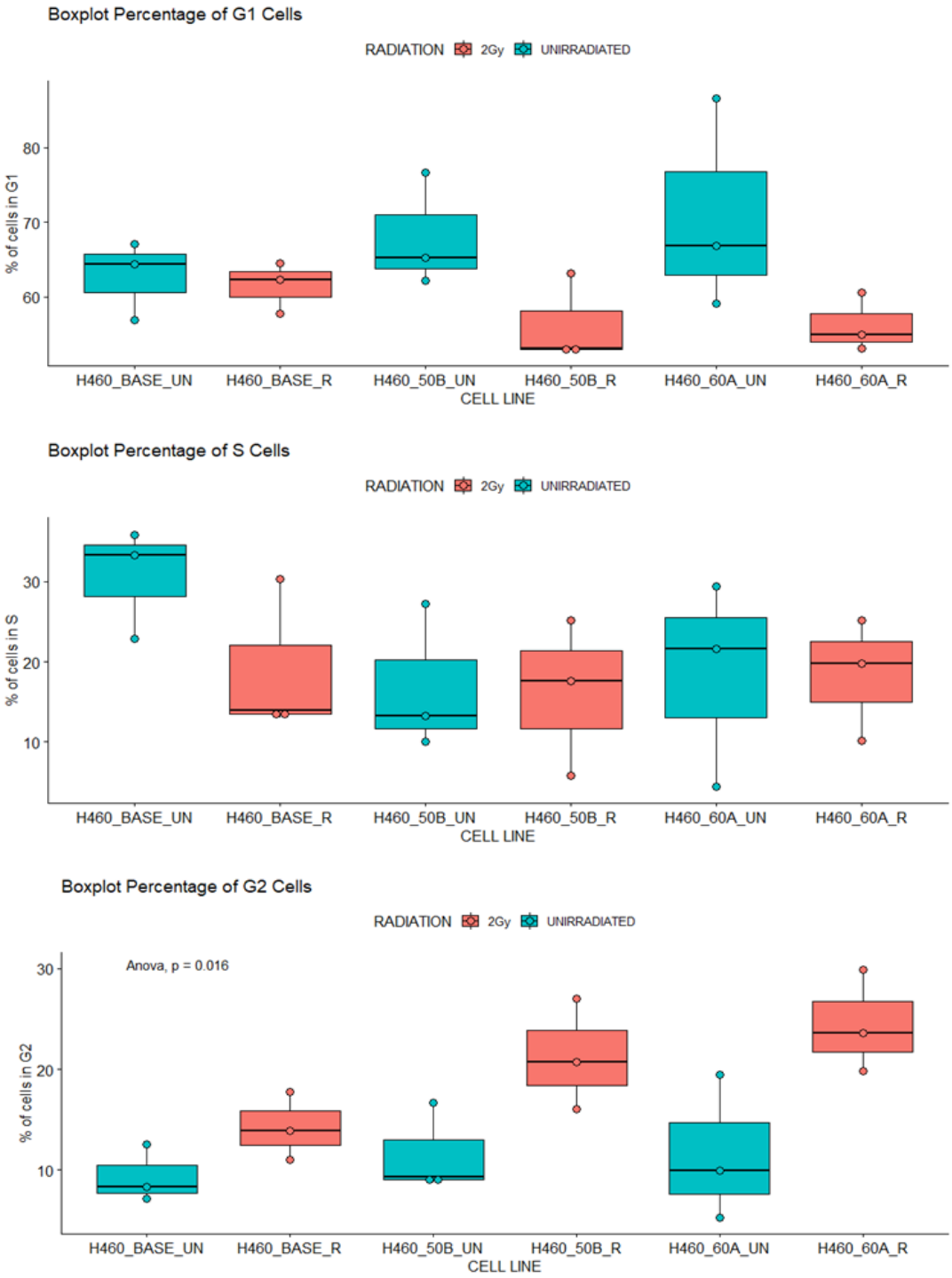


Figure 3.5 Boxplot of percentage of Cells in stages of Cell Cycle pre- and post-radiation.

Cell cycle analysis was performed using flow-cytometry to identify relative proportions of each cell line, pre and 24 hours post 2Gy irradiation. Mean of 3 biological replicates calculated. Analysis of Variance demonstrates significance in percentage difference in G2 between irradiated and unirradiated samples ($p = 0.016$).

Significant differences in the proportion of cells in G2 phase was identified pre-and post-radiation, with an average of 9% of cells in G2 in unirradiated H460 parental line, compared

to 14% 24 hours post 2Gy further irradiation. Radioresistant cell lines demonstrated a greater proportion of cells in G2 pre- and post- further 2Gy irradiation, 11.6 vs 21.2% and 11.5% vs 24.4% for H46050B and H460 60A cell lines respectively. Tables 3.6 and 3.7 and Figure 3.5. The overall difference between the proportion of cells in G2 pre- and post-radiation was significant by ANOVA p = 0.016 however this difference between individual lines was non-significant H460 parent 0.2, H460 50B 0.2, and H460 60A (0.1 Appendix).

Table 3.6 Cell Cycle Proportions Pre Radiation

CELL LINE	REPLICATE	RADIATION	G1 (%)	S (%)	G2 (%)
H460_BASE_UN	1	UNIRRADIATED	56.9	35.8	8.35
H460_BASE_UN	2	UNIRRADIATED	64.4	33.3	12.6
H460_BASE_UN	3	UNIRRADIATED	67.1	22.9	7.13
			62.8	30.66667	9.36
H460_50B_UN	1	UNIRRADIATED	76.7	10	9.28
H460_50B_UN	2	UNIRRADIATED	65.3	13.2	16.7
H460_50B_UN	3	UNIRRADIATED	62.2	27.2	8.77
			68.06667	16.8	11.58333
H460_60A_UN	1	UNIRRADIATED	66.9	21.6	19.5
H460_60A_UN	2	UNIRRADIATED	86.6	4.46	5.31
H460_60A_UN	3	UNIRRADIATED	59.1	29.4	9.93
			70.86667	18.48667	11.58

Table 3.7 Cell Cycle Proportions Post-Radiation

CELL LINE	REPLICATE	RADIATION	G1 (%)	S (%)	G2 (%)
H460_BASE_R	1	2Gy	62.3	13.9	13.9
H460_BASE_R	2	2Gy	57.8	30.3	11
H460_BASE_R	3	2Gy	64.5	13	17.8
			61.53333	19.06667	14.23333
H460_50B_R	1	2Gy	53.1	25.2	16.1
H460_50B_R	2	2Gy	63.2	5.79	27
H460_50B_R	3	2Gy	53	17.6	20.7
			56.43333	16.19667	21.26667
H460_60A_R	1	2Gy	53.1	25.2	19.8
H460_60A_R	2	2Gy	60.6	10.1	29.9
H460_60A_R	3	2Gy	55	19.8	23.6
			56.23333	18.36667	24.43333

Tables 3.6 and 3.7 demonstrate the differences in the mean of 3 biological replicate experiments, in the proportion of cells in each phase of the cell cycle as measured by flow cytometry comparing pre- and 24 hours post- repeat irradiation with 2Gy.

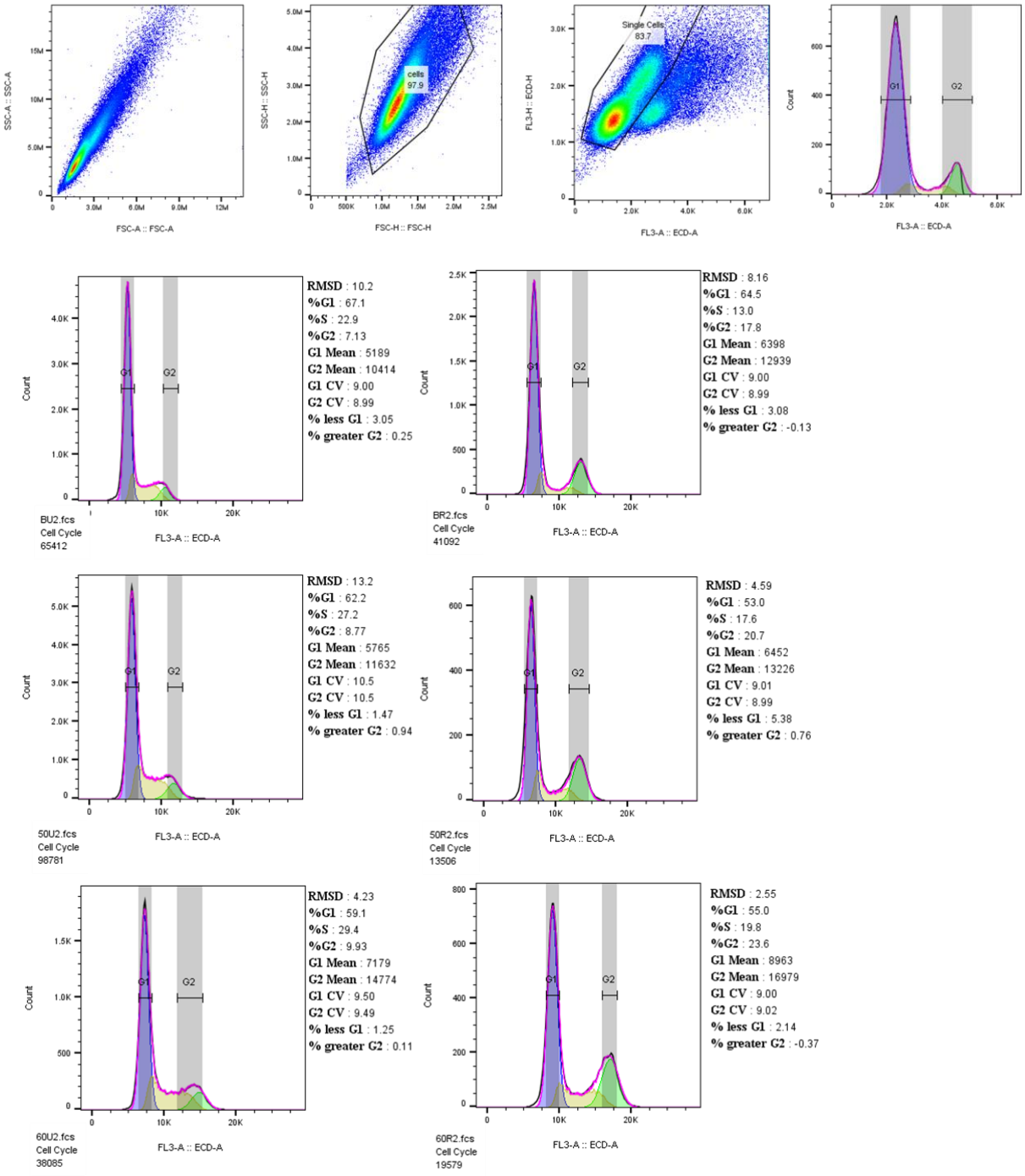


Figure 3.7 Cell Cycle Analysis H460, H460 60A, H460 50A, pre and post radiation.

Flow Cytometry of Propidium Iodide stained H460 parental, H460 50B and H460 60A cell lines pre- and 24 hours post further 2 Gy of radiation demonstrating an increase in the proportion of cells in G2 for each cell line. Top panel demonstrates gating of single cells for cell cycle analysis.

3.19 Characterisation of radioresistant cell lines proliferation H460

To assess whether there were differences in proliferation between the derived radioresistant lines (H460 50B and parental cell line (H460) proliferation assays were performed using CCK8 assay. Radioresistant cell lines were demonstrated to be slower proliferating compared to parental cell lines (Table 3.8, Figure 3.8).

Table 3.8 Proliferation Assay - H460 Baseline, H460 60A H460 50B

CELL LINE	TIME	AVERAGE ABSORBANCE	Standard Deviation
H460	0	0.33	0.003
	24	0.724	0.204
	48	2.152	0.21
	72	2.90	0.45
H460 50B	0	0.330	0.01
	24	0.385	0.05
	48	1.14	0.25
	72	2.19	0.397
H460 60A	0	0.200	0.03
	24	0.37	0.065
	48	0.903	0.065
	72	1.86	0.08

Proliferation assays were performed using 3x biological replicates for each cell line, each biological replicate contained 3x technical replicates. Average absorbance and Standard deviations were calculated.

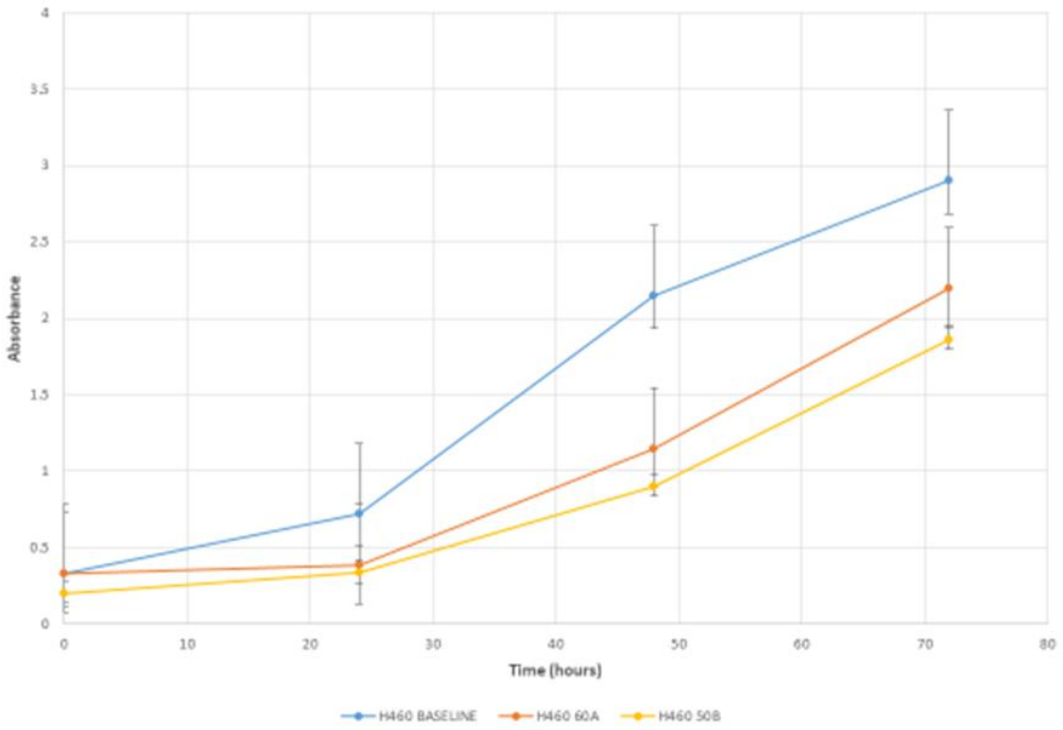


Figure 3.8 Proliferation Assay -
H460 Baseline, H460 60A, H460 50B cell lines were incubated in CCK8 and y axis demonstrates Absorbance (460 nm), x axis demonstrates time. H460 baseline demonstrated more rapid proliferation than irradiated cell lines.

3.20 Characterisation of radioresistant cell lines senescence H460

To further investigate the findings during the generation of radioresistant cell lines Figure 3.1 of the development of senescent cells, Beta-Galactosidase assays were performed to investigate the proportion of senescent cells in radioresistant and parental cell lines pre- and 24 hours post a further irradiation with 2Gy. The results demonstrate a difference in the proportion of senescent cells between pre- and post-irradiation in the radioresistant cells lines bordering on significant. (Figure 3.9). Differences between the percentage of senescent cells between the parental and radioresistant cell lines prior to repeat irradiation was not significant.

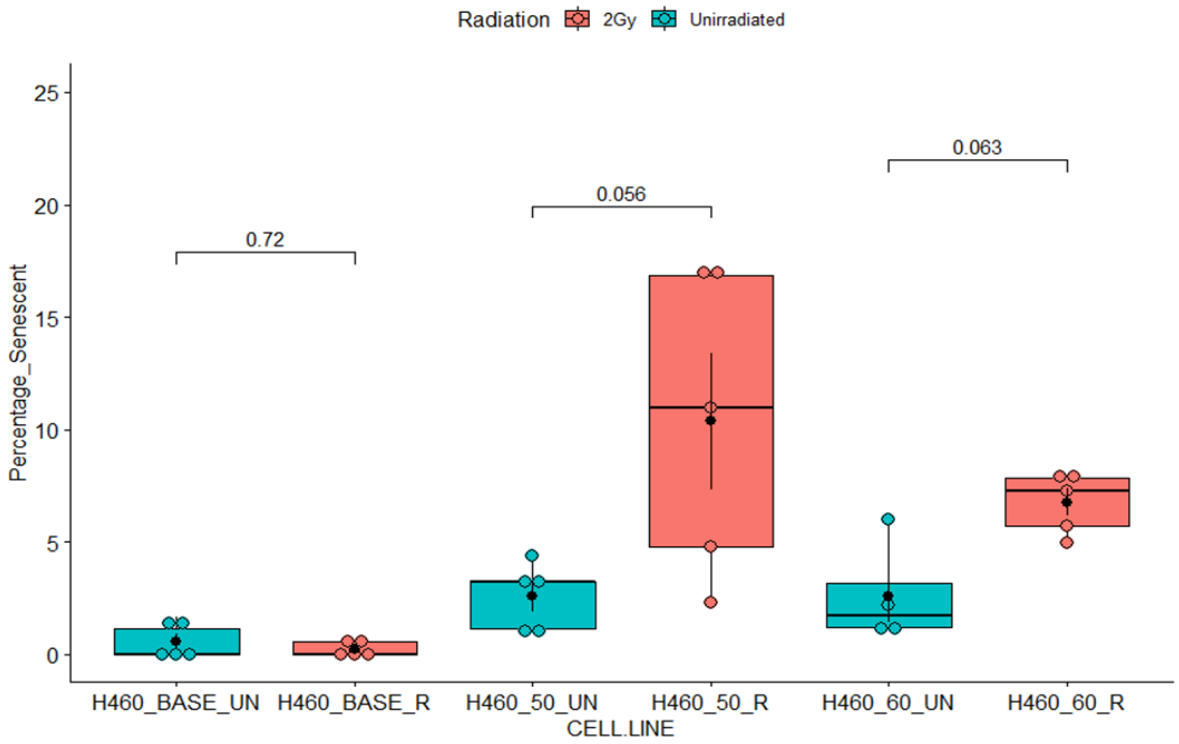


Figure 3.9 Senescence Assay (Beta Galactosidase)
Cell line comparisons pre- and 24 hours post- irradiation demonstrated non-significant difference in percentage of senescent cells. 3x biological replicates were performed per condition.

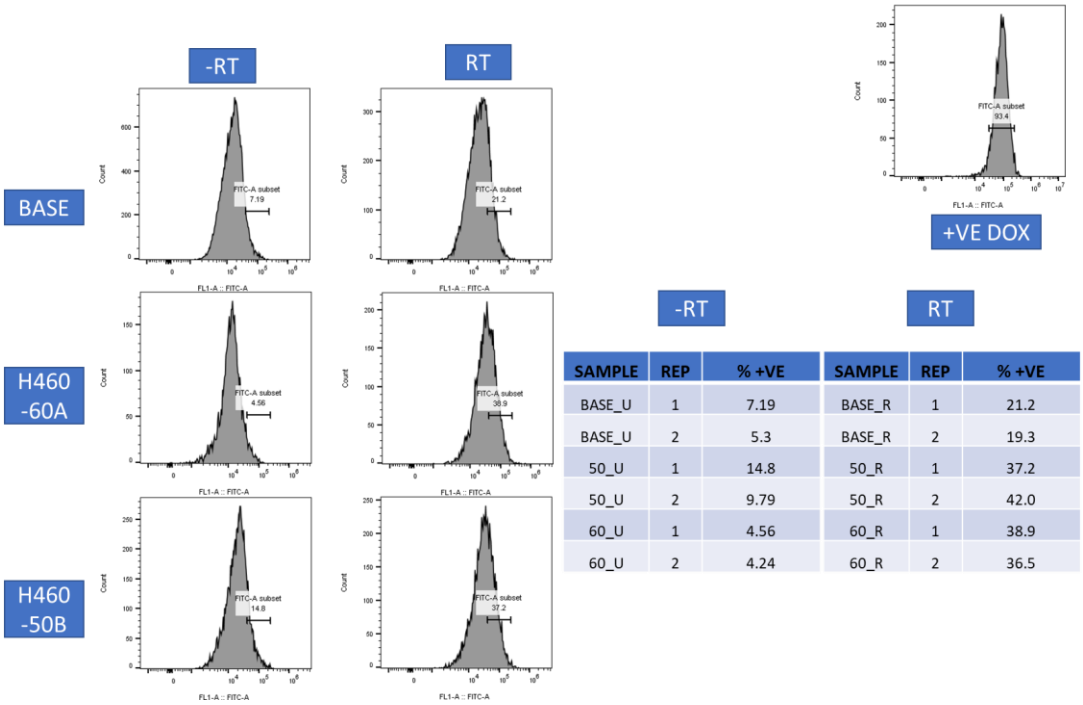
Table 3.9 Senescence Assay (Beta Galactosidase)

	% Senescent Cells	
Cell Line	**Non-Irradiated**	**2Gy**
H460 BASE	0.45	0.23
H460_50	2.69	10.96
H460_60	2.82	7.07

3.21 Characterisation of radioresistant cell lines senescence flow cytometry

To further investigate these findings, Flow Cytometry with 5-Dodecanoylaminofluorescein Di-β-D-Galactopyranoside (C12 FDG) was conducted on parental and radioresistant H460 cell lines pre- and 24 hours post- a further 2 Gy irradiation utilising Doxorubicin treated cells as a positive control (see Methods).

These results demonstrate similar findings to the Beta Galactosidase senescence assay with an increase in the proportion of senescent cells between cell lines pre- and 24 hours post further 2Gy irradiation. The experiment was repeated and the mean of the results of 2 biological replicate experiments was obtained (Figure 3.10, Table 3.10).



SAMPLE	REP	% +VE	SAMPLE	REP	% +VE
BASE_U	1	7.19	BASE_R	1	21.2
BASE_U	2	5.3	BASE_R	2	19.3
50_U	1	14.8	50_R	1	37.2
50_U	2	9.79	50_R	2	42.0
60_U	1	4.56	60_R	1	38.9
60_U	2	4.24	60_R	2	36.5

Figure 3.10 Senescence Analysis - C12 FDG Flow Cytometry
Flow Cytometry with 5-Dodecanoylaminofluorescein Di-β-D-Galactopyranoside (C12 FDG) was conducted on parental and radioresistant H460 cell lines pre- and 24 hours post- a further 2 Gy irradiation utilising Doxorubicin treated cells as a positive control. Two experimental replicates were performed for each condition.

The proportion of senescent cells pre-radiation is again seen to be higher in the radioresistant cell lines compared to the parental cell lines both before and after repeat irradiation. This is consistent with the findings of the previous Beta Galactosidase senescence assay, suggesting that there is a higher proportion of senescent cells within the radioresistant cell lines and that a higher proportion of cells become senescent after repeat irradiation compared to the parental cell lines.

Table 3.10 Senescence Analysis - C12 FDG Flow Cytometry

SAMPLE	RADIATION DOSE	% SENESCENT (mean of 2 replicates)
H460 BASELINE	Nil	6.25
	2Gy	20.25
H460 50	Nil	12.3
	2Gy	39.6
H460 60	Nil	4.4
	2Gy	37.7

3.22 Characterisation of radioresistant cell lines senescence stemness assay H460

In order to assess whether there was a higher proportion of stem-like cells in the radioresistant cell lines a ALDH1 Assay was conducted on one of the radioresistant cell lines H460-60A. This assay did not demonstrate any increase in ALDH1 activity in the radioresistant cell line suggesting that there is not an increase in stem-like phenotype within the H460 radioresistant cell lines.

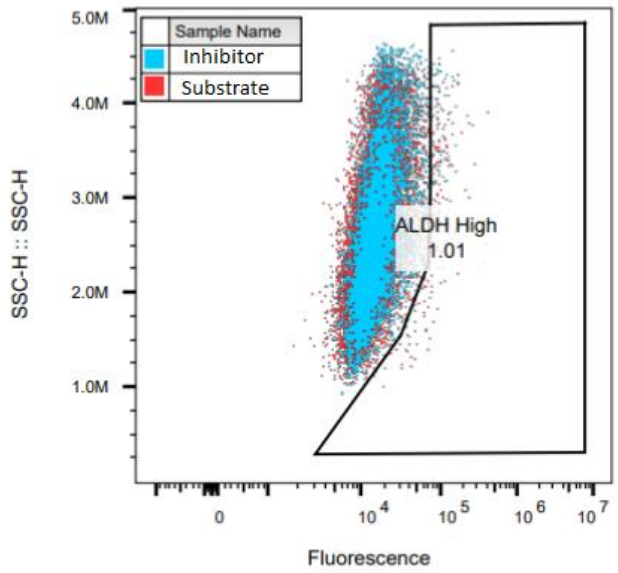


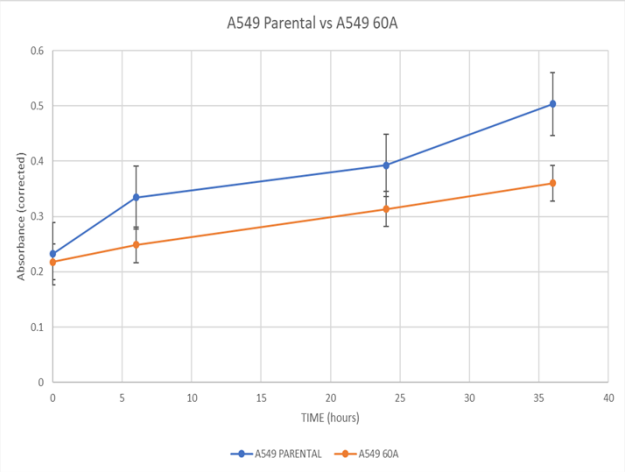
Figure 3.11 ALDH1 assay H460-60

Flow Cytometry of H460-60 cells using ALDEFLUOR kit demonstrates no high fluorescence population comparing substrate and inhibitor

3.23 Characterisation of radioresistant cell lines proliferation A549

A549 radioresistant cells A549 60A were demonstrated to have reduced proliferation measured by CCK8 assay than their parental cell lines A549 Figure 3.9, Table 3.9. The slower proliferation observed in radioresistant cell lines complements the finding in the comparison of H460 cell lines suggesting that radioresistant cell lines are slower growing.

Table 3.11 A549 and A549 60A Cell Line Proliferation



A549 BASELINE	MEASURED ABSORBANCE			AVERAGE	SD
TIME	1	2	3		
0	0.253	0.219	0.226	0.232667	0.017954
6	0.352	0.308	0.343	0.334333	0.023245
24	0.394	0.38	0.403	0.392333	0.01159
36	0.438	0.533	0.539	0.503333	0.05666

A549 60A	MEASURED ABSORBANCE			AVERAGE	SD
TIME	1	2	3		
0	0.225	0.236	0.193	0.218	0.022338
6	0.281	0.247	0.218	0.248667	0.031533
24	0.305	0.367	0.269	0.313667	0.049571
36	0.321	0.397	0.363	0.360333	0.03807

Figure 3.12 A549 and A549 60A Cell Line Proliferation

3.24 Ecological Modelling A549 growth

Utilising data obtained from the cell line experiments for each cell line growth rate over 24 hours (0.62 for A549 parental and 0.51 for A549 60A) and SF2 0.52 for A549 parental and 0.79 for A549 60A) an ecological model was constructed to model tumour heterogeneity with respect to cell proliferation and radiosensitivity. This model relied on several assumptions;

i) that one sub-population is more rapidly growing and initially larger in number,

ii) the sub-population growing more rapidly has a negative effect on the other sub-population near carrying capacity.

iii) the slower growing sub-population is more radioresistant.

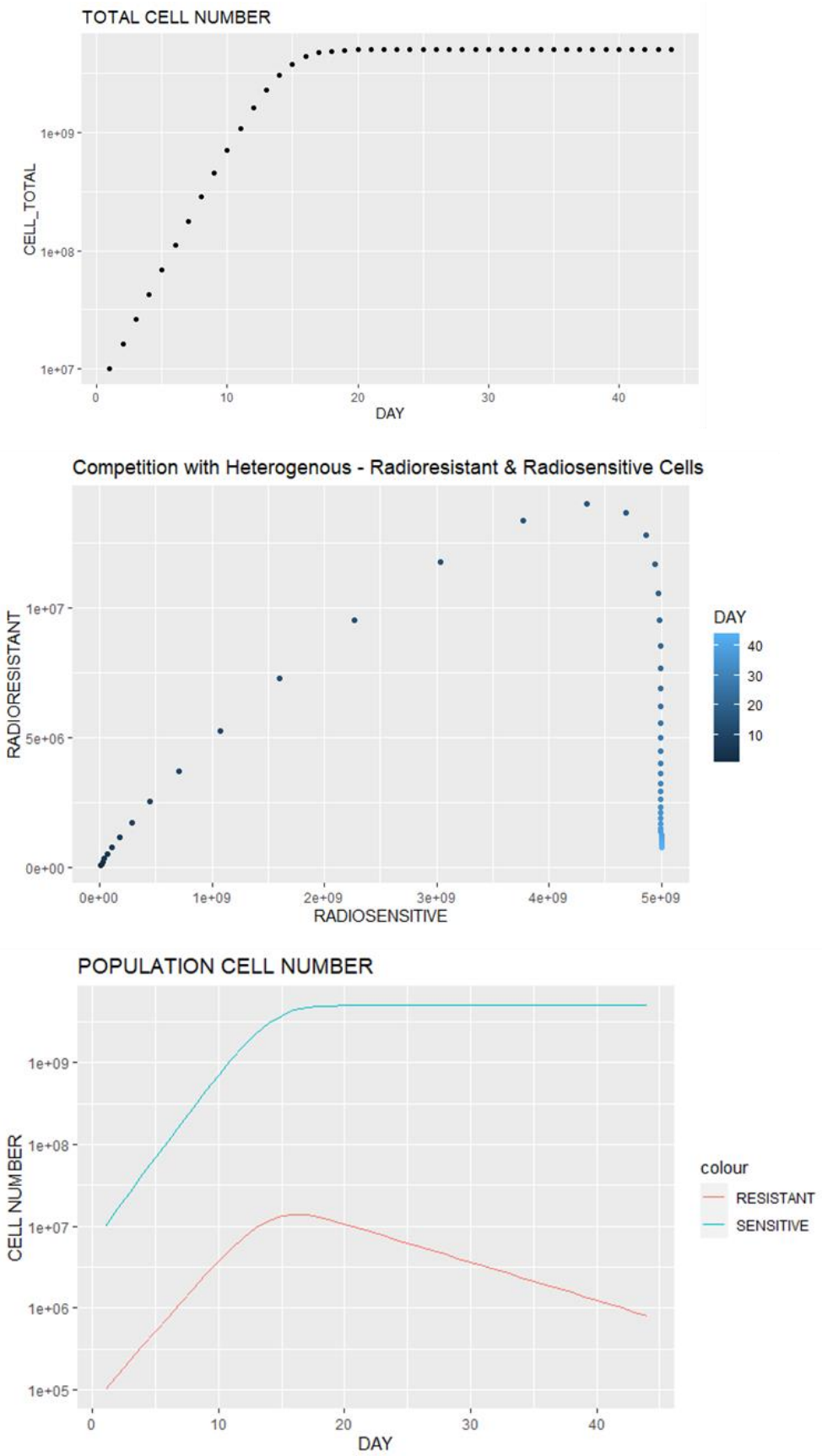


Figure 3.13 Ecological Modelling A549 Growth Phase

Top panel shows total cell number over time, middle panel shows radioresistant (y axis) and sensitive populations (x axis) over each timepoint, bottom panel cell numbers of each population over time

Ecological modelling of a heterogeneous tumour population identifies that the sub-population of slower growing radioresistant cells in initial growth conditions, are outcompeted by the more rapidly growing radiosensitive cells once the population is sufficiently large to reach carrying capacity (i.e. limitations of growth resources). It can be demonstrated that total cell number increases over time, nominally 40 days, until it reaches a maximum population based on carrying capacity (Figure 3.13 Top panel). The relative cell numbers of each population over time (radioresistant on the y axis, radiosensitive on the x axis) demonstrate a steady increase in both populations until carrying capacity is reached followed by an increase in rapidly growing radiosensitive cells and a decrease in slower growing radioresistant cells (Figure 3.13 Middle Panel). The cell number of each population over time, again demonstrating that the rapidly growing radiosensitive cells continue to slowly increase whilst the number of radioresistant slow growing cells continues to decrease over time (Figure 3.13 Bottom Panel), a finding which should be taken in the context of the top panel whereby total cell number remains constant, i.e. the proportion of cells is changing but the overall population size remains the same.

3.25 Ecological Modelling A549 under fractionated radiation

To model the effect of radiation further terms were introduced to model the reduction in cell numbers of each subpopulation through exposure to radiation. To model clinically relevant doses this was modelled as 2 Gy per fraction over 33 consecutive days (66Gy total) utilising the experimentally derived values for SF2 of 0.52 for A549 parental and 0.79 for A549 60A to reduce the relevant population by each day.

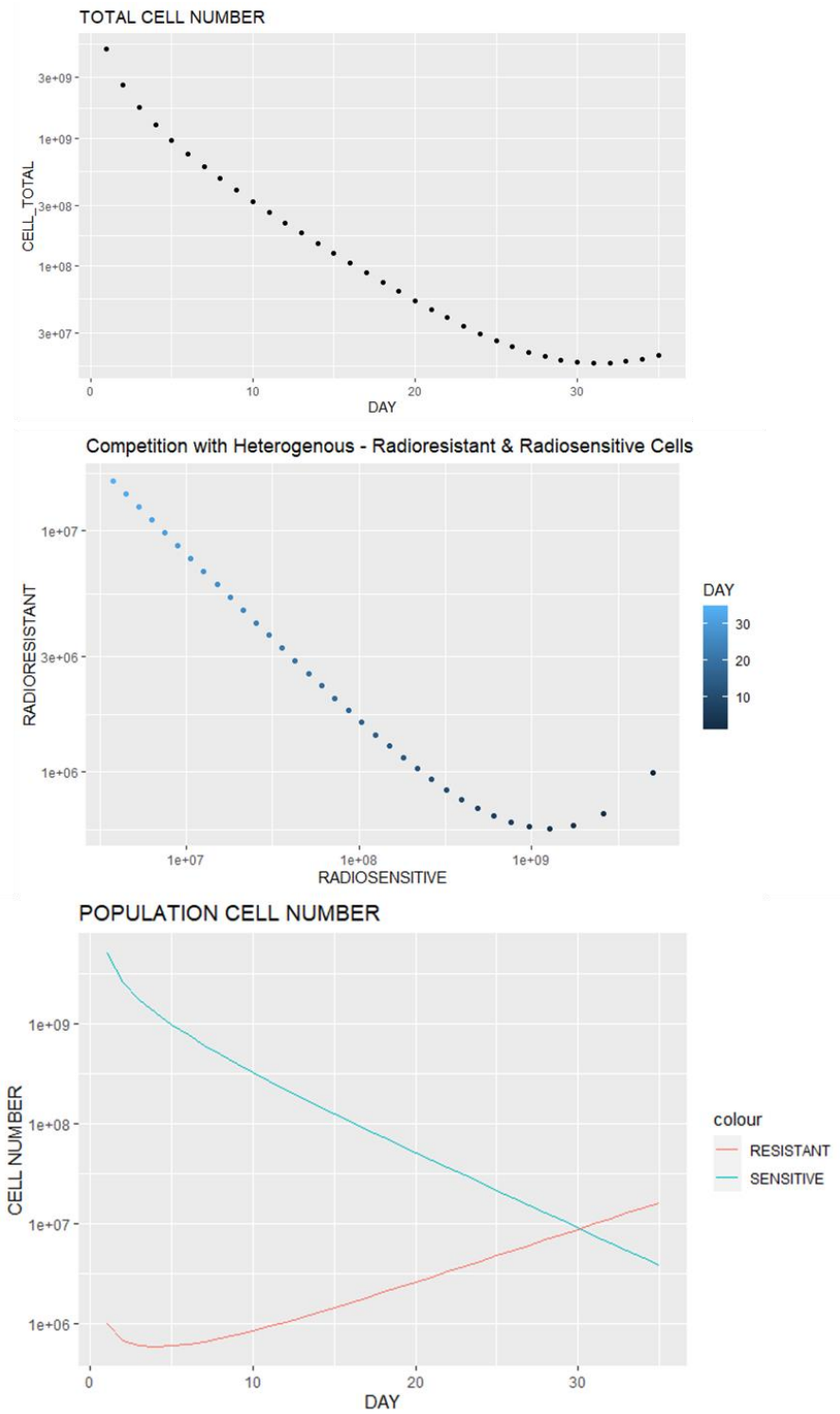


Figure 3.14 Ecological Modelling A549 during fractionated radiation
Top panel shows total cell number over time, middle panel shows radioresistant (y axis) and sensitive populations (x axis) over each timepoint, bottom panel cell numbers of each population over time

Figure 3.14 Top Panel demonstrates the total cell numbers over the course of treatment decreases in a logarithmic manner over time. The middle panel demonstrates relative cell

numbers of each population over time (radioresistant on the y axis, radiosensitive on the x axis) demonstrating an initial decrease in both populations followed by an increase in radioresistant cells and an ongoing decrease in radiosensitive cells. The bottom panel of Figure 3.14 demonstrates the population cell numbers of both radioresistant and radiosensitive sub-populations over time. Notably the radiosensitive cells demonstrate a logarithmic decrease during treatment, however the radioresistant cells, despite initially decreasing, actually increase in number throughout the course of treatment. Taken into context with the top panel, whilst the total population of cells is decreasing the smaller radioresistant sub-population is actually increasing in number.

3.26 Chapter Summary

The experimental results have demonstrated that radioresistant cell lines can be successfully generated from repeat irradiation using clinically relevant doses and fractionation (Tables 3.2 -3.4). Experimental results for parental cell lines (SF2) are demonstrated to be within one standard deviation of previously published results in the literature (Table 3.5). The degree of radioresistance measured as Surviving Fraction at 2 Gy (SF2) induced by irradiation differed, between H460 60A (0.48 +/- 0.015) and H460 50B (0.57 +/- 0.017), of note H460 50B cell line only received 50 Gy irradiation, due to reduced cell numbers during repeat irradiation this was terminated one week earlier than planned. Despite similar SF2s radioresistant cell line H460 60A survival curves are statistically significant to parental lines by analysis of variance (ANOVA) with a p value of 2.2×10^{-6} (Figure 3.3).

Cell cycle analysis identified a greater mean percentage of cells were in the more radioresistant G2 phase in radioresistant lines than parental, 11.5% vs 9.3% (Table 3.6) at baseline (pre-radiation) and 24 hours post further 2Gy irradiation (21.2% and 24.4% vs 14.2%).

The proportion of senescent cells in the radioresistant and parental cell lines pre- and 24 hours post further irradiation, identifies an increase in the percentage of senescent cells in both radioresistant (H460 50B 12.3-39.6%, H460 60A 4.4-37.7%) and parental (6.25 - 20.25%) cell lines by C12FDG Flow Cytometry (Table 3.10), with a higher percentage of senescent cells in the radioresistant cells than parental at both timepoints. Similar findings also seen in Beta-Galactosidase experiments (Figure 3.9)

The ALDH1 assay did not demonstrate any increase in ALDH1 activity in the radioresistant cell line suggesting that there is not an increase in stem-like phenotype within the H460 radioresistant cell lines.

Radioresistant cell lines were demonstrated to have a slower proliferation rate by CCK8 assay than radiosensitive cell lines in both H460 and A549 cell lines (Figure 3.8, 3.11)

Ecological modelling demonstrates that the slower growing radioresistant population is outcompeted by rapidly growing radiosensitive populations in normal growth conditions when approaching carrying capacity, but under daily irradiation conditions, the number of radioresistant cells can actually increase during treatment (Figure 3.14).

3.27 Chapter Discussion

The experimental results have demonstrated that radioresistant cell lines were successfully generated from repeat irradiation using clinically relevant doses and fractionation (Tables 3.2 -3.4). To place the experimental results in the context of the existing literature, previously published results were identified and a comparison performed which demonstrated that the experimentally derived results for parental cell lines (SF2) are within one standard deviation of those in the literature (Table 3.5). This provides support that the experimental methods employed produce results that are consistent with established findings.

The degree of radioresistance measured as Surviving Fraction at 2 Gy (SF2) induced by irradiation differed, between H460 60A (0.48 +/- 0.015) and H460 50B (0.57 +/- 0.017), of note H460 50B cell line only received 50 Gy irradiation, due to reduced cell numbers during repeat irradiation this was terminated one week earlier than planned. Despite similar SF2s radioresistant cell line H460 60A survival curves are statistically significant to parental lines by analysis of variance (ANOVA) with a p value of 2.2 x10-6 (Figure 3.3).

Given the duration of the radiation treatment (2Gy per day monday to friday over 6 weeks) a potential confounding factor explaining differences between parental and radioresistant lines could be due to prolonged passage, which has previously been reported in the literature as contributing towards radioresistance in A549 cells[142]. The parental lines were however passage matched to the cells undergoing irradiation, which during irradiation did not reach confluence, and underwent regular media change only (see Methods).

Cell cycle analysis identified a greater mean percentage of cells were in the more radioresistant G2 phase in radioresistant lines than parental, 11.5% vs 9.3% (Table 3.6) at baseline (pre-radiation) and 24 hours post further 2Gy irradiation (21.2% and 24.4% vs 14.2%). The findings for radioresistant cell lines are compatible with previous studies investigating the response of fibroblasts (IMR90 cell line) 24 hours following 5Gy irradiation

where 28.26% of cells were found to be in G2 phase[143] and SQ-5 human lung squamous cell carcinoma treated with daily 2Gy irradiation for 14 days were also found to demonstrate prominent G2/M accumulation[144]. H460 proliferation assays further support this finding demonstrating lower rates of proliferation in the radioresistant cell lines than radiosensitive (Figure 3.8). Previous studies have demonstrated that radioresistant cells demonstrate a prolongation of G2/M phase during cell cycle in response to IR and lower proliferation rates. Previous studies by Peng demonstrated that in Nasopharyngeal radioresistant cell lines (CNE2-R), irradiation induced a G2/M phase arrest which was recoverable, in contrast to radiosensitive NPC cells (CNE2), where the arrest was permanent[145]. This study further identified that radioresistant cells whilst maximally accumulating in G2 arrest at 20 hours post-irradiation, were able to progress through the cell cycle at 28 hours whereas radiosensitive cells were not, with a higher proportion of cells undergoing apoptosis[145]. Taken together these results suggest that the radioresistant cell lines can undergo G2 arrest, repair sub-lethal damage and re-enter the cell cycle, in contrast to the radiosensitive cells.

Taking forward the idea that tumour populations may be heterogeneous in both terms of radioresistance and growth rates, with the slower growing cell lines being radioresistant, a competitive ecological model was used to identify the difference between population growth with and without fractionated radiotherapy. This demonstrated that whilst slower growing radioresistant populations are outcompeted by rapidly growing radiosensitive populations at carrying capacity under normal growth conditions, under the survival pressure of daily irradiation, the number of radioresistant cells can actually increase during treatment (Figure 3.11). In the context of the total cell number decreasing over the course of treatment, this could potentially be a mechanism of treatment failure shortly following treatment in the clinical setting, particularly in the case of lung cancer. The polyclonal nature of lung cancer primaries has been extensively demonstrated, with widespread intratumor heterogeneity of somatic copy-number alterations and heterogenous driver mutations found in more than 75% of the tumours including genes involved in chromatin modification and DNA damage

response and repair[146]. Such tumour heterogeneity, taken together with finite resources available to a growing tumour due limitation to blood supply and diffusion of nutrients (i.e. carrying capacity) provide a plausible opportunity for competitive ecological models to model the situation in clinical practice. This has been demonstrated in experimental studies where mixed populations of prostate cancer cell lines (consisting of radioresistant and parental cell lines) grown in spheroids, were demonstrated to have reduced radiation sensitivity[147]. Ecological modelling predicted that these changes were the result of competition between populations resulting in differential population growth dependent on the population heterogeneity. One key addition of the ecological model outlined in Figure 3.14 is that it has been designed to incorporate not only the growth in normal conditions, but also the growth and competitive interactions under daily fractionated radiotherapy, as is the case in clinical practice. This enables the model to estimate the changing proportions of radioresistant and radiosensitive cell subpopulations throughout the course of 30 days of treatment and define the relative proportions of each subpopulation following treatment. Other work to model tumour heterogeneity utilising a non-ecological model revealed similar findings through a modified standard radiation response curve accounting for variations in the repopulation rates of the tumour cell subpopulations across a distribution of radiosensitivity[55].

Previous work has identified radioresistance as being caused by radiation-induced changes of stem-like cell content[148]. This was not demonstrated to be the case in the H460 60A cell-line with ALDH+ assay not demonstrating a notable stem-like population (Figure 3.12). During irradiation to generate radioresistant cell lines, cells were noted to undergo morphological changes which were confirmed to be due to cellular senescence with B-Galactosidase staining (Figure 3.1). This finding led to the measurement of the proportion of senescent cells in the radioresistant and parental cell lines pre- and 24 hours post further irradiation, identifying an increase in the percentage of senescent cells in both radioresistant (H460 50B 12.3-39.6%, H460 60A 4.4-37.7%) and parental (6.25 - 20.25%) cell lines by

C12FDG Flow Cytometry (Table 3.10), with similar findings also seen in Beta-Galactosidase experiments (Figure 3.9)

Senescence is known to be induced by ionising radiation and believed to confer resistance to therapy via a number of mechanisms including epigenetic modification via chromatin modulation and is associated with immune stimulation following DNA damage in the secretory phenotype[149-154]. One of the recent key findings is that senescence, which was previously considered to be a permanent process, can be dynamic[153]. The H3K9-active demethylases JMJD2C and LSD1 have been demonstrated to mediate escape from oncogene induced senescence in both fibroblast and melanocyte models[155], the H3K9 histone methyltransferase Suv39h1 or p53 is required for senescence, specifically for the maintenance of therapy-induced senescence (TIS) in various malignancies[156]. It can be noted that the cell lines H460 and A549 are both P53 wild-type and HT29 has a p53 missense mutation (Table 3.1).

When considering the limitations of the experimental design one of the most notable is that these are cell line studies, utilising adherent two-dimensional growth in flasks. This is not representative of the clinical situation where tumours exist as three-dimensional structures with complex micro-environments and diffusion gradients of both nutrients and oxygen. Studies have demonstrated increased chemotherapy resistance and radioresistance of colorectal cell lines and lung adenocarcinoma cell lines when grown in spheroids[157,158]. Another limitation of this study is that the ecological model utilised an assumption for the competition ratio, although it can be noted the other parameters for growth and SF2 were experimentally derived. To further this work and to further validate the ecological model findings further experimental work could be undertaken utilising cell culture in spheroids, not only to more accurately recreate clinical conditions but to further quantify the growth parameters utilised in competition modelling through experimentally derived competition ratios.

One slight area of inconsistency in the results relates to the increased proportion of senescent cells and the increase in percentage of cells in G2 following radiation, as senescence was classically defined as an irreversible cell cycle arrest in G1 phase. Increasing evidence suggests that senescence may possibly result in G2 arrest however an additional experiment to measure cell cycle and senescence simultaneously would have been the only way to answer that question. The fact that the radioresistant cells are present in a higher proportion 24 hours following radiation in the more radiosensitive G2 phase is an area of inconsistency in the data which could potentially have been addressed by such a combined experiment.

The aim of this Chapter was to generate radioresistant cell lines using clinically relevant dose and fractionation regimes and identify baseline characteristics of radioresistant cell lines compared to parental cell lines. The results in Chapter 3 demonstrate this aim has been achieved and radioresistant cell lines successfully generated using clinically relevant dose and fractionation regimes. The baseline characteristics of radioresistant cell lines (growth rate/cell cycle/senescence) were identified and further characterised in response to further radiation (2Gy) and compared to parental cell lines. This additional characterisation provides an additional understanding of potential mechanisms of radio-resistance e.g. induction of senescence. Competition modelling incorporating experimentally derived parameters demonstrates a potential mechanism whereby tumour heterogeneity with a slower growing more radioresistant sub-population can become the dominant population of tumour cells under the selective pressure of daily fractionated radiation.
The radioresistant cell-lines derived in this Chapter are further investigated at the molecular level through RNA sequencing and methylation analysis in Chapter 4 to identify molecular features associated with radioresistance.

CHAPTER 4

4.1 Radioresistant cell line gene expression and methylation analysis

4.11 Chapter 4 Aims:

Previous review of the literature identified significant variation in the methodology of studies exploring gene expression characteristics of radioresistant cell lines, leading to a minimal overlap of findings between studies. This informed the design of an experiment whereby gene expression changes associated with radioresistance could be identified at baseline and in a dynamic manner following repeat irradiation through a time-course experiment. This Chapter aims to identify molecular features associated with a radioresistant phenotype through gene expression and methylation analysis, and identify the predictive value of features associated with radioresistant cell line phenotype in clinical cohorts

4.12 Radioresistant cell line experiment overview

RNA samples from generated radioresistant cell lines pre- and post-repeat irradiation as per the cell line experimental Method, briefly this involved samples from parental and radioresistant cell lines at baseline (Time 0) and 2, 6, 12 and 24 hours post further 2Gy irradiation with 3 biological replicates. 3’ RNA sequencing of all samples was performed in a single cohort. Methylation studies at baseline were also performed for a subset of radioresistant and parental cell lines.

4.12 Radioresistant cell line experiment sequencing Quality Control

Quality Control metrics were obtained for samples using FastQC and comparison metrics for different mapping pipelines and techniques obtained (Appendix Table 2.1).

In a comparison of multiple alignment pathways used for comparison, the STAR + trim method was deemed to be the optimal method producing the highest number and percentage of uniquely mapped reads, and fewest samples with <1 million reads.

A comparison of Differential Expression techniques was also performed on the data with significant overlap in results seen, therefore DESEQ2 was used for the primary analysis (Appendix 2.1).

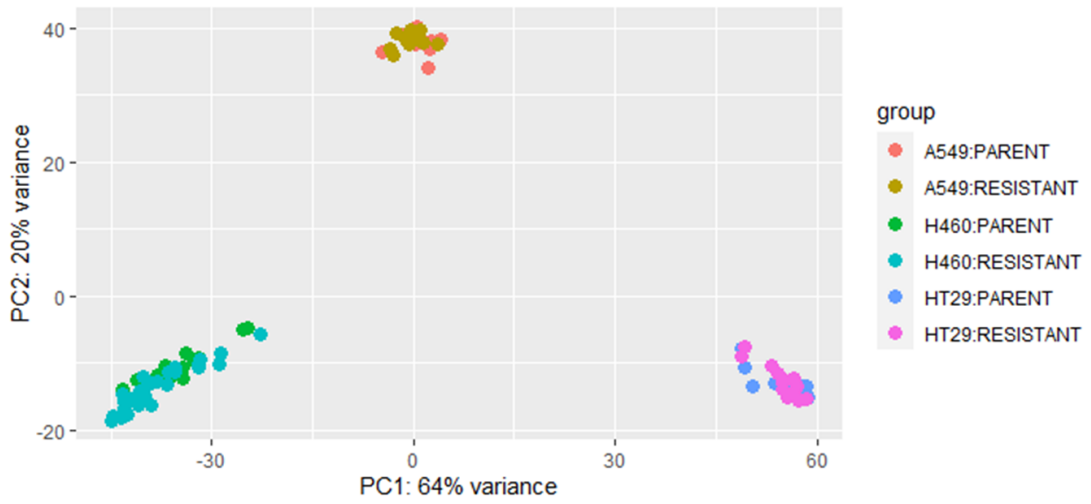


Figure 4.1 3' RNAseq Analysis Principal Component Analysis All Cell Lines
PCA plot of all samples demonstrating clustering by cell-line

Initial comparison of all cell lines via Principal Component Analysis (PCA) suggests that Radioresistant cells lines cluster closely with parental cell lines, suggesting that the majority of differences between cell lines is due to baseline differences, rather than changes due to radioresistance. This is consistent with the literature review findings suggesting that the majority of genes identified in previous studies comparing different cell lines with different radiosensitivities, were due to the intrinsic differences between cell lines. In order to

investigate whether later Principal Components were able to separate radioresistant from parental lines PC3-6 were plotted (Figure 4.2).

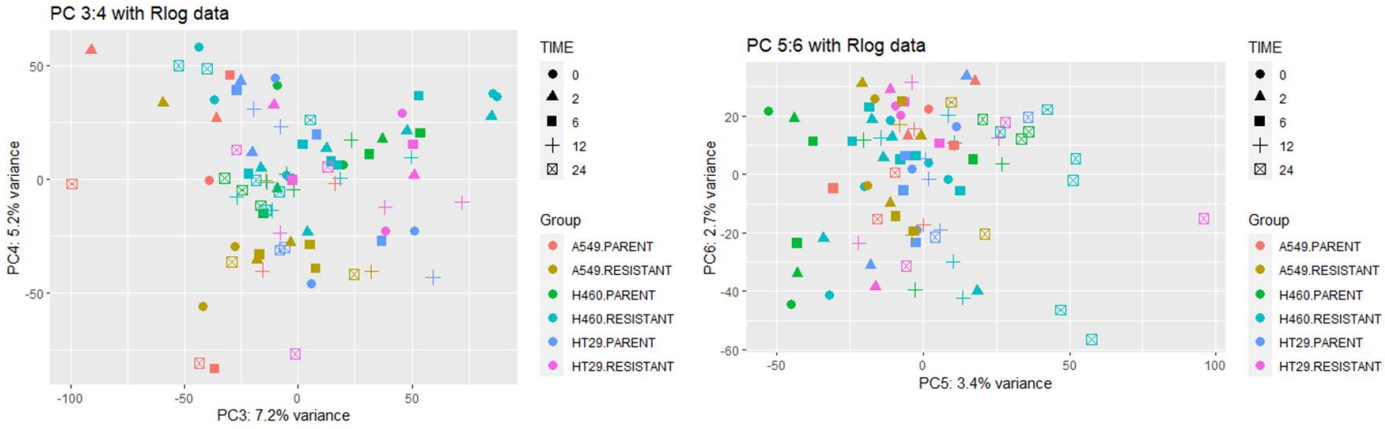


Figure 4.2 3' RNAseq Analysis Principal Component Analysis (PC3-6)
Principal components 3:4 (left) and 5:6 (right) for all samples.

There is no clear separation of parental and resistant cell lines across all cell lines, in the first 6 Principal Components (PCs). PC6 (Figure 4.2 right panel) demonstrates a clear separation of the majority of samples from Time 24 (crossed boxes) to the right. Based on the findings at PCA further analysis between resistant and parental samples was done on an individual cell line basis.

4.13 H460 cell line sequencing Principal Component Analysis

Comparisons between the parental H460 and 2 radioresistant lines H460-60A and H460-50B were conducted using Principal Component Analysis (PCA) to explore the gene expression data and group factors.

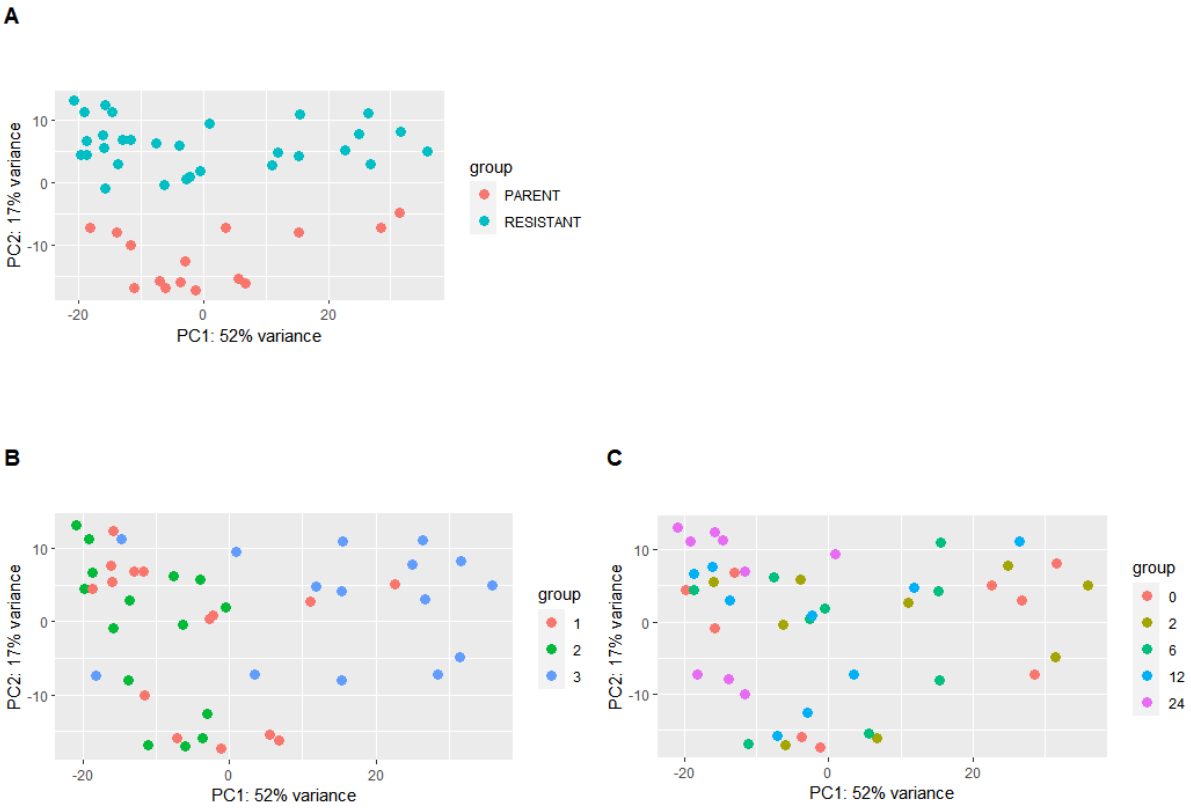


Figure 4.3 3' RNAseq Analysis PCA - H460 Cell Lines
Principal Component Analysis with sample groups identified by; A) Class (Parent, Resistant), B) Batch (biological replicate) and C) Time (0, 2, 6, 12, 24 hours post irradiation.

The PCA demonstrates clear separation between Parent and Resistant samples by the second Principal Component (Panel A) accounting for 17% of variation. Variation is also demonstrated between samples by biological replicate (Panel B). To take these effects into account they were included in the analysis model for differential expression. Variation due to time is less clear (Figure 4.3 Panel C), most Time 24 samples cluster to the left of PC1.

4.14 H460 Pre-Radiation (Time 0) Resistant vs Parent

Initial comparison was made between the Radioresistant and radiosensitive cell lines pre-radiation (Time 0), before comparisons at each point (Time 2, 6, 12, 24) after radiation.

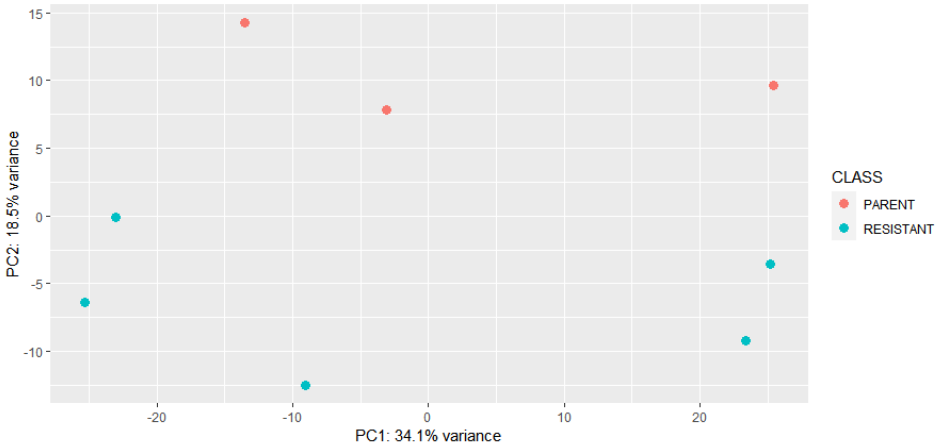


Figure 4.4 Principal Component Analysis H460 Time 0
Principal Component 1 (x-axis) vs Principal Component 2 (y axis) plot H460 Parent vs Resistant

Analysis of the pre-radiation samples (Time 0) confirms the findings of the group PCA with Parental and Resistant samples separating on the second Principal Component (PC2). Given the findings of PCA factors to account for batch (biological replicate) and class (parent vs resistant) were included in the analysis model for differential expression analysis.

4.15 Differential Expression Analysis at Baseline

Differential expression Analysis using DESEQ2 between Resistant and Parental cell line at time 0 (pre-radiation) identified 352 significant genes at an adjusted p value of <0.05 (Figure 4.5). Of these 276 genes were increased when comparing resistant to parental samples defined by Log Fold Change (LFC) > 1, and 76 decreased, defined by Log Fold Change (LFC) < -1 (Figure 4.5 Volcano Plot).

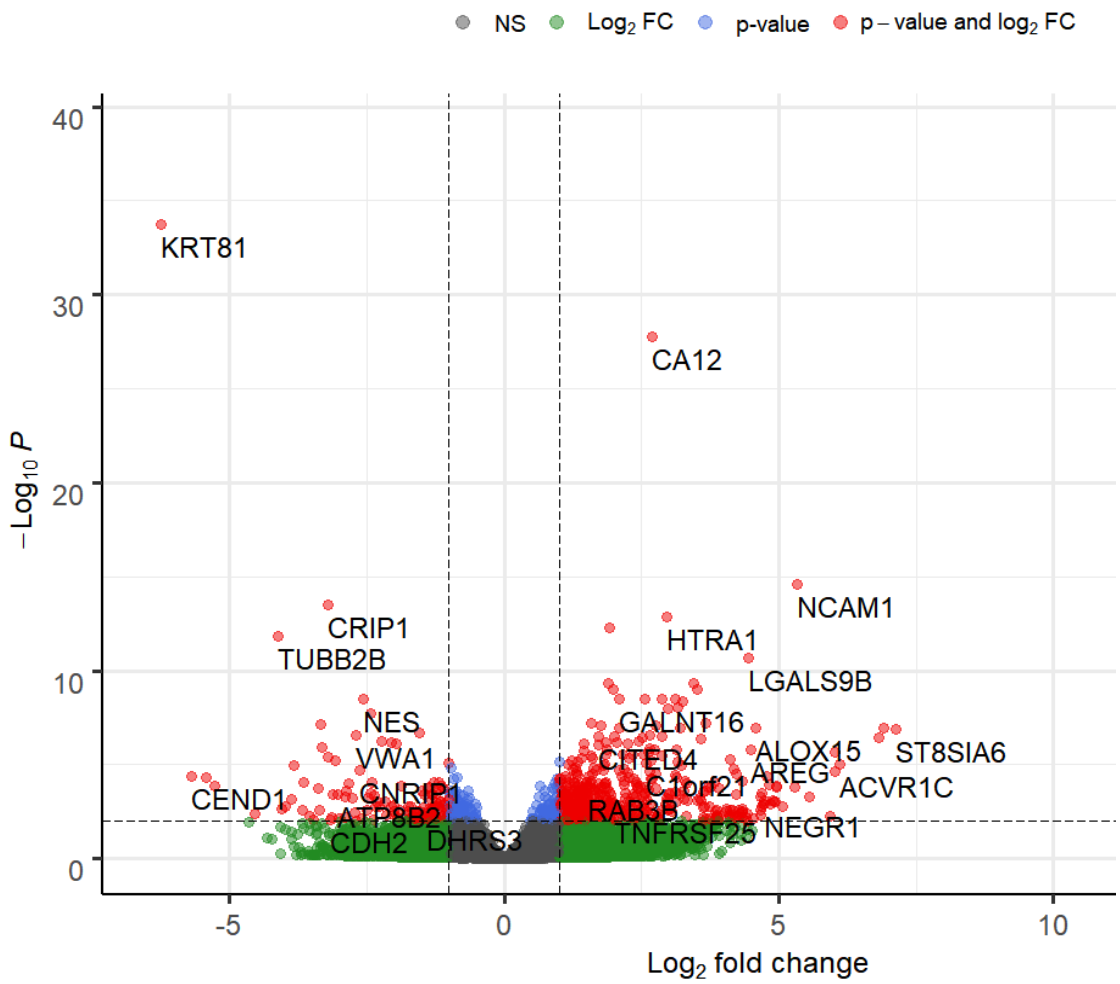


Figure 4.5 Volcano Plot of Differentially expressed genes Resistant vs Parent(Time 0)

Volcano plot comparing resistant to parental cell lines, 257 genes increased (LFC >1), 55 genes (LFC < -1) decreased at adjusted p valuej <0.05

In order to further investigate the biological roles of the significantly differentially expressed genes, pathway enrichment was performed using Over-Representation Analysis (ORA) of Gene Ontology (GO) genesets. Analysis was performed separately on increased and decreased genesets defined by Log Fold Change as previously (LFC >1/ LFC < -1).

For the 55 genes decreased in radioresistant samples no pathways were significantly enriched at an adjusted pvalue of <0.05. For the 257 genes increased in radioresistant samples, pathways that were significantly enriched (adjusted pvalue of <0.05) included Extracellular matrix, Growth Factor binding, Antigen binding and MHC protein complex

(Figure 4.6 Panel A). To further explore whether the overlap between pathways was due to the inclusion of the same group of genes in multiple different pathways, a heatplot was generated (Figure 4.6 Panel A). This demonstrates that although there are clearly overlapping genesets e.g. extracellular matrix, pathways are enriched due to expression of distinct sets of genes related to each pathway. Full enrichment analysis results including significance, number of overlaps and overlapping genes is provided in Appendix Table.

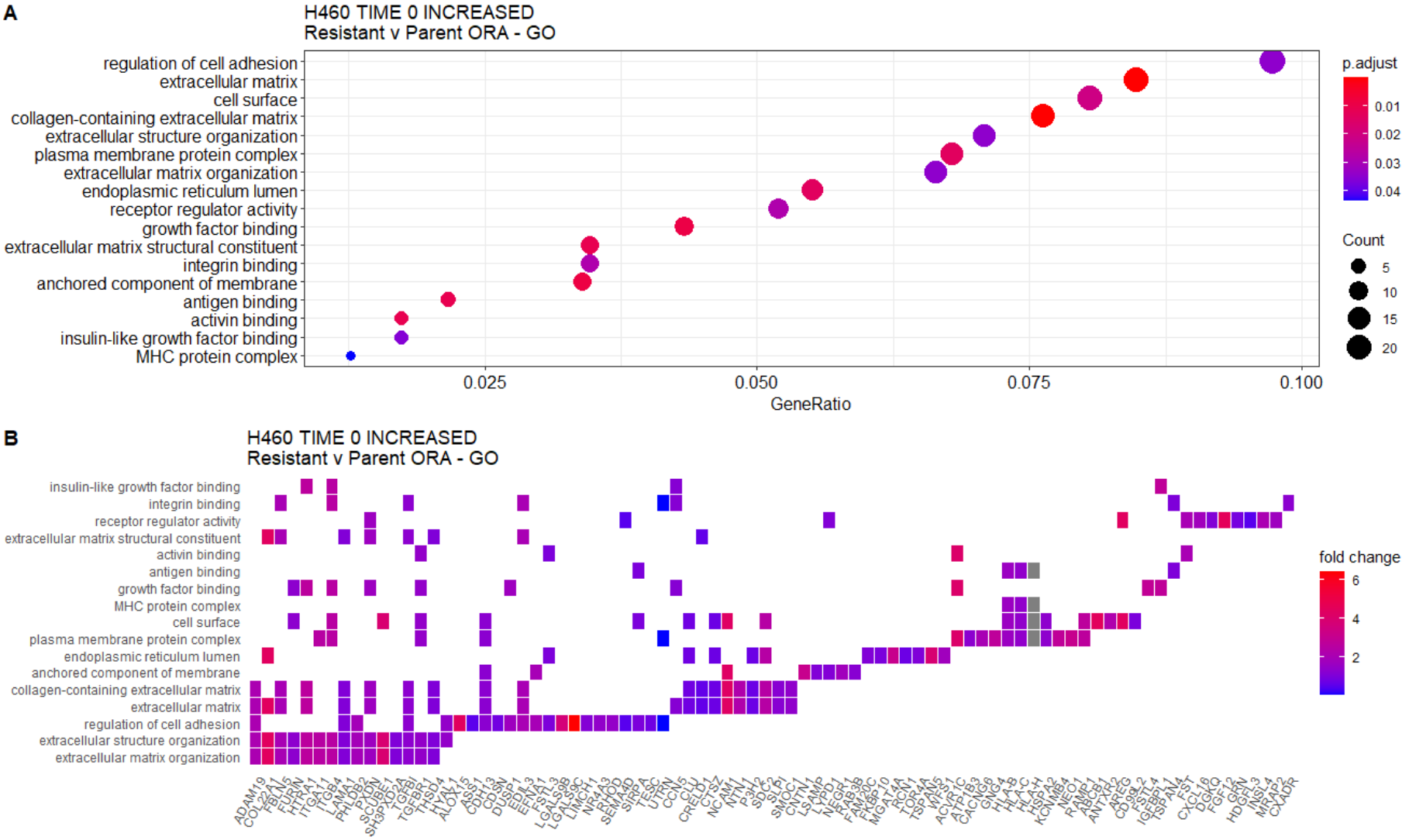


Figure 4.6 ORA Pathway Analysis H460 TIME 0 - Increased Genes Panel A - Dotplot of enriched pathways genes increased in radioresistant (LFC>1), Panel B - Heatplot demonstrating genes contributing to enrichment per pathway, genes displayed on x axis, pathways on y axis

4.16 H460 2 hours post-radiation (TIME 2) Resistant vs Parent

Differential expression analysis between resistant and parental cell lines 2 hours after radiation identified 158 genes significantly (padj<0.05) increased in resistant cell lines (Log Fold Change (LFC) > 1) and 42 genes significantly (padj<0.05) decreased (LFC < -1). The majority of the increased genes, 124 genes (78.5%) were also increased at Time 0, of the decreased genes 20 (47.6%) were also decreased at Time 0. Pathway enrichment was again performed using Over-representation Analysis of Gene Ontology (GO) genesets. Analysis was performed separately on significantly increased and decreased genes with no decreased pathways significant (padj <0.05). Pathways that were significantly enriched in the genes increased in radioresistant samples at 2 hours post radiation included extracellular structure organisation, MHC Protein complex, tissue morphogenesis, regulation of ion transport and positive regulation of cytokine secretion (Figure 4.6).

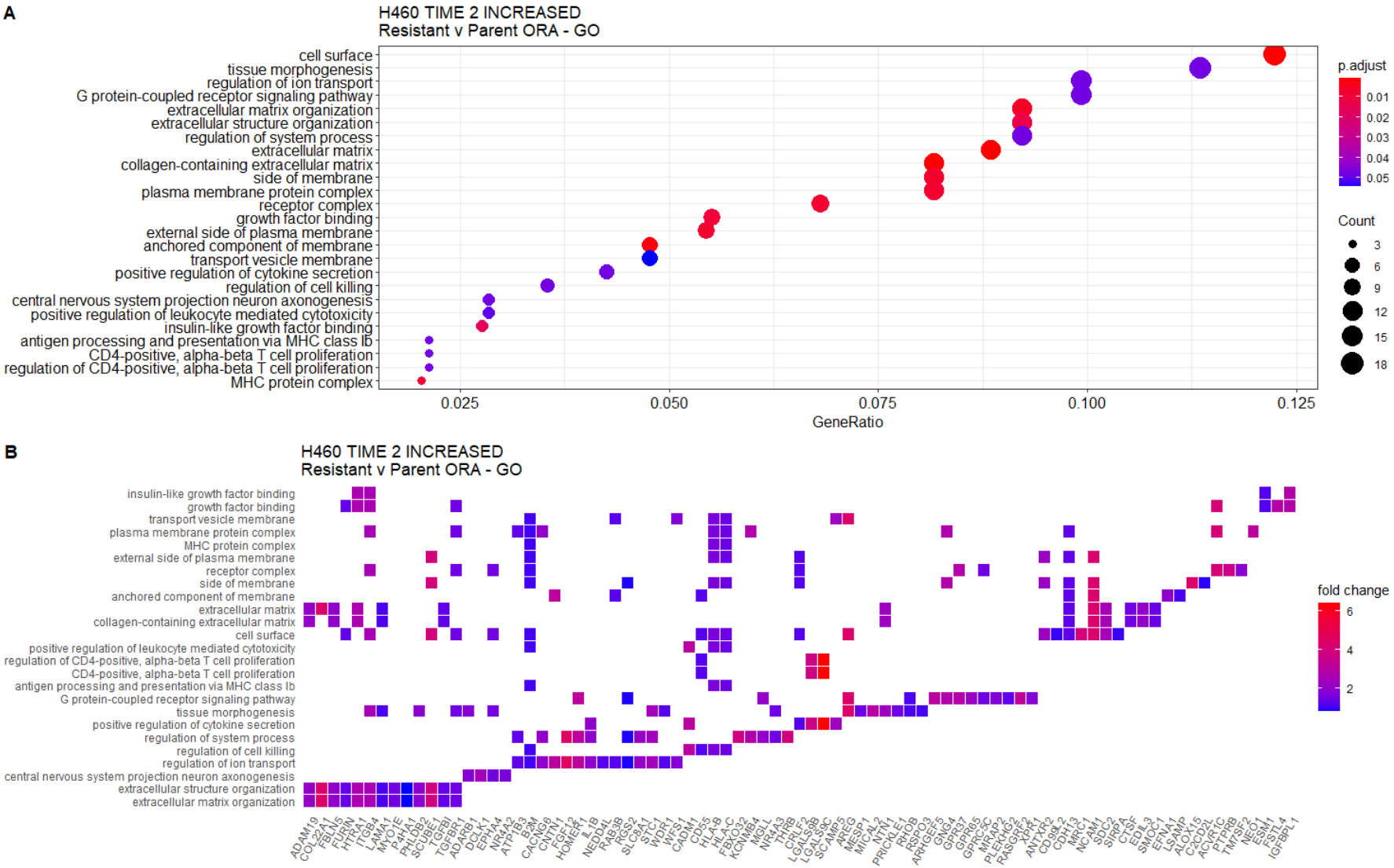


Figure 4.7 ORA Pathway Analysis H460 TIME 2 Increased genes Panel A - Dotplot of enriched pathways genes increased in radioresistant (LFC>1),Panel B - Heatplot demonstrating genes contributing to enrichment per pathway(genes x axis, pathways on y axis, foldchange colour)

Pathways can again be seen to be enriched due to expression of distinct sets of genes related to each pathway, with some areas of overlap and multiples of similar pathways e.g. extracellular matrix, regulation of CD4+ve, alpha beta T cell proliferation.

4.17 H460 6 hours post-radiation (TIME 6) Resistant vs Parent

Differential expression analysis between resistant and parental cell lines 6 hours after radiation identified 144 genes significantly (padj<0.05) increased in resistant cell lines (LFC > 1) and 42 genes significantly (padj<0.05) decreased (LFC < -1).

The majority of the significantly increased genes, 95 genes (65.9%) were also increased at Time 2, of the significantly decreased genes 18 (40.9%) were also significantly decreased at Time 2. Pathways analysis was performed separately on increased and decreased genes with 2 pathways significantly (padj <0.05) enriched in the decreased genes polymeric cytoskeletal fibre and intermediate filament (Figure 4.7 Panel A). Both pathways relate to cytoskeletal structural components and there is complete overlap of the component genes.

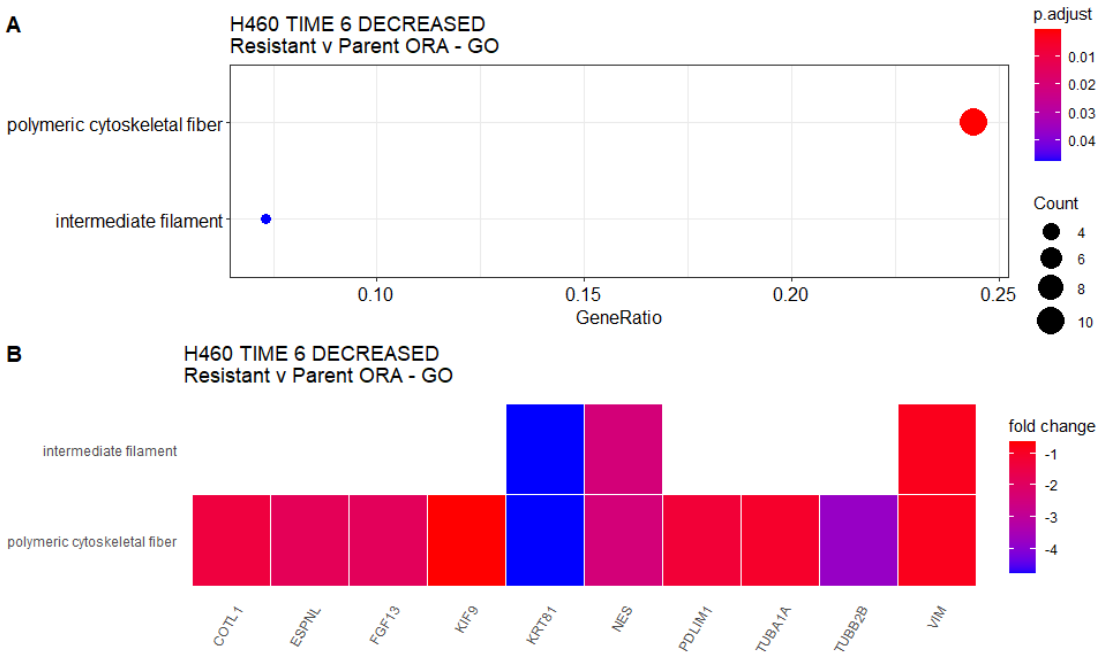


Figure 4.8 ORA Pathway Analysis H460 TIME 6 - Decreased genes

Panel A - Dotplot of enriched pathways genes decreased in radioresistant at 6 hours post radiation (LFC> -1, adjusted p value < 0.05), Panel B - Heatplot demonstrating genes contributing to enrichment per pathway (genes on x-axis, pathways y-axis)

Pathways that were significantly enriched in the genes increased in radioresistant samples at 6 hours post radiation included response to BMP, cellular response to growth factor stimulus, and triglyceride catabolic process (Figure 4.8 Panel A)

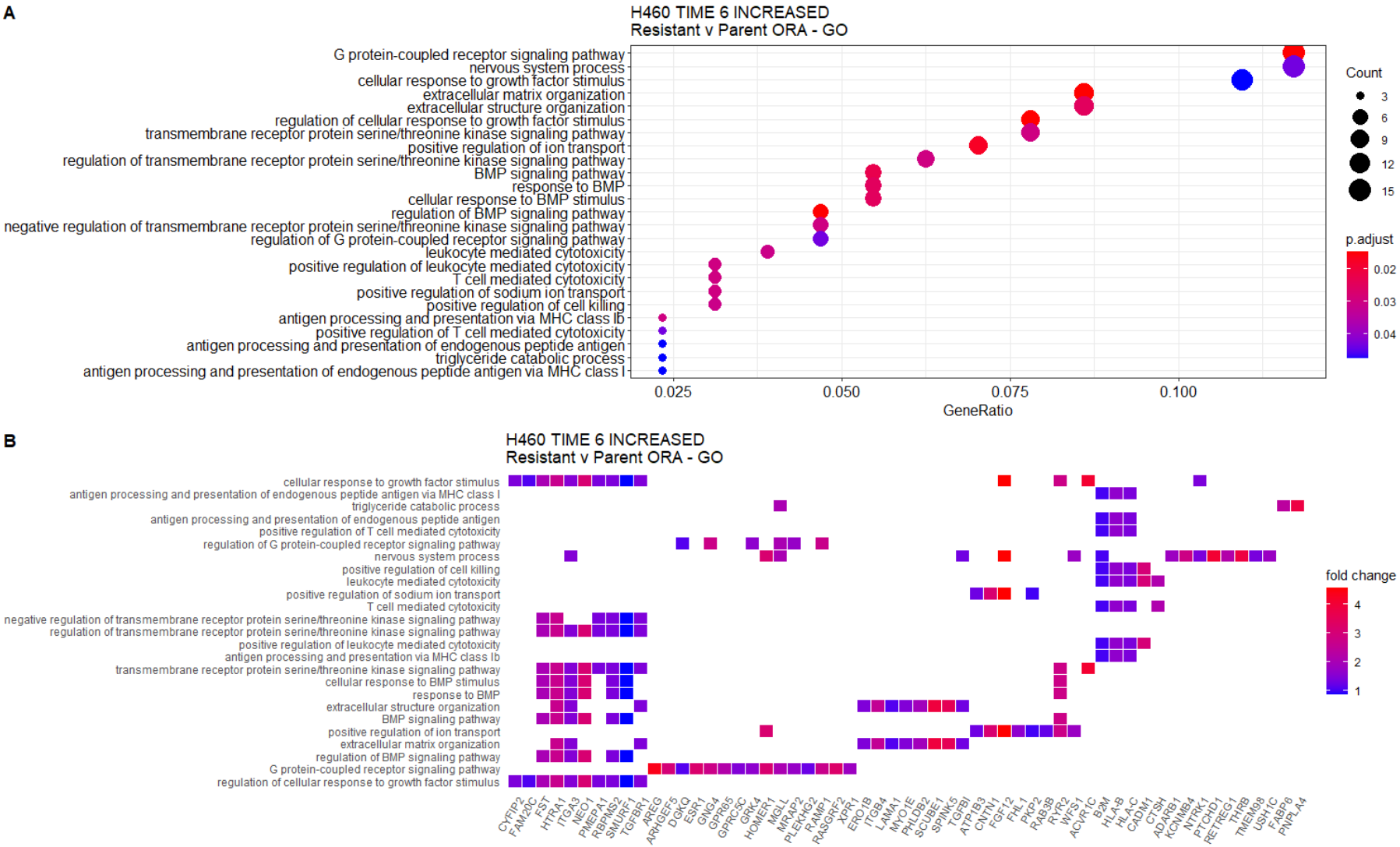


Figure 4.9 ORA Pathway Analysis H460 TIME 6 - Increased genes

Panel A - Dotplot of enriched pathways genes increased in radioresistant at 6 hours post radiation (LFC> 1, p adj < 0.05), Panel B - Heatplot demonstrating genes contributing to enrichment per pathway (genes on x-axis, pathways y-axis)

4.18 H460 12 hours post-radiation (TIME 12) Resistant vs Parent

Differential expression analysis between resistant and parental cell lines 12 hours after radiation identified 127 genes significantly (padj<0.05) increased in resistant cell lines (LFC > 1) and 50 genes significantly (padj<0.05) decreased (LFC < -1). The majority of the significantly increased genes, 77 genes (60.6%) were also increased at Time 6, of the significantly decreased genes 22 (44%) were also significantly decreased at Time 6. Pathways analysis identified enrichment of genes significantly decreased in resistant samples in pathways associated with extracellular matrix and intrinsic components of plasma membrane (Figure 4.9 Panel A). Significant overlap of genes between pathways is evident.

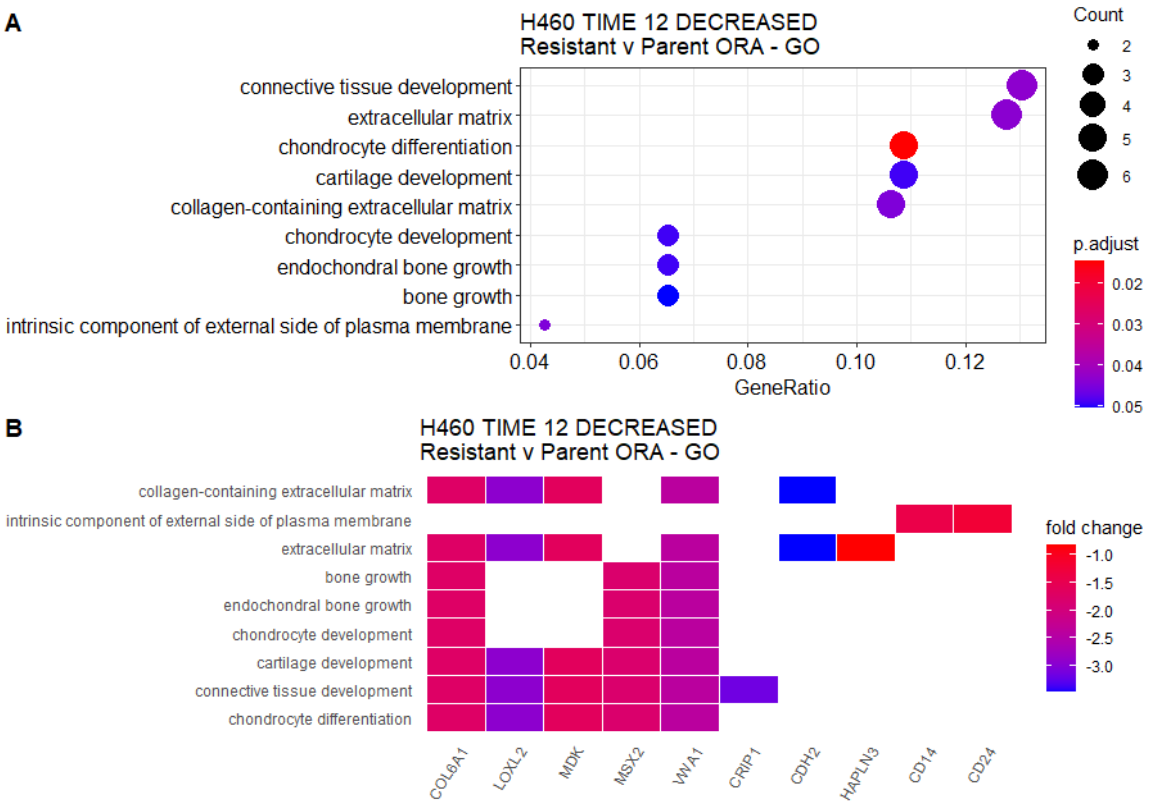


Figure 4.10 ORA Pathway Analysis H460 TIME 12 - Decreased genes
Panel A - Dotplot of enriched pathways genes decreased in radioresistant at 12 hours post radiation (LFC< -1, pad < 0.05), Panel B - Heatplot demonstrating genes contributing to enrichment per pathway (genes on x-axis, pathways y-axis)

Pathways analysis identified enrichment of genes significantly increased in resistant samples in pathways associated with BMP signalling, regulation of actin filament, and activin binding (Figure 4.10 Panel A

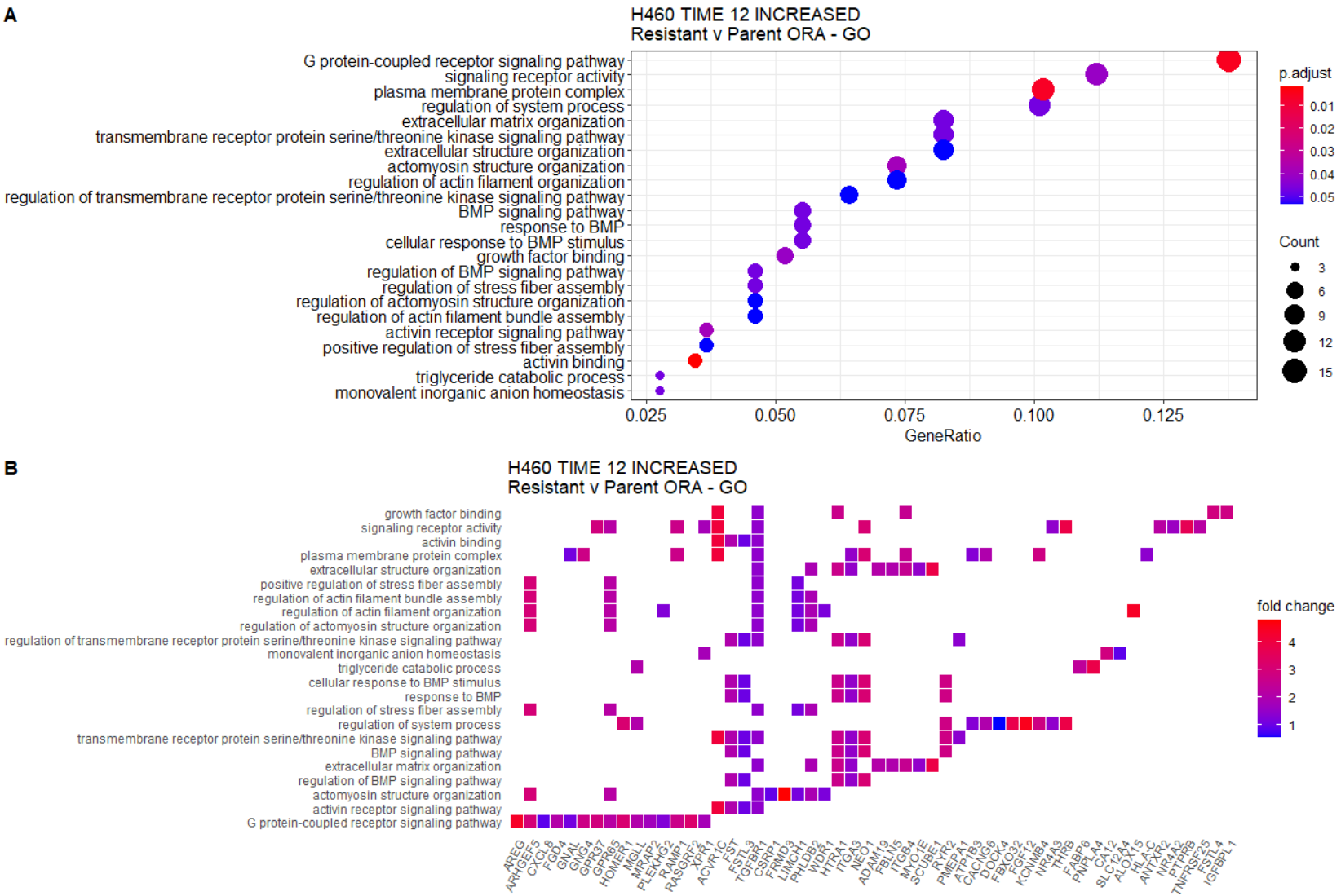


Figure 4.11 ORA Pathway Analysis H460 TIME 12 - Increased genes

Panel A - Dotplot of enriched pathways genes increased in radioresistant at 12 hours post radiation (LFC> 1, pad <0.05), Panel B - Heatplot demonstrating genes contributing to enrichment per pathway (genes on x-axis, pathways y-axis)

4.19 H460 24 hours post-radiation (TIME 24) Resistant vs Parent

Differential expression analysis between resistant and parental cell lines 24 hours after radiation identified 191 genes significantly (padj<0.05) increased in resistant cell lines (LFC > 1) and 93 genes significantly (padj<0.05) decreased (LFC < -1). Of the significantly increased genes, 81 genes (42.4%) were also increased at Time 12, of the significantly decreased genes 30 (32.2%) were also significantly decreased at Time 12.
Pathway analysis identified that the 93 genes significantly decreased in resistant samples were over-represented in pathways associated with blood vessel development, epithelial migration and regulation of Wnt signalling with significant overlap of genes for multiple pathways (Figure 4.11 Panel A,B).

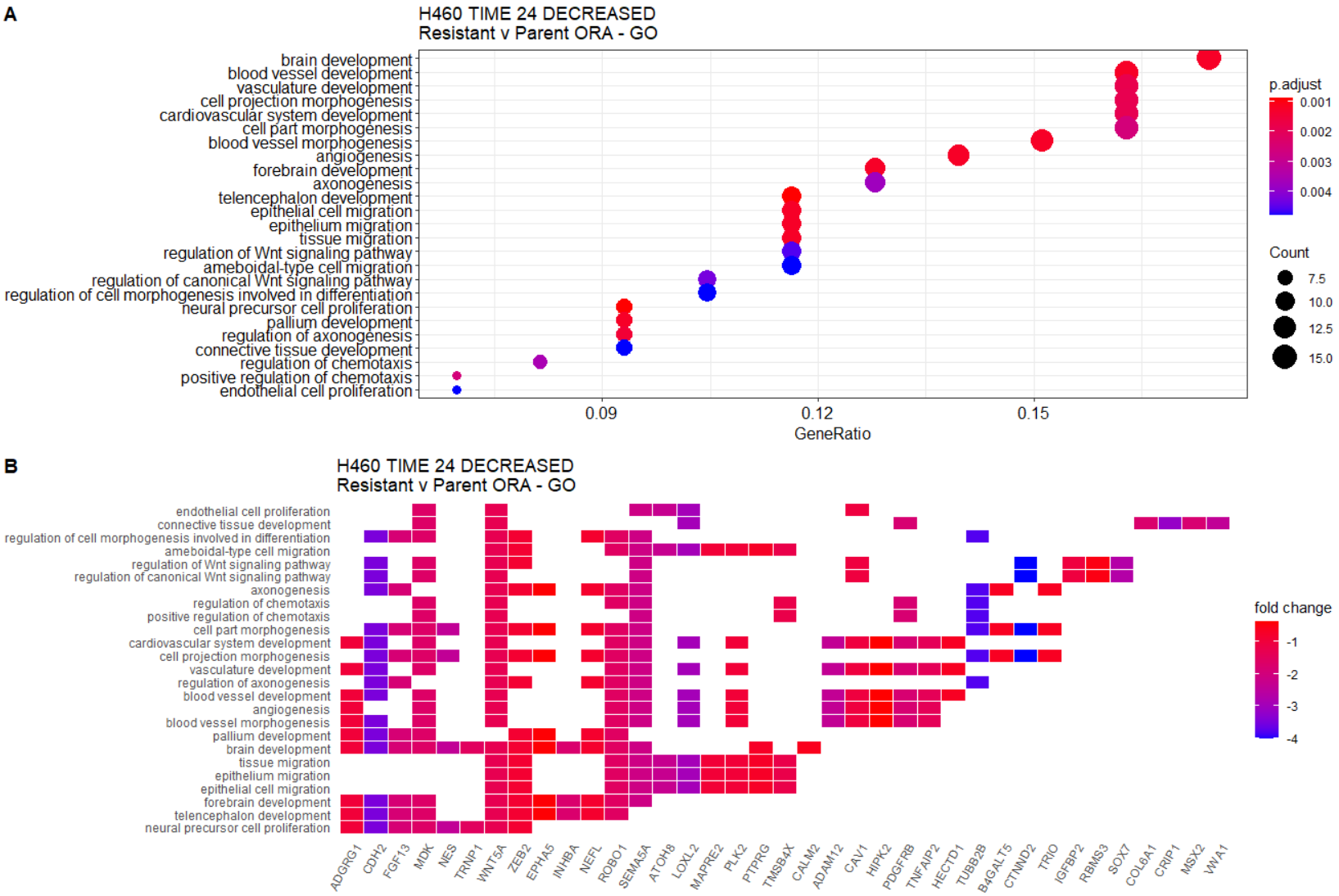


Figure 4.12 ORA Pathway Analysis H460 TIME 24 - Decreased genes Panel A - Dotplot of enriched pathways genes decreased in radioresistant at 24 hours post radiation (LFC< -1, padj <0.05), Panel B - Heatplot demonstrating genes(x axis)contributing to enrichment per pathway (y axis)

The 191 significantly increased genes enriched for pathways associated with G-protein receptor signalling pathways, plasma membrane complex and ion binding and transmembrane transporter activity (Figure 4.12 Panel A). Approximately 6 distinct groups of non-overlapping genes representing multiple pathways (Panel B) representing G-protein receptor signalling pathways, plasma membrane protein complex, receptor regulator activity, ion channel binding and ion transmembrane transporter activity.

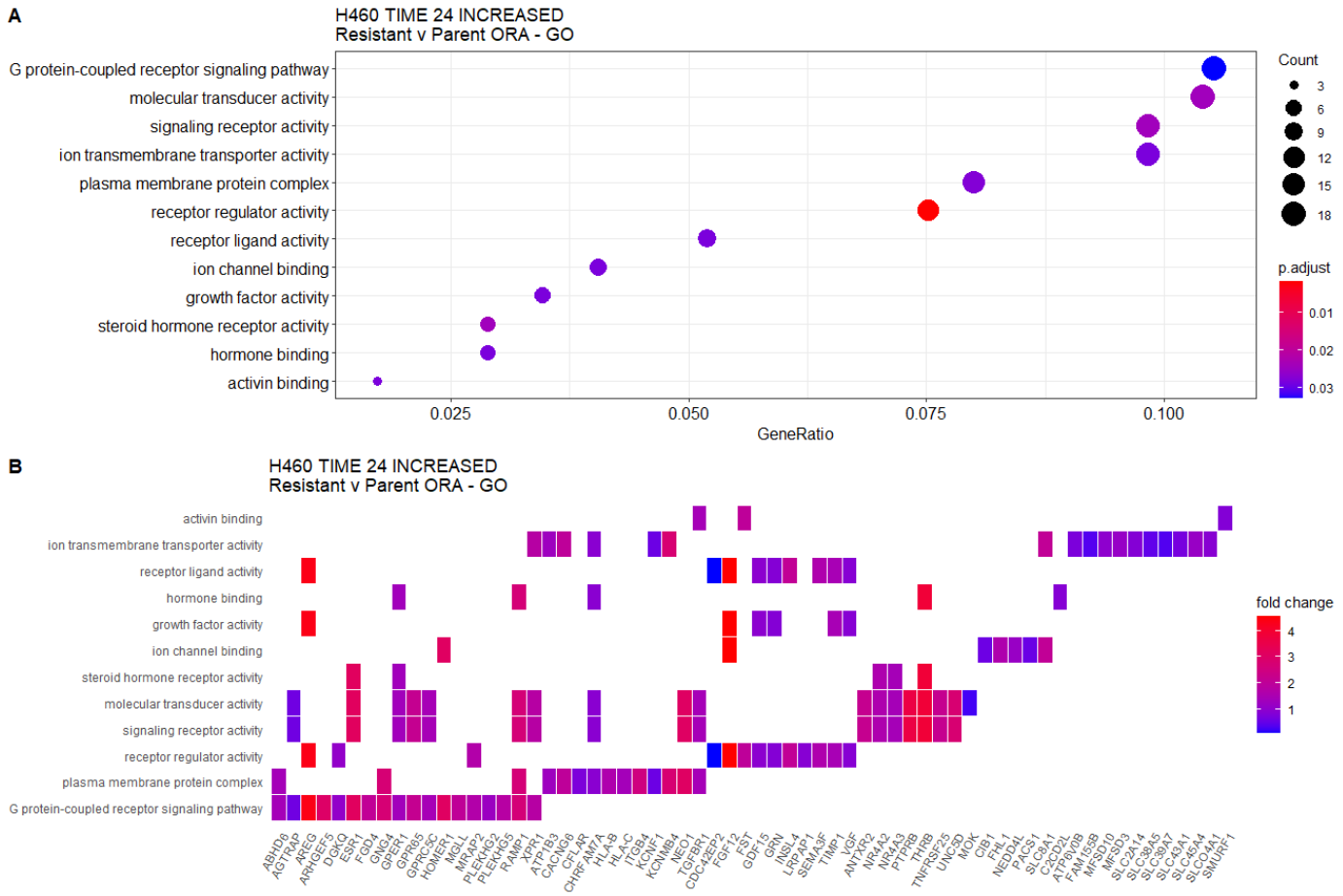


Figure 4.13 ORA Pathway Analysis H460 TIME 24 - Increased genes

Panel A - Dotplot of enriched pathways genes increased in radioresistant at 24 hours post radiation (LFC> 1, padj < 0.05), Panel B - Heatplot demonstrating genes (x axis) contributing to enrichment per pathway (y axis).

In order to further investigate the number of genes that were differentially expressed at each timepoint a further analysis was performed to compute the overlaps between differential expression in resistant and parent samples at each timepoint. The analysis identified 52

genes that were consistently increased in Resistant vs Parent at all timepoints and 11 genes consistently decreased (Figure 4.13). The maximum number of uniquely increased genes was seen at baseline Time 0 and Time 24 with 98 and 68 genes respectively. The maximum number of uniquely decreased genes was seen at Time 2 and Time 24 with 25 and 50 genes respectively (Figure 4.13).

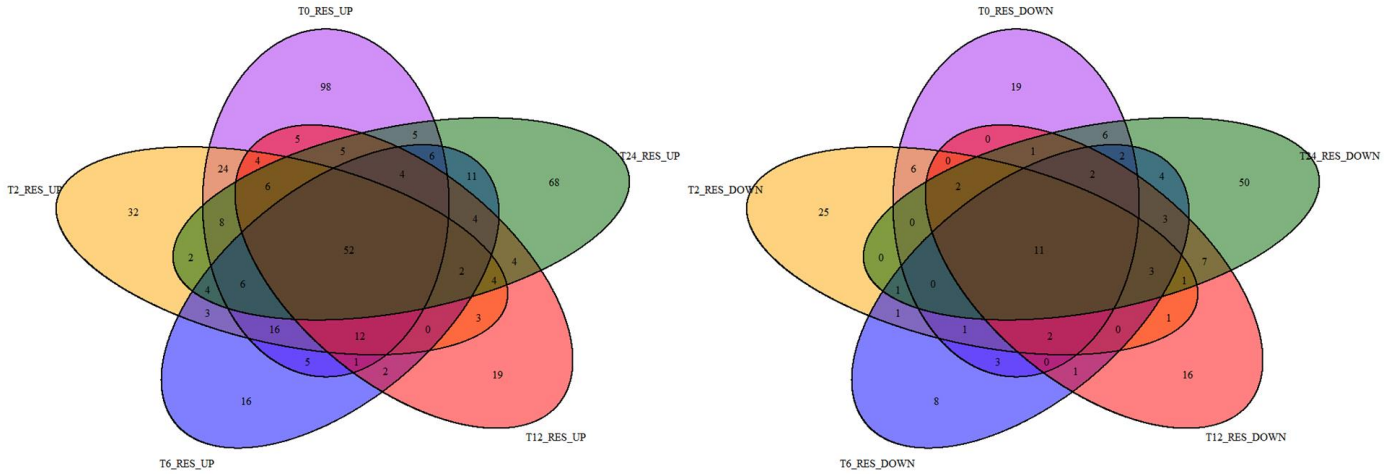


Figure 4.14 Venn Diagram of Differentially Expressed genes at each timepoint
Left - Overlaps of Differentially expressed genes Increased in Radioresistant at each timepoint, Right - Overlaps of Differentially expressed genes Increased in Radioresistant at each timepoint

The 52 significantly increased genes enriched for pathways associated with G-protein receptor signalling pathways and catabolic pathways (Figure 4.15 Panel A), with distinct groups of non-overlapping genes representing multiple pathways (Figure 4.15 Panel B). The 11 decreased genes just met significance (padj <0.05) for intermediate filament pathway.

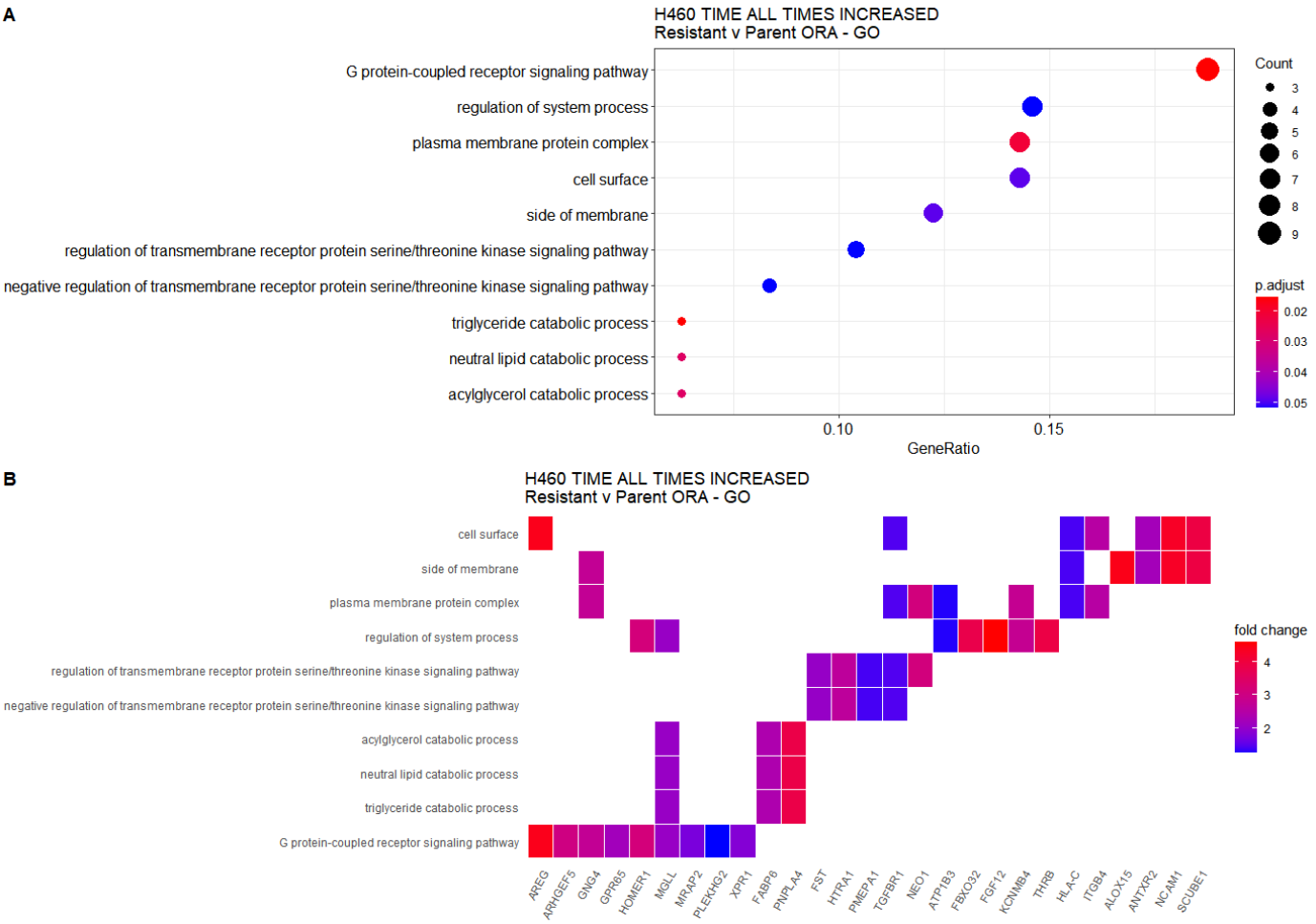


Figure 4.15 ORA Pathway Analysis of commonly Increased genes at each timepoint

Panel A - Dotplot of enriched pathways genes increased in radioresistant at 24 hours post radiation (LFC> 1, padj <0.05), Panel B - Heatplot demonstrating genes (x axis) contributing to enrichment per pathway (y axis).

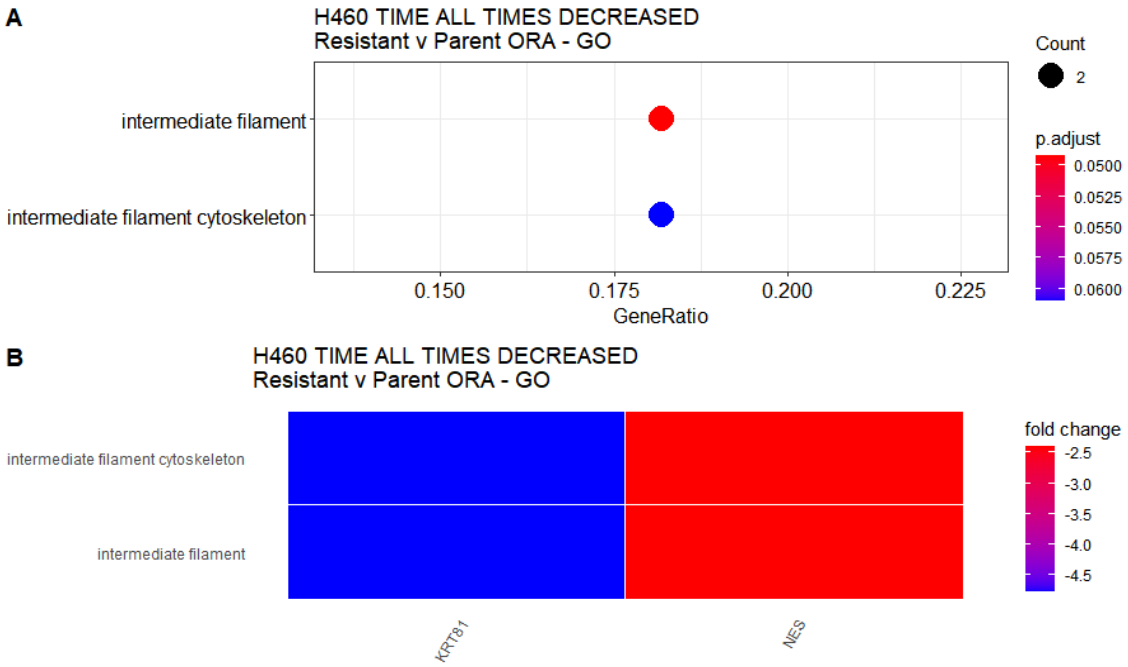


Figure 4.16 ORA Pathway Analysis of commonly Decreased genes at each timepoint
Panel A - Dotplot of enriched pathways genes decreased in radioresistant at 24 hours post radiation (LFC> 1, padj <0.05), Panel B - Heatplot demonstrating genes (x-axis) contributing to enrichment (y-axis) per pathway.

4.20 H460 Uniquely Increased games in Radioresistant cells after 24 hours

Given the previous findings of most differentially expressed genes being at base-line (pre-radiation) and 24 hours post-radiation, a further differential expression analysis was conducted to compare the differentially expressed genes between resistant cell lines 24 hours after radiation and at baseline, with the differentially expressed genes of the parental samples 24 hours after radiation and at baseline i.e. (Time 24 Resistant vs Time 0 Resistant) compared to Time 24 Parent vs Time 0 Parent, p adj < 0.05).

This demonstrates 628 genes that are significantly increased in expression in the resistant cells 24 hours after radiation that are not significantly increased in the Parental cells 24 hours after radiation. The overlaps with baseline differences between Parental and Radioresistant cells (Time 0) were also computed (Appendix 2.2).

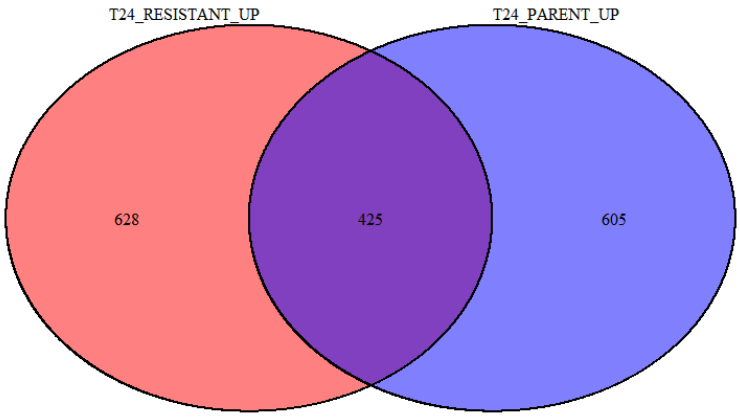


Figure 4.17 Overlap of Increased genes Resistant Time 24 vs Time 0, with Parent Time 24 vs Time 0.

Venn diagram illustrating overlap of genes that are significantly (padj < 0.05) differentially expressed genes 24 hours following radiation, between radioresistant (red) and radiosensitive (blue) cell-lines.

To further investigate whether this group of genes, that are exclusively significantly increased in radioresistant cells 24 hours after radiation, pathway enrichment analysis was performed. This identified pathways associated with response to Endoplasmic reticulum stress, Ras signalling transduction, macroautophagy and autophagy as being significantly enriched, suggesting that these pathways are activated in the radioresistant cells but not the parental cells twenty four hours after repeat irradiation.

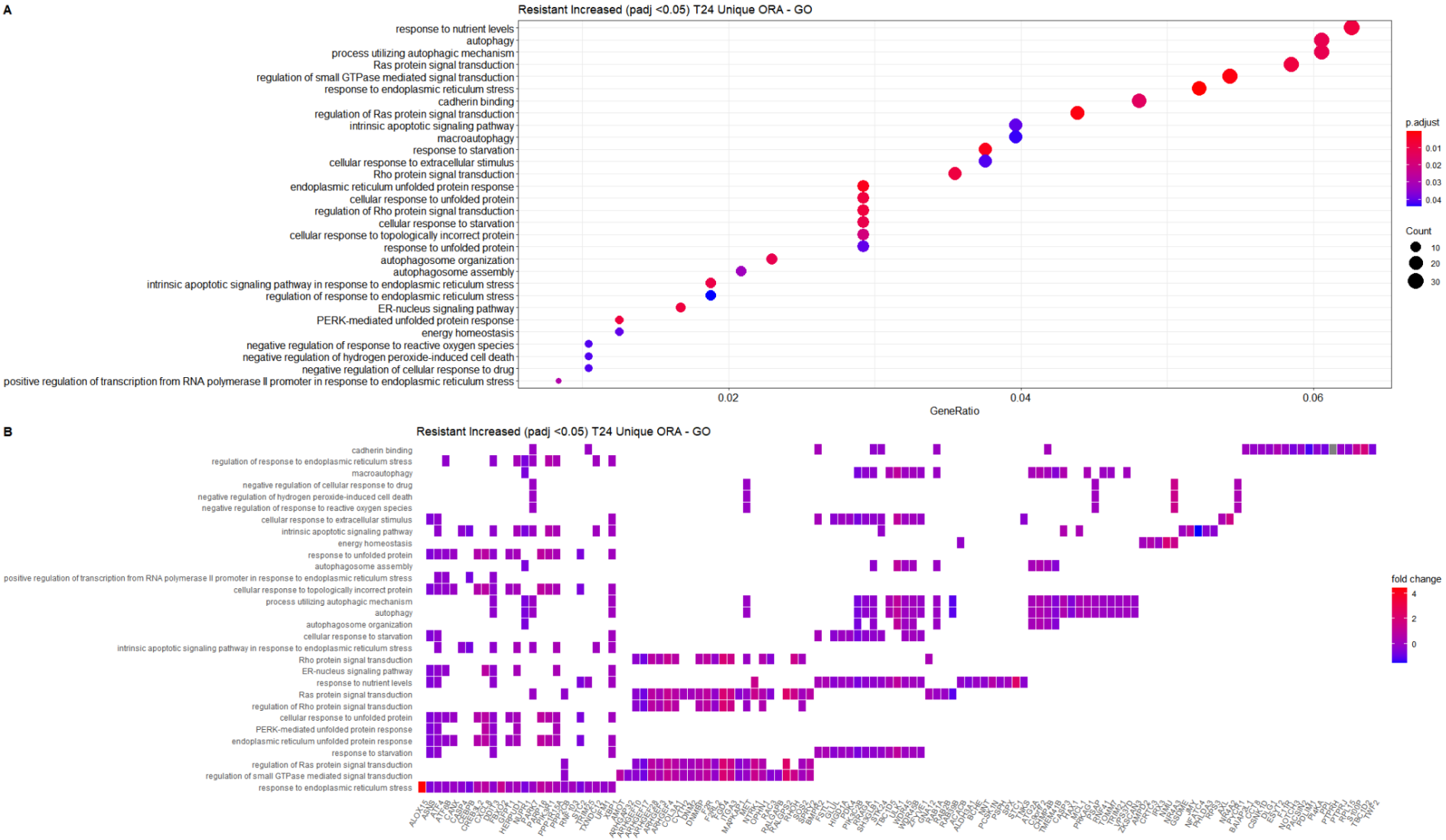
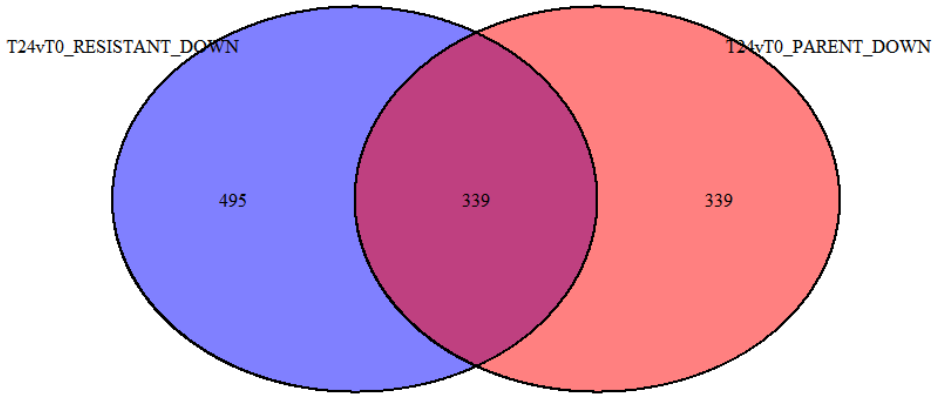


Figure 4.18 ORA Pathway Analysis of genes Increased in only Radioresistant cells after 24 hours Panel A - Dotplot of enriched pathways genes increased in only radioresistant at 24 hours post radiation (LFC> 1, padj <0.05), Panel B - Heatplot demonstrating genes (x axis) contributing to enrichment per pathway (y axis).

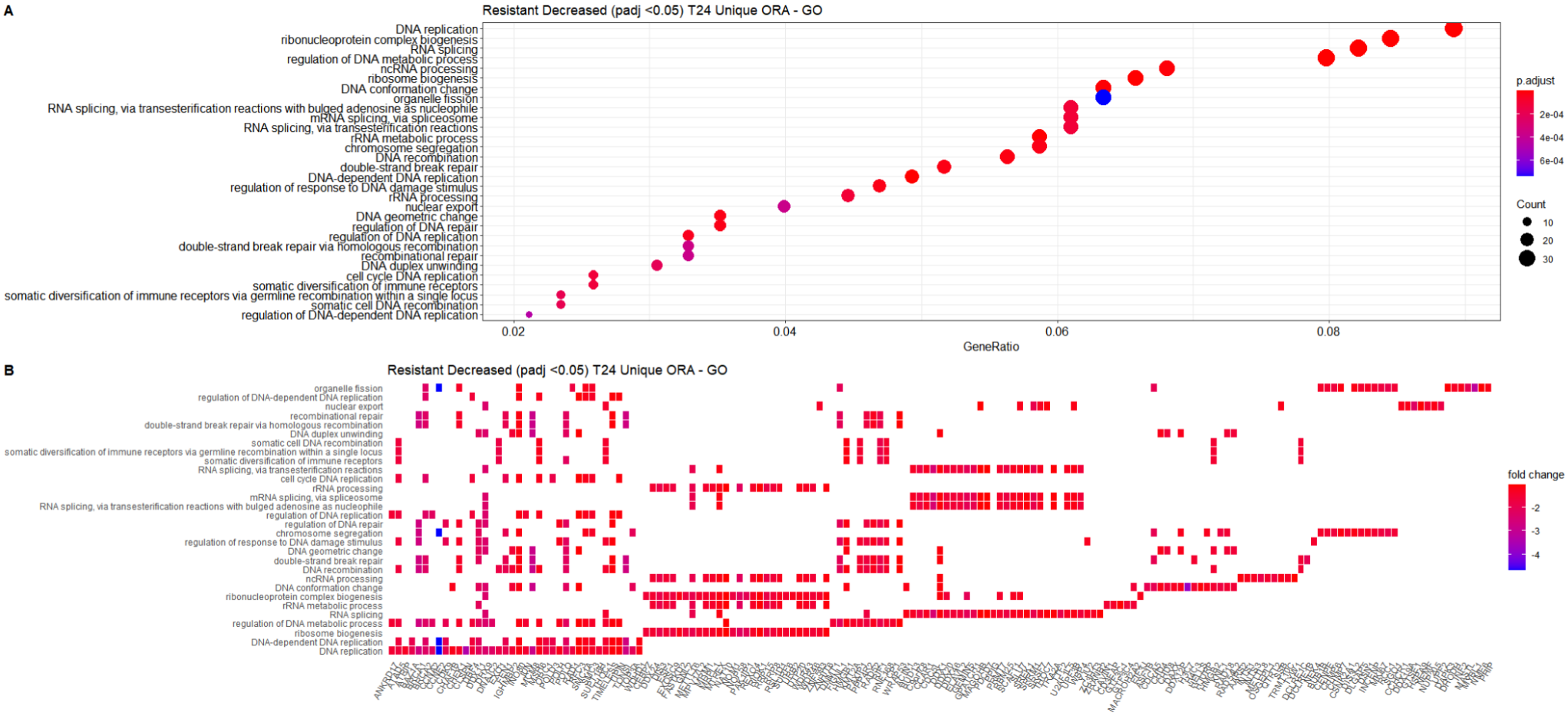
The same analysis demonstrated 339 genes that are significantly decreased in expression in the resistant cells 24 hours after radiation compared to baseline that are not significantly decreased in the Parental cells 24 hours after radiation compared to baseline.

Figure 4.19 Overlap of Decreased genes Resistant Time 24 vs Time 0, with Parent Time 24 vs Time 0.



Pathway analysis of these 339 genes decreased at 24 hours after radiation identifies pathways that are significantly enriched for including DNA replication, ribosome biogenesis, double strand break repair via homologous recombination and RNA splicing.

Figure 4.20 ORA Pathway Analysis of genes Decreased in only Radioresistant cells after 24 hours



Panel A - Dotplot of enriched pathways genes increased in only radioresistant at 24 hours post radiation (LFC> 1), Panel B - Heatplot demonstrating genes contributing to enrichment per pathway. Panel A - Dotplot of enriched pathways genes increased in only radioresistant at 24 hours post radiation (LFC> 1, padj <0.05), Panel B - Heatplot demonstrating genes (x axis) contributing to enrichment per pathway (y axis).

This suggests that there is less expression of genes involved in these pathways in radioresistant cells compared to parental cells twenty four hours compared to baseline after repeat irradiation.

4.21 H460 Cell Line Analysis utilising a time interaction model

Based on the previous differential expression findings of differences in gene expression between radioresistant cell lines and parental cell lines not only across time but also in a time dependent manner, where expression changes over time are different between parent and resistant samples (Figures 4.17, 4.19) a further differential expression analysis was conducted using a term to account for interaction over time (Methods). This enabled the utilisation of the full dataset to explore changes over time between parental and resistant cell lines.

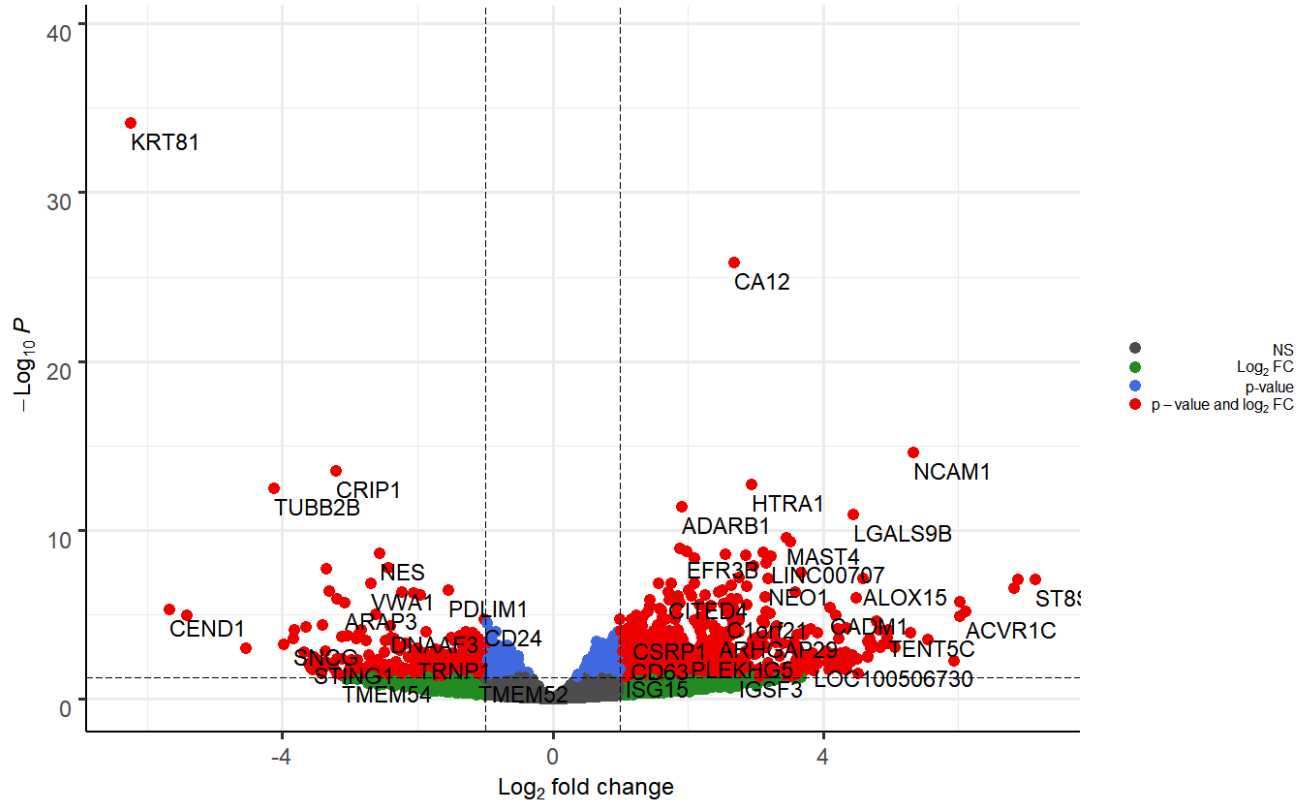


Figure 4.21 Volcano Plot of differential expression between Radioresistant and Parent cell lines (Time interaction model)
Log fold change y-axis, Log p value y-axis, red = log fold change < -1 to > 1 and padj <0.05

Differential expression analysis using a time interaction model, identified 356 significantly differentially expressed genes (padj <0.05) This was further refined to 269 genes increased (>LFC 1) and 60 genes decreased (<LFC -1), of which 126 of the genes increased in radioresistant samples, were significant at padj <0.01 and 26 decreased (Figure 4.21).

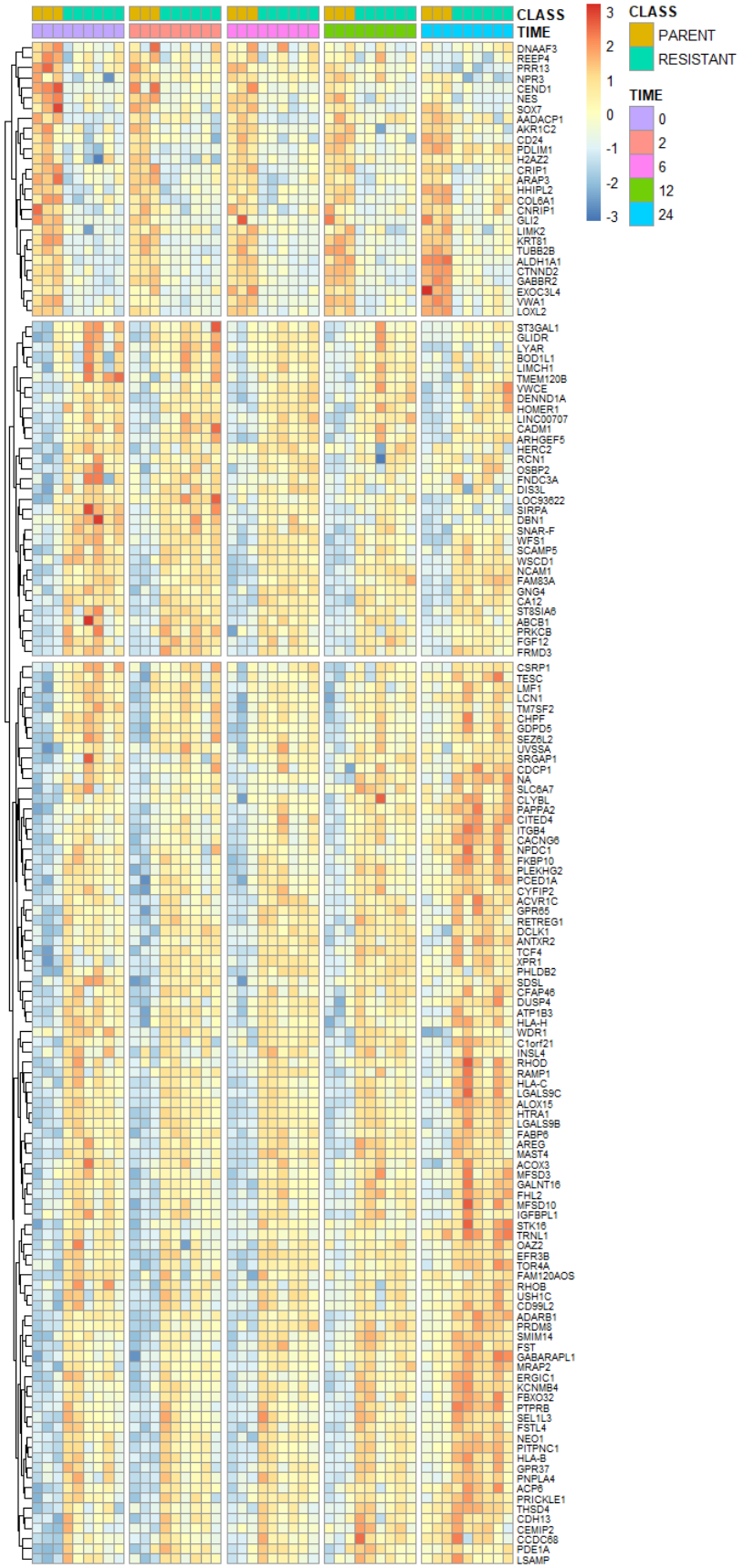


Figure 4.22 Heatmap of significantly (padj <0.01) differentially expressed genes between Radioresistant and Parent cell lines (Time interaction model)
Genes are displayed on y-axis, cell lines ordered by time displayed on x-axis

The heatmap Figure 4.22 demonstrates the significantly differentially expressed genes between radioresistant and parent cell lines and is displayed in time order to allow visualisation of the differing expression levels at different timepoints. Gene expression is scaled across rows, and genes on the horizontal axis are clustered using hierarchical clustering and partitioned based on the top 3 clusters. The top third portion of the heatmap demonstrates genes that are decreased in radioresistant samples across all times. The initial cluster of 7 genes are demonstrated to have increased expression at Time 0 with minimal expression at Time 24. The middle cluster of 11 genes demonstrate relatively constant differential expression and the lower 9 genes demonstrate increased expression at Time 24 in the Parental cell line.

The middle portion of the heatmap demonstrates genes that are increased in radioresistant vs parental samples, it can be noted that these genes have their highest expression at early timepoints Time 0 and Time 2. The lowest portion demonstrates genes that are increased in radioresistant vs parental samples however they have maximal expression at the later timepoint Time 24. This heatmap demonstrates the dynamic nature of gene expression changes over time.

Over representation pathway analysis yielded similar results to those previously presented with increased pathways associated with extracellular matrix, G-protein receptor signalling pathways, plasma membrane protein complex. (Appendix 2.3, 2.4)

4.22 Geneset Enrichment Analysis H460 Resistant vs Parent Time Interaction

Geneset Enrichment Analysis (GSEA) is a computational method that determines whether an a priori defined set of genes shows statistically significant, concordant differences between two biological states. Figure 4.23 demonstrates the top 25 significantly (padj <0.05)

differentially expressed Reactome pathways between Radioresistant and Parent cell lines (Time interaction model, divided into those that are suppressed in the Resistant samples and those that are activated. Pathways that are significantly suppressed include DNA replication, Mitotic Spindle Checkpoint and G2/M transition.

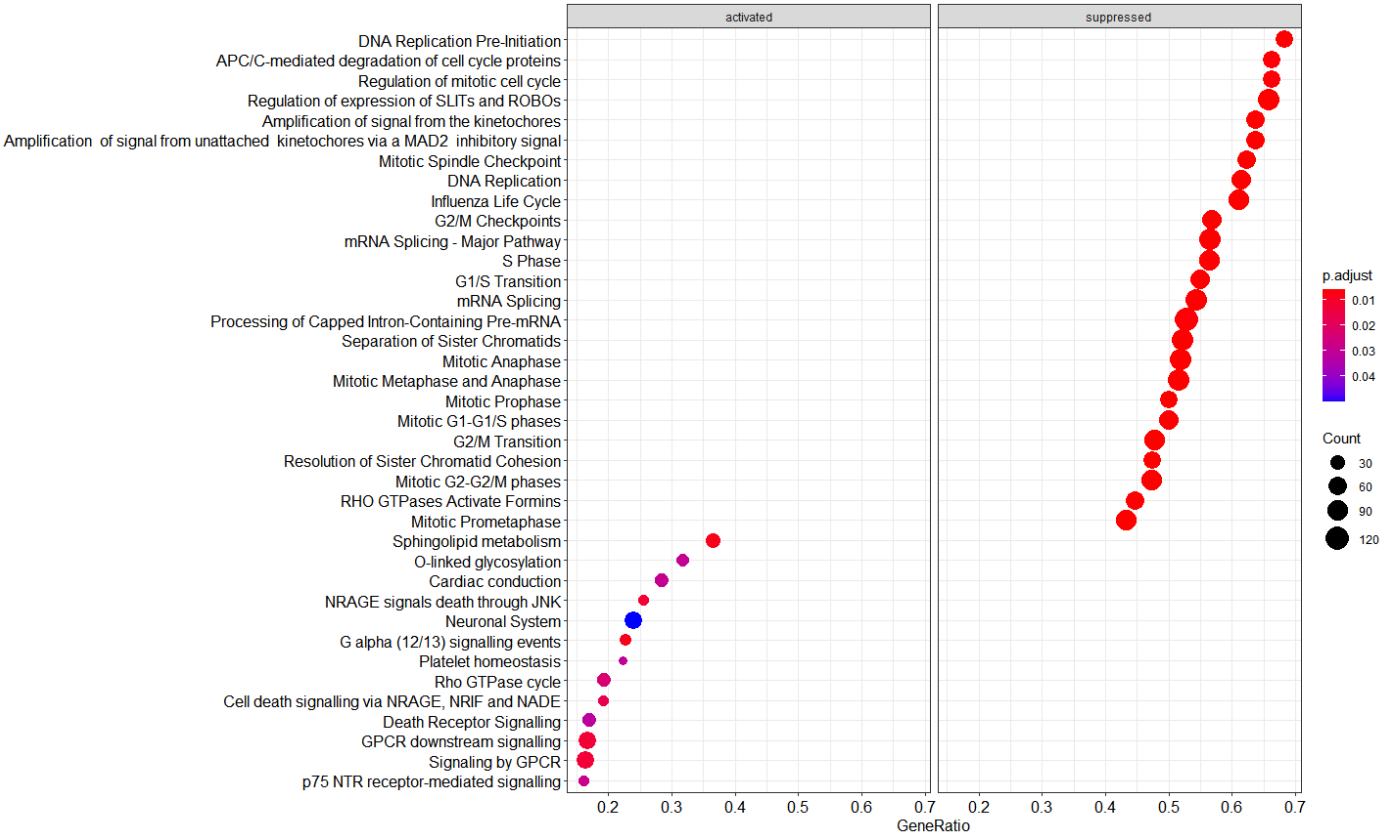


Figure 4.23 Dotplot of significantly (padj <0.05) differentially expressed Reactome pathways by GSEA between Radioresistant and Parent cell lines (Time interaction model)
Pathways are demonstrated on the x-axis, left panel demonstrates activated pathways, right panel demonstrates suppressed pathways, colours demonstrate adjusted p value, size shows number of genes

Pathways that are activated include Death Receptor Signalling, Sphingolipid metabolism, Cell death signalling via NRAGE, NRIF and NADE and p75 NTR receptor-mediated signalling.

Figure 4.24 demonstrates the top 25 significantly (padj <0.05) differentially expressed GO pathways between Radioresistant and Parent cell lines (Time interaction model), divided into those that are suppressed in the Resistant samples and those that are activated. Pathways

that are significantly suppressed include Spliceosome, mitotic spindle organisation, mitotic nuclear division and chromatin assembly. Pathways that are significantly activated include the intrinsic component of the membrane, endoplasmic reticulum and golgi apparatus part. The two GSEA analyses utilised different curated genesets Reactome and GO, however both display similarity in the suppressed pathways in radioresistant cells of pathways related to mitosis and mitotic cell division.

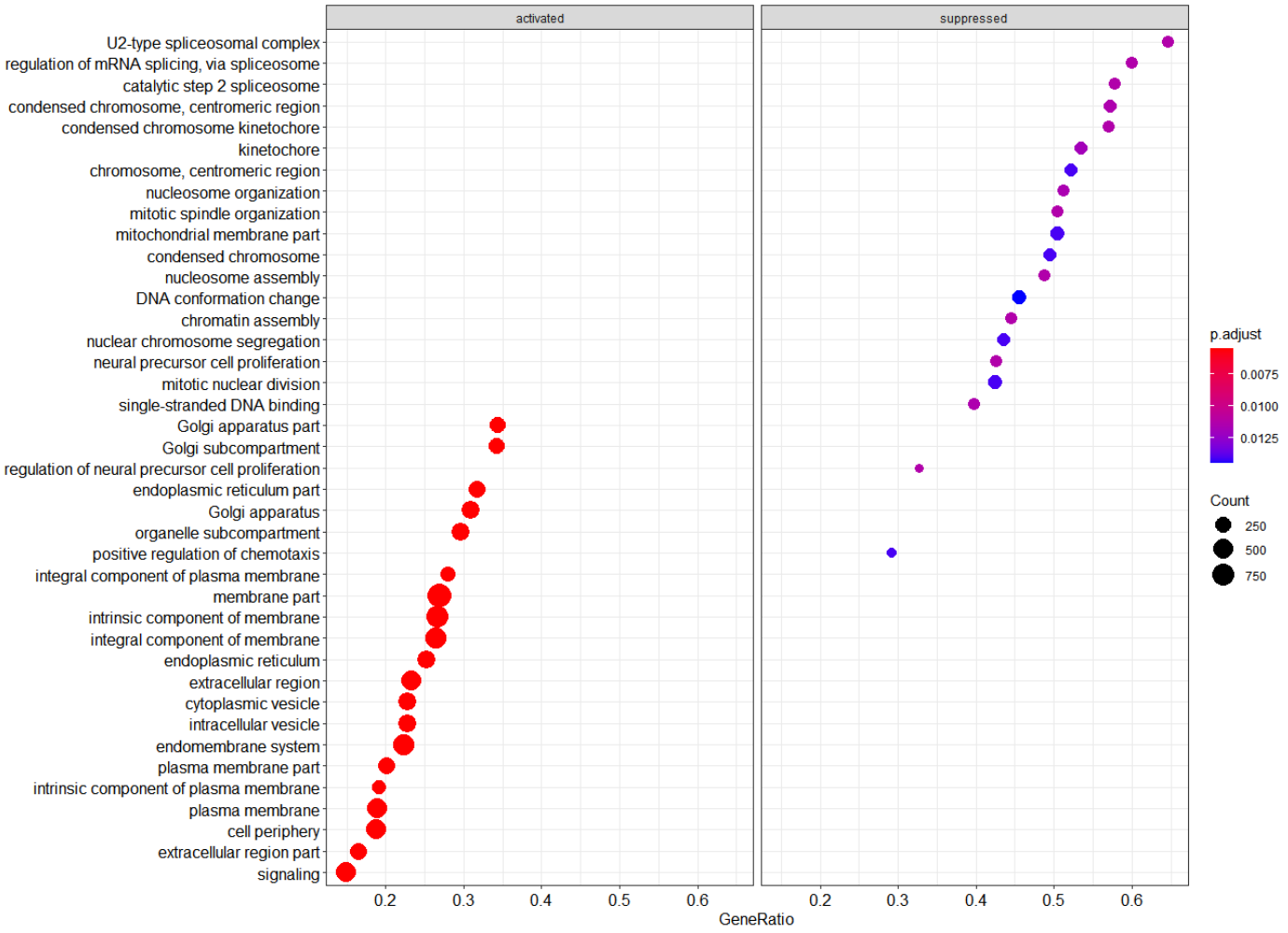


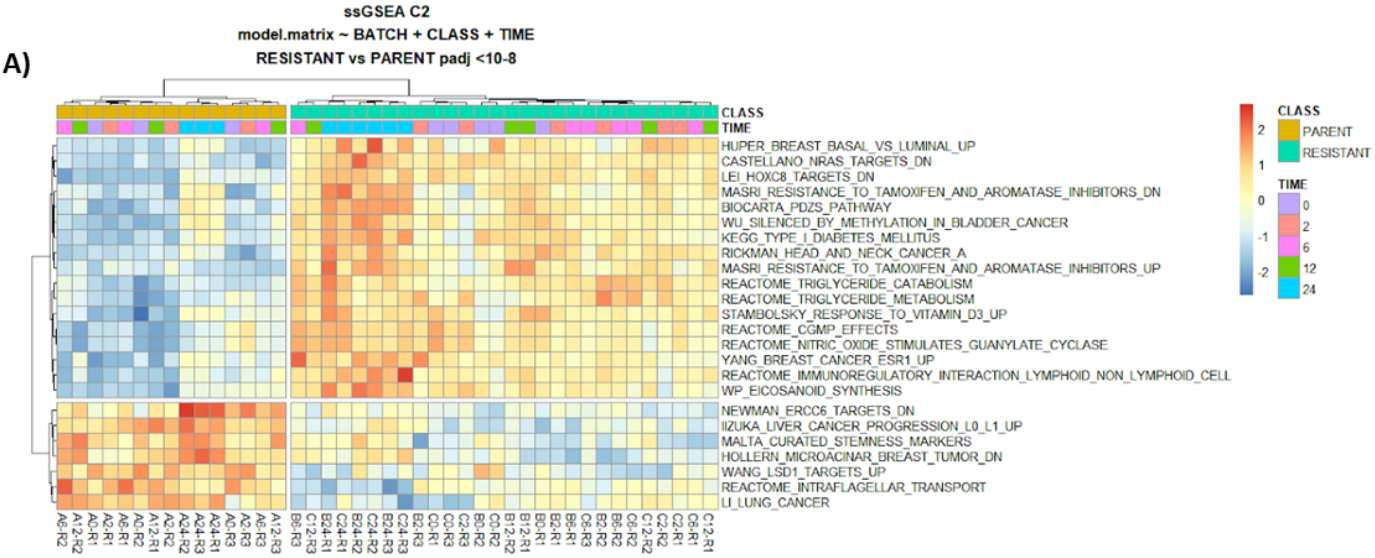
Figure 4.24 Dotplot of significantly (padj <0.05) differentially expressed GO pathways by GSEA between Radioresistant and Parent cell lines (Time interaction model)
Pathways are demonstrated on the x-axis, left panel demonstrates activated pathways, right panel demonstrates suppressed pathways, colours demonstrate adjusted p value, size shows number of genes

In order to further investigate pathways that are differentially expressed between Radioresistant and Parental cell lines, single sample gene set enrichment analysis was performed, using the C2 Curated Genesets from MSigDB. This identified signatures

associated with NRAS targets down, HOXC8 targets down, as well as Immuno-Regulatory elements between a lymphoid and non-lymphoid cell, triglyceride metabolism and eicosanoid synthesis as increased in the radioresistant samples, silenced by methylation in bladder cancer (Figure 4.25).Pathways associated with Stemness and ERCC6 targets down were increased in the parental cell lines compared to resistant cell lines (Figure 4.25).

The process was repeated with Hallmarks genesets which revealed significantly differentially expressed pathways of Angiogenesis, Estrogen response, UV Response and Hypoxia increased and Notch Signalling, IL6-Jak-Stat3 signalling and G2M Checkpoint reduced (Appendix 2.6)

A)



B)

Pathway	logFC	AveExpr	t	P.Value	adj.P.Val	B
CASTELLANO_NRAS_TARGETS_DN	0.08642	0.063831	20.36852	2.07E-22	1.01E-18	40.87483
LEI_HOXC8_TARGETS_DN	0.06307	0.132561	15.82286	1.40E-18	3.42E-15	31.94367
RICKMAN_HEAD_AND_NECK_CANCER_A	0.029908	-0.05569	13.1221	6.73E-16	1.10E-12	25.67702
IIZUKA_LIVER_CANCER_PROGRESSION_L0_L1_UP	-0.0416	0.16099	-12.3788	4.27E-15	5.22E-12	23.80365
REACTOME_TRIGLYCERIDE_CATABOLISM	0.045726	0.232139	12.28923	5.35E-15	5.22E-12	23.57325
BIOCARTA_PDZS_PATHWAY	0.047561	0.130834	12.21952	6.39E-15	5.22E-12	23.39329
REACTOME_IMMUNOREGULATORY_INTERACTIONS_BETWEEN_A_LYMPHOID_AND_A_NON_LYMPHOID_CELL	0.027025	0.08489	11.84565	1.67E-14	1.17E-11	22.41757
MASRI_RESISTANCE_TO_TAMOXIFEN_AND_AROMATASE_INHIBITORS_DN	0.050773	0.062411	11.2127	8.88E-14	5.43E-11	20.72519
WU_SILENCED_BY_METHYLATION_IN_BLADDER_CANCER	0.046016	-0.02372	10.69816	3.58E-13	1.95E-10	19.31138
KEGG_TYPE_I_DIABETES_MELLITUS	0.049998	0.030605	10.57021	5.10E-13	2.49E-10	18.95451
REACTOME_CGMP_EFFECTS	0.072695	-0.22651	10.47239	6.68E-13	2.97E-10	18.68025
YANG_BREAST_CANCER_ESR1_UP	0.034523	0.054927	10.38496	8.52E-13	3.30E-10	18.43406
REACTOME_NITRIC_OXIDE_STIMULATES_GUANYLATE_CYCLASE	0.06811	-0.24893	10.37492	8.76E-13	3.30E-10	18.40573
HUPER_BREAST_BASAL_VS_LUMINAL_UP	0.046941	0.153803	10.28457	1.13E-12	3.94E-10	18.15019
WP_EICOSANOID_SYNTHESIS	0.04415	-0.05014	10.23416	1.30E-12	4.24E-10	18.00716
MALTA_CURATED_STEMNESS_MARKERS	-0.04321	0.023754	-10.065	2.09E-12	6.40E-10	17.52472
REACTOME_INTRAFLAGELLAR_TRANSPORT	-0.02341	0.063196	-9.85413	3.81E-12	1.07E-09	16.91838
REACTOME_TRIGLYCERIDE_METABOLISM	0.030621	0.137119	9.843054	3.93E-12	1.07E-09	16.88638
NEWMAN_ERCC6_TARGETS_DN	-0.0457	0.03841	-9.53529	9.52E-12	2.45E-09	15.99083
STAMBOLSKY_RESPONSE_TO_VITAMIN_D3_UP	0.0315	0.072061	9.45878	1.19E-11	2.91E-09	15.76636
WANG_LSD1_TARGETS_UP	-0.02952	0.017125	-9.38857	1.46E-11	3.40E-09	15.55975
MASRI_RESISTANCE_TO_TAMOXIFEN_AND_AROMATASE_INHIBITORS_UP	0.05048	-0.20052	9.358123	1.59E-11	3.54E-09	15.46996
LI_LUNG_CANCER	-0.01839	0.434005	-9.261	2.12E-11	4.50E-09	15.18277
HOLLERN_MICROACINAR_BREAST_TUMOR_DN	-0.05382	0.196559	-8.98447	4.78E-11	9.75E-09	14.35893

C)

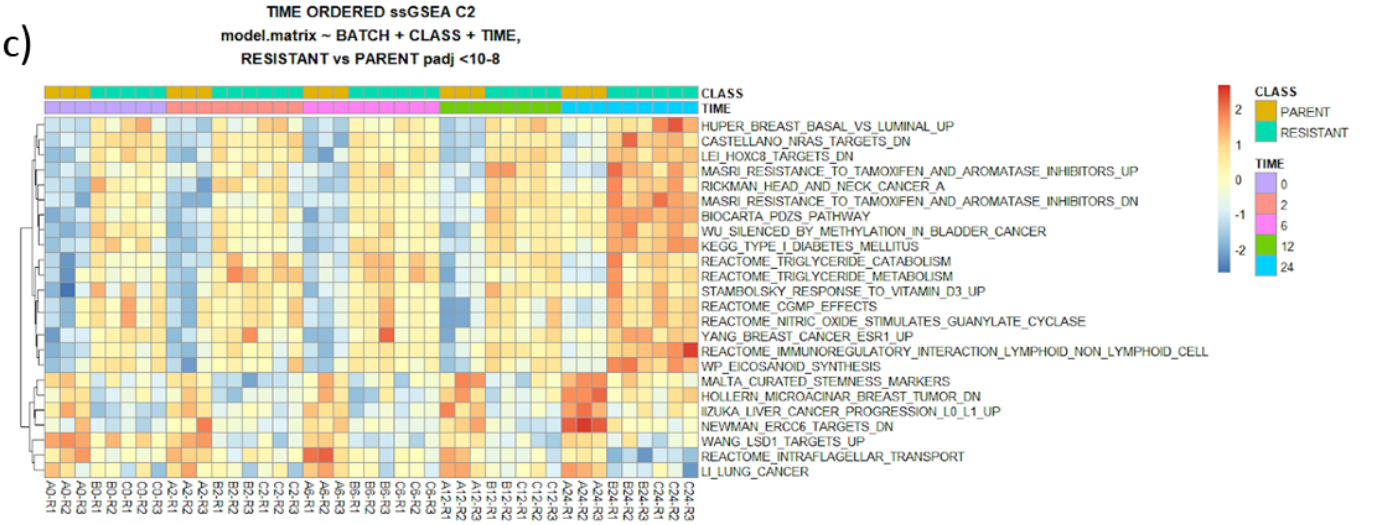


Figure 4.25 Single sample Gene Set Enrichment Analysis ssGSEA Heatmap
Significant (padj < 0.05) Curated ssGSEA Pathways, Panel A Heatmap of significantly differentially enriched Pathways between resistant and parent, pathways on y-axis, samples on x-axis, Panel B Table of top 20 differentially enriched pathways and Panel C Heatmap as per A with samples in time order. Enriched pathways are in red, reduced are in blue and this is dynamic. Visualised pathways are significantly differentially enriched (padj <0.001).

To further explore gene expression changes over time ssGSEA was performed and comparisons were made between Time 2 (2 hours after radiation) and Time 0 (baseline pre-radiation. The significantly (padj <0.05) differentially expressed pathways include pathways that have previously been demonstrated to be induced in response to Ionising Radiation, Response to IR pathways (by Warters, Smirnov, Warters and Macaeva), response to Cisplatin and P53 targets. This provides some degree of affirmation that the experiment is demonstrating results consistent with previous investigations. The expression of these pathways appears to be higher in the parental cell lines than the radioresistant cell lines (Figure 4.26).

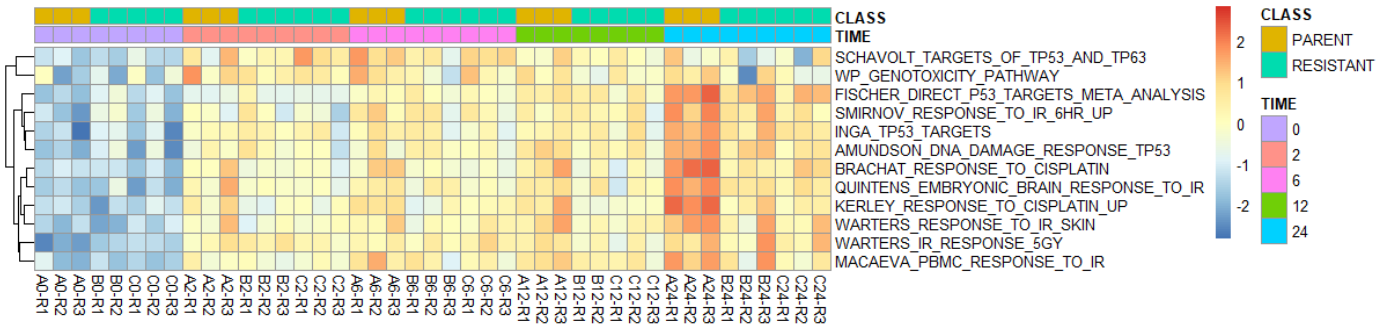


Figure 4.26 Heatmap of Single sample GSEA C2 Time 2: Time0 Significant

Heatmap of significantly differentially (padj < 0.05) enriched pathways between Time 2 and Time 0, pathways on y-axis, samples on x-axis by cell type and in time order

The analysis was repeated to compare Time 24 to baseline which again demonstrated the previously identified pathways as being significantly increased at 24 hours, as well as additional pathways including Fridman Senescence UP and Yamashita Methylated in Prostate Cancer (Figure 4.27).

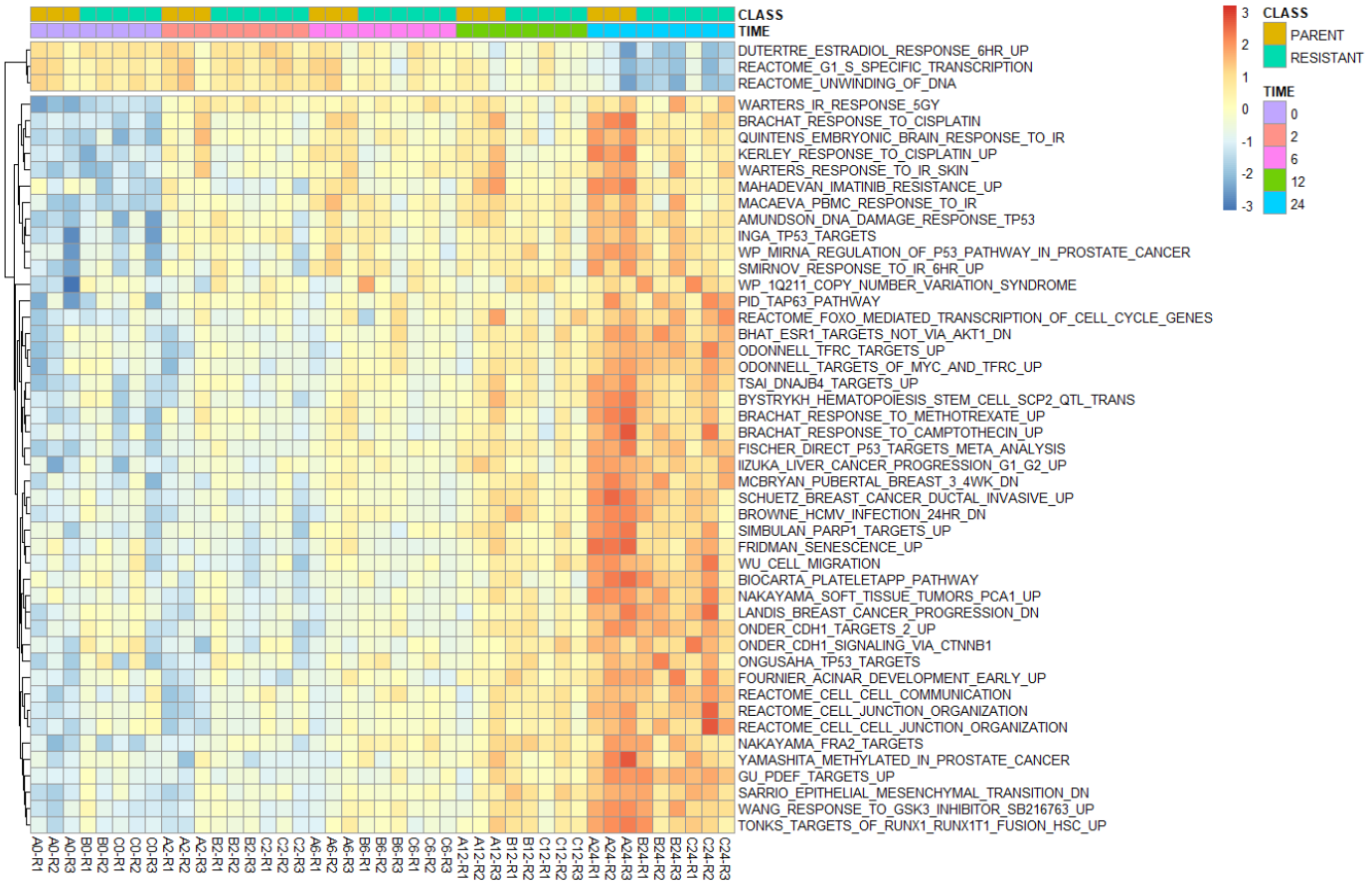
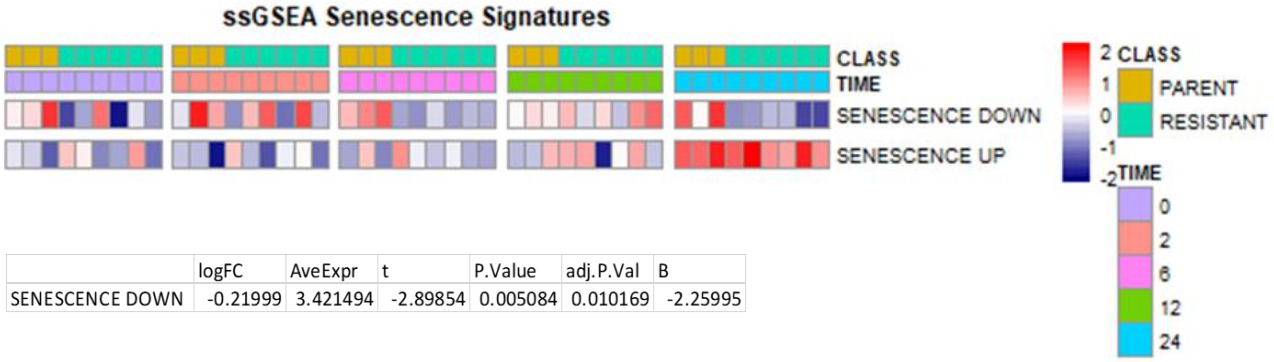


Figure 4.27 Heatmap of Single sample GSEA C2 Time 24: Time 0

Heatmap of significantly differentially (padj < 0.05) enriched pathways between Time 2 and Time 0, pathways on y-axis, samples on x-axis ordered by cell type and in time order. Significantly enriched pathways at 24 hours are seen in red (bottom panel), reduced are in blue(top panel).

The ssGSEA results of an increased senescence gene expression profile at 24 hours following radiation mimic the findings of the cell line experiments in Results Chapter 1 which demonstrated an increase in senescence in both parental and resistant cell lines 24 hours following radiation. This was confirmed with 2 separate techniques C12 FDG and Beta-Galasctosidase staining. To further explore gene expression of senescence a recently reported transcriptomic signature of senescence was identified which was generated by taking the overlapping sets of genes that were increased and decreased in a number of cell types confirmed to be senescent due to either IR or Doxorubicin induction. The published signatures were applied using ssGSEA (using Ensembl based mapping counts) and differential expression analysis between Resistant and Parental samples performed.



	logFC	AveExpr	t	P.Value	adj.P.Val	B
SENESCENCE DOWN	-0.21999	3.421494	-2.89854	0.005084	0.010169	-2.25995

Figure 4.28 Heatmap of single sample GSEA of Senescence Signatures
Heatmap of single sample gene gene set enrichment analysis of senescence pathways, pathways on y-axis, samples on x-axis ordered by cell type and in time order
Table of significantly differentially (padj < 0.05) enriched senescence pathways between resistant and parental cell lines.

The ssGSEA analysis demonstrates that the Senescence down geneset which is previously demonstrated to be reduced in senescent cells is differentially expressed between the parental and radioresistant lines (padj = 0.01) which can be clearly seen at Time 24 (Figure 4.26, and heatmap Appendix 2.5). As per previously the Senescence Up signature is increased in expression at 24 hours, which supports the previous findings in Chapter 1 of increased senescence at 24 hours post IR. The differential expression of the senescence down genes potentially supports the findings that a greater proportion of radioresistant cells were senescent 24 hours following radiation.

4.23 Time interaction model with shrinkage H460

To further explore the interactions between conditions a time reduced interaction model was used incorporating a generalised linear model shrinkage technique (apeglm) which uses Bayesian shrinkage estimators for effect sizes, using approximation of the posterior for individual coefficients of differential expression. This identified 6 genes that are significantly decreased (padj <0.01) and 20 genes significantly increased Figure 4.29.

Figure 4.30 demonstrates the relevant log fold changes of these significant genes identified for each possible comparison e.g. left side demonstrates each time comparison, right side Resistant Samples at each timepoint.

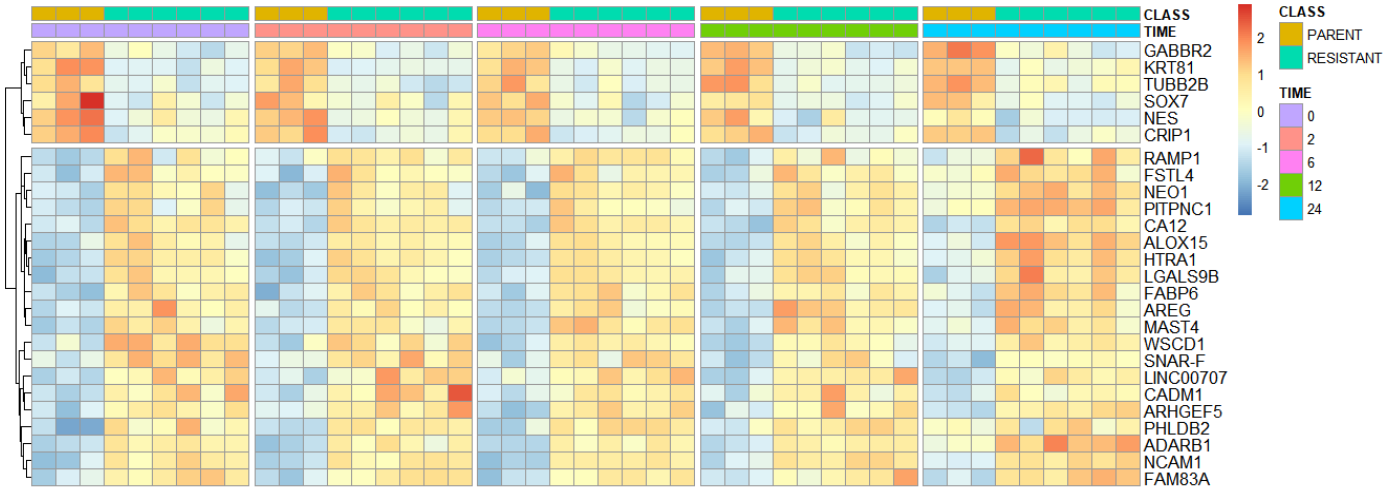


Figure 4.29 Heatmap of significantly (padj <0.01) differentially expressed genes between Radioresistant and Parent cell lines (Time interaction model) with shrinkage

Heatmap of significantly differentially (padj < 0.05) expressed genes between resistant and parental cell lines, cells on y-axis, samples on x-axis by cell type and in time order. Genes in the top panel are decreased (blue cells) in resistant cell lines and increased (red cells) in the bottom panel.

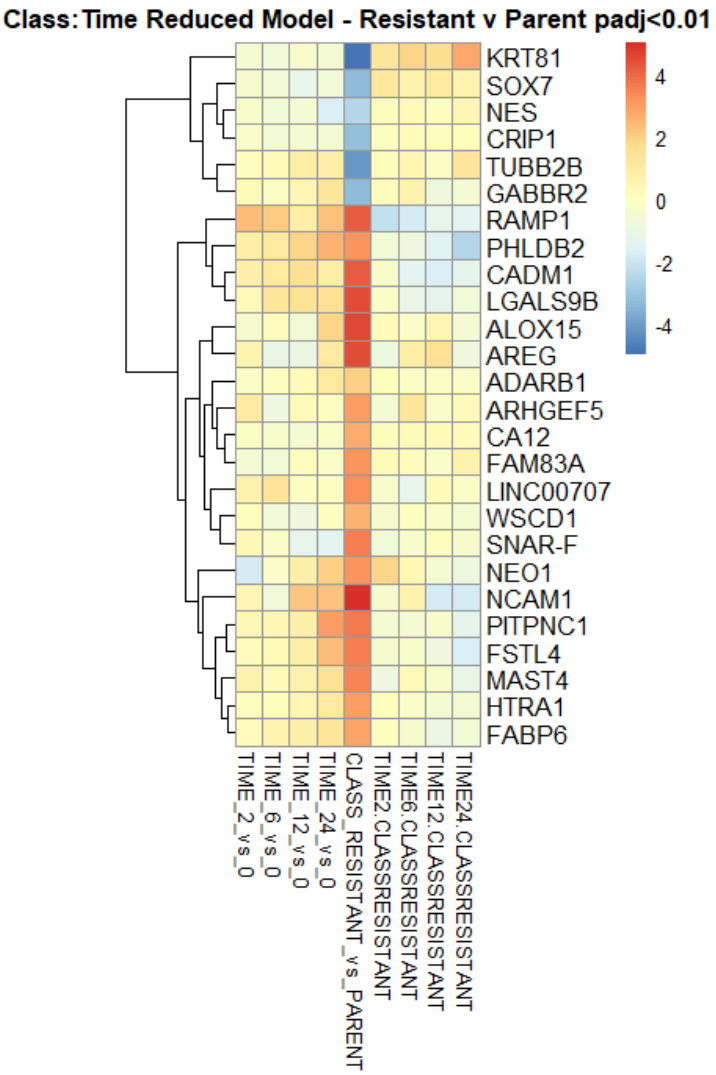


Figure 4.30 Heatmap of significantly (padj <0.01) differentially expressed genes between Radioresistant and Parent cell lines (Time interaction model) with shrinkage

Genes displayed on y axis, with comparisons on x axis, colour demonstrates log fold change

This compact list of genes is taken forward to explore whether they are prognostic in a cohort of radiotherapy treated lung cancer patients (4.29 External Prognostic Validation).

HT29 Cell Line Analysis

The HT29 cell line analysis was conducted as per methods section. Principal Component analysis did not demonstrate any clear separation of Resistant vs Stable samples or across

times. (Figure 4.31). This suggested that the utility of any differential expression analysis may be relatively limited.

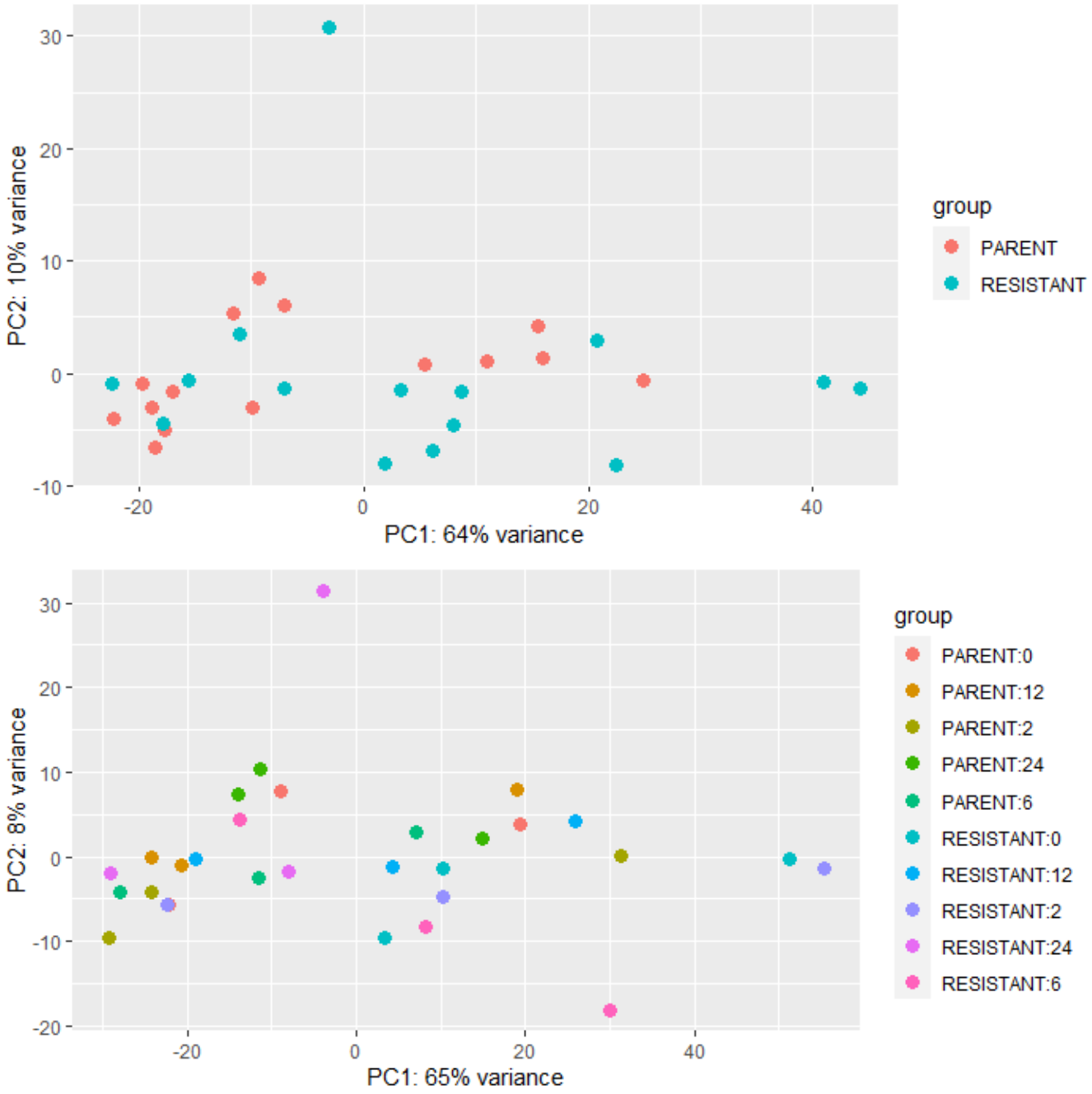


Figure 4.31 Principal Component Analysis HT29
Top plot demonstrates PCAs with parental and resistant samples, Bottom plot demonstrates individual timepoints for each

Utilizing the same method as previously a time reduced interaction model was used incorporating a generalised linear model shrinkage technique, to identify significantly (padj

<0.05) differentially expressed genes between resistant and parent samples across all time points (Figure 4.32)

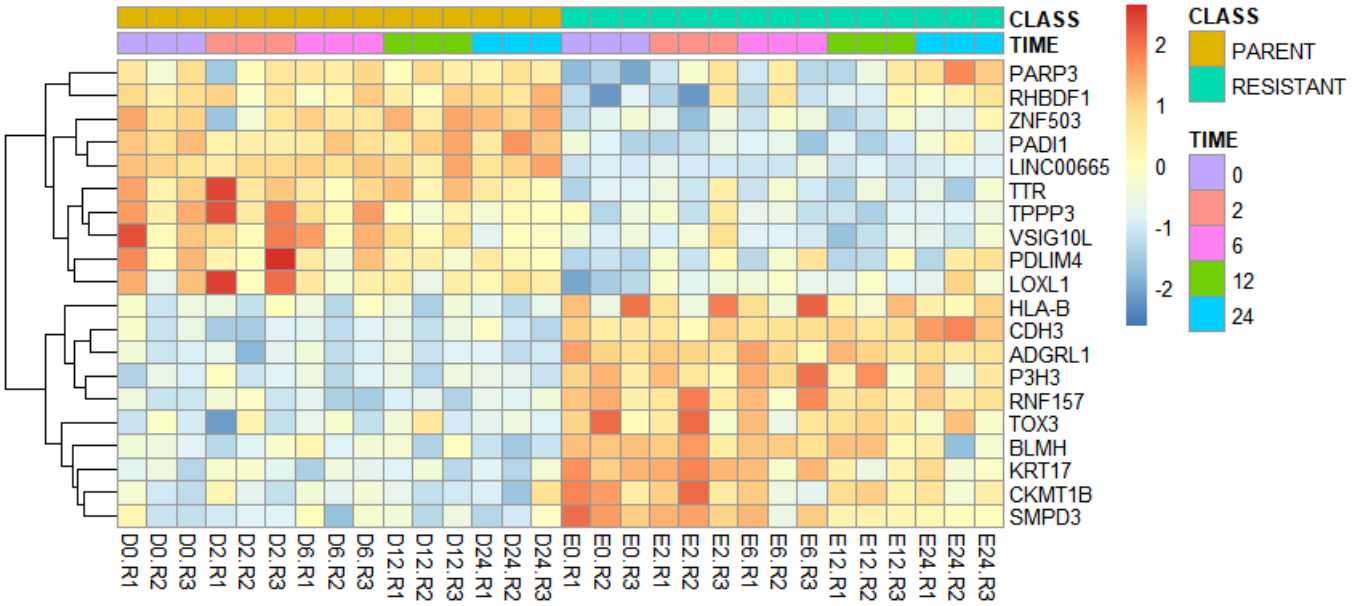


Figure 4.32 Heatmap of significantly (padj <0.05) differentially expressed genes between HT29 Radioresistant and Parent cell lines (Time interaction model) with shrinkage

Heatmap of significantly differentially (padj < 0.05) expressed genes between resistant and parental cell lines, cells on y-axis, samples on x-axis by cell type and in time order

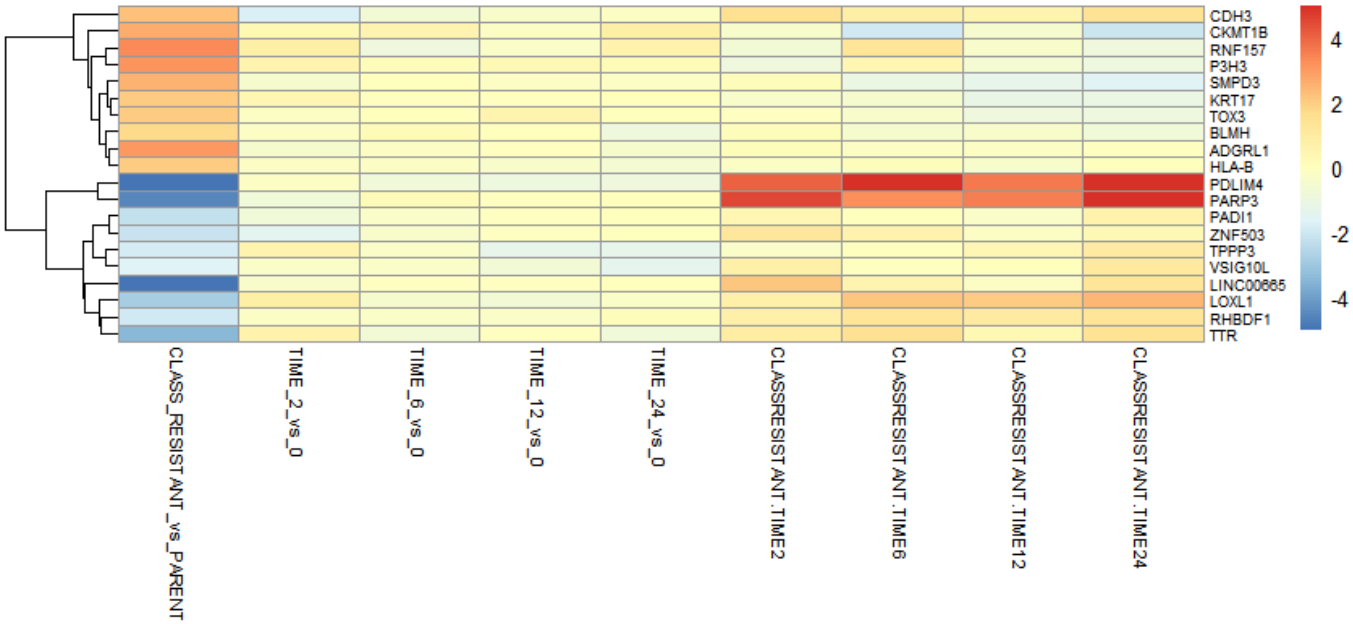


Figure 4.33 Heatmap of significantly (padj <0.05) differentially expressed genes between HT29 Radioresistant and Parent cell lines (Time interaction model) with shrinkage

Colours demonstrate log fold changes of the genes (y-axis) for each of the comparisons (x-axis)

Expression of the significantly differentially expressed genes in the parental (orange) and radioresistant (green) samples demonstrates clearly the differences in expression levels at each timepoint (Figure 4.32). The heatmap in Figure 4.33 demonstrates the relevant log fold changes of the genes (rows) for each of the comparisons (columns).

4.25 A549 Cell Line Analysis

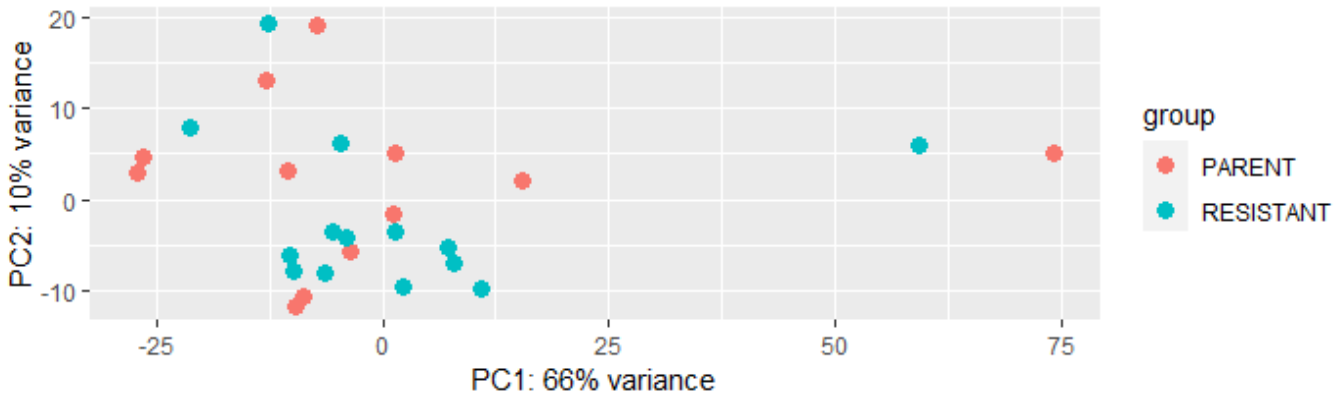


Figure 4.34 Principal Component Analysis A549

Principal Component Analysis did not demonstrate any clear separation between Radioresistant and parental cell lines, unfortunately multiple samples from Time 0 of the parental cell line failed sequencing thus preventing any meaningful differential expression analysis.

4.26 Methylation analysis

Methylation analysis was conducted as per Methods section. Quality Control metrics were generated and are available in the Appendix. Samples were run over 2 Methylation Chips therefore Principal Component Analysis was conducted to ensure that this was not a technical factor for differences seen between radioresistant and parental samples. The PCA plot Figure 4.35 demonstrates clear separation between the HT29 and H460 cell lines across PC1, which is expected as this mirrors the previous PCA findings for gene expression Figure 4.1, whereby the greatest difference is between differing underlying cell types. Principal Component 2 demonstrates clear separation of the H460 Radioresistant and Parental Cell Lines.

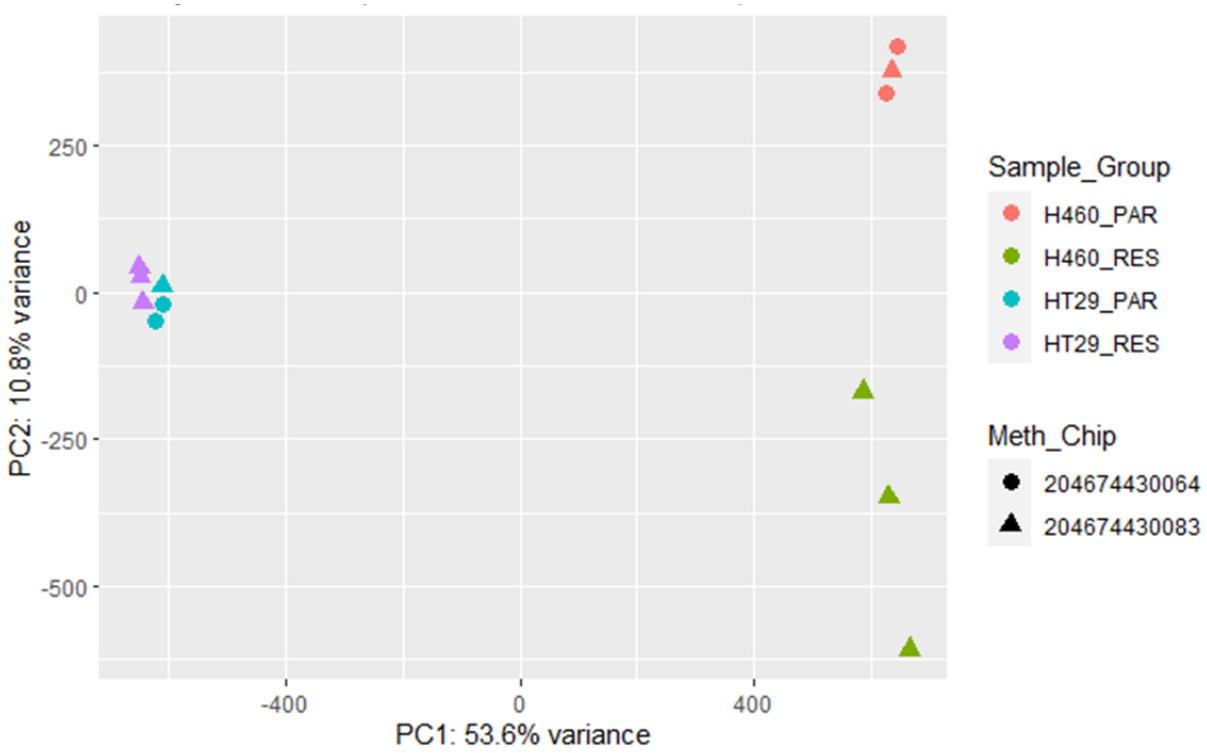


Figure 4.35 Principal Component Analysis Methylation Methylation PCA plot using filtered and normalised (Mvals). Cell types separate by PC1, H460 resistant and parental separate by PC2.

Differential Methylation analysis was performed for both cell lines to compare resistant to parental samples to identify differentially methylated probes and regions as per Methods Section. The samples used for methylation analysis are all baseline pre-radiation samples (Time 0).

4.27 Differential methylation analysis H460

The top Differentially Methylated Region (DMR) corresponds to FAM83A, which was also one of the most differentially expressed genes (Figure 4.29). Figure 4.36 demonstrates on the left the Average Meta (Methylation) for each probe within the FAM83A gene, the average Betas for FAM83A can be seen to be lower and this is statistically significant (Wilcoxon p = 0.042). The right-hand panel demonstrates the RNAseq experiment data showing the mean normalised gene expression for FAM83A across all timepoints. FAM83A was noted to be

one of the most significantly (padj <0.01) differentially expressed genes between resistant and parent using the time interaction model and shrinkage (Figure 4.26).

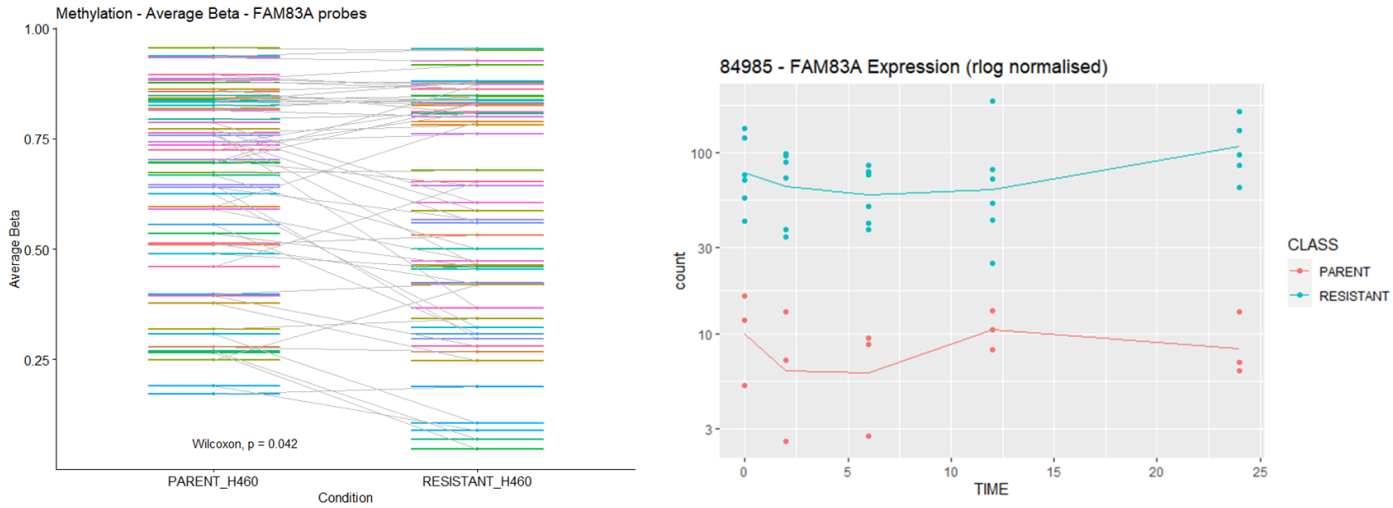


Figure 4.36 Differentially Methylated Probes and 3' RNAseq expression

Left Methylation per FAM83A probe (Betas) comparing Parent and Resistant Wilcoxon p = 0.042, , Right FAM83A RNAseq expression over time is significantly increased in parental cell lines (rlog normalised values, padj <0.01)

FAM83A was the top Differentially Methylated Region (DMR), the DMR plot (Figure 4.37) demonstrates the relative genomic location of the probes associated with FAM83A, and the smoothed group means (bottom) can be seen to differ between the Radioresistant and Parental samples. The top 25 DMRs are listed in the Appendix 2.12.

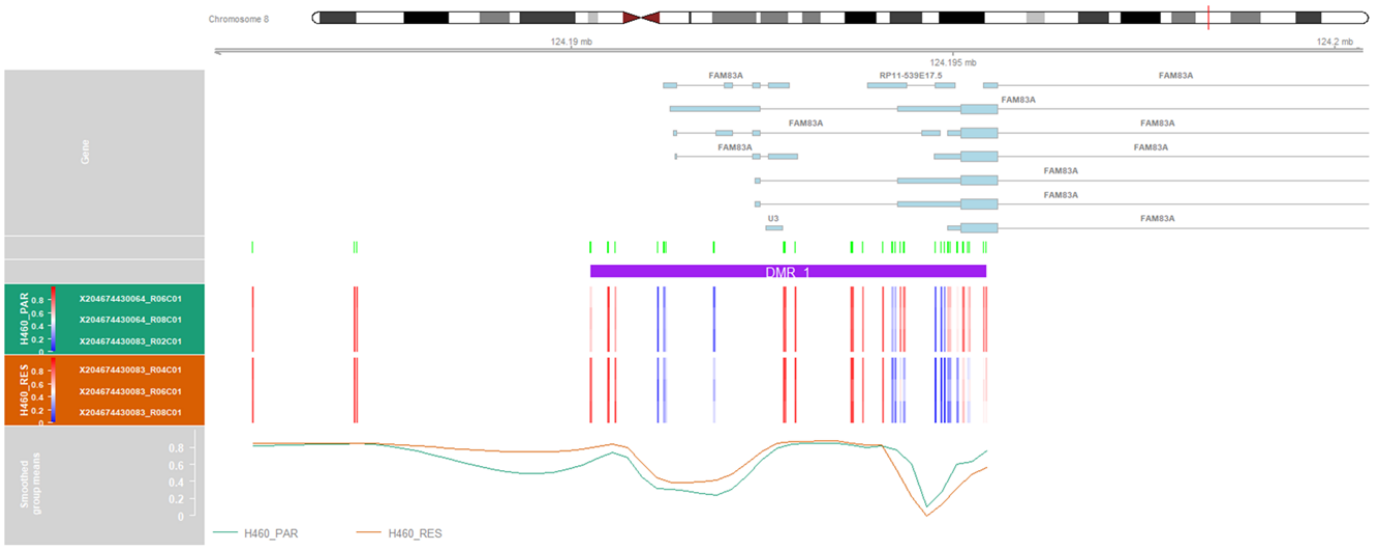


Figure 4.37 Differentially Methylated Region Plot FAM83A

Top panel demonstrates genomic location of the probes associated with FAM83A, Middle demonstrates relative methylation of each probe, Bottom demonstrates smoothed group means per probe, demonstrating differential methylation.

In order to further investigate the pathways associated with differentially methylated genes, pathway analysis was performed with Over Representation using the C2 Curated Genesets from MSigDB. This demonstrated that the hypomethylated genes in Radioresistant samples compared to parental samples enriched for pathways including UV response via ERCC3 Down, TP53 targets, SUZ12 targets, PRC2 targets, Apoptosis by Doxorubicin, BMP targets and multiple pathways associated with H3K27ME3 (Figure 4.38). Pathway analysis of genes that were hypermethylated in Radioresistant compared to parental samples included TP53 targets, SUZ12 targets, PRC2 targets, pathways associated with Prostate cancer (Liu, Schaeffer) and again multiple pathways associated with H3K27ME3 (Figure 4.39).

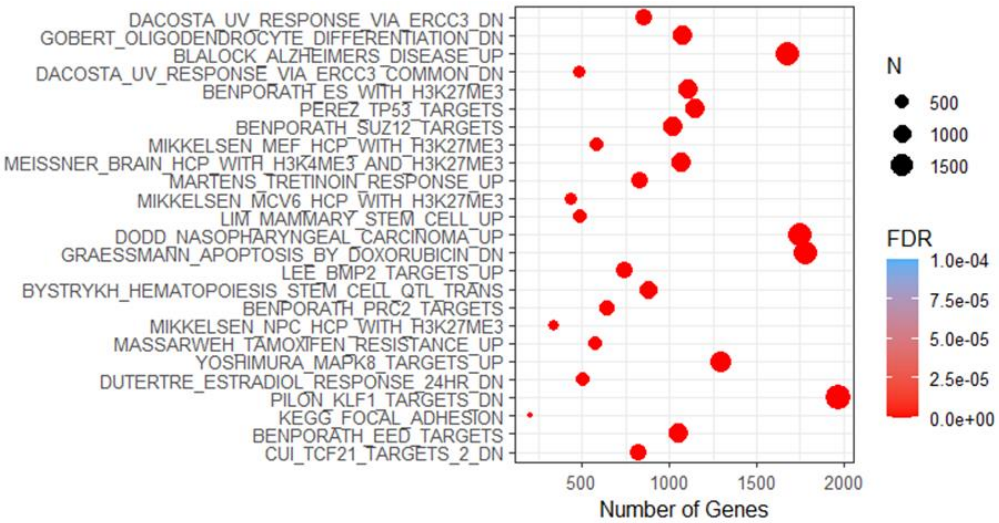


Figure 4.38 Decreased Methylation Over Representation Pathway Analysis Resistant vs Parent H460

Pathways are demonstrated on y-axis, number of hypo-methylated genes on x-axis, significance by False Discovery Rate demonstrated by colour (FDR < 0.01)

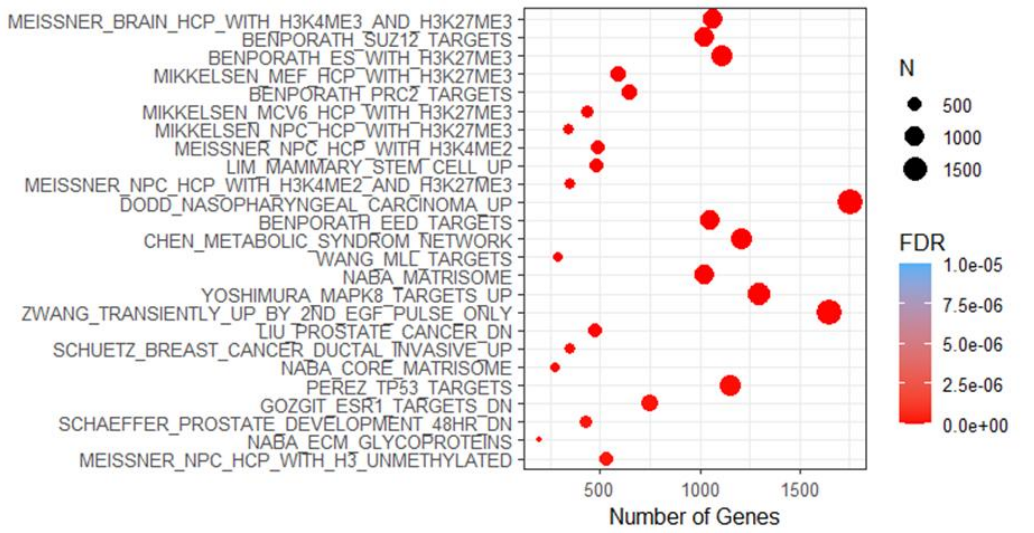


Figure 4.39 Increased Methylation Over Representation Pathway Analysis Resistant vs Parent H460

Pathways are demonstrated on y-axis, number of hyper-methylated genes on x-axis, significance by False Discovery Rate demonstrated by colour (FDR < 0.01)

Similar methylation analysis was conducted with HT29 cells to compare radio-resistant with parental cell lines which demonstrates similar pathways as being differentially methylated (Appendix 2.7). To further explore whether radioresistant cell lines had similar methylation events a Methylation Density Map (Figure 4.40) was constructed to visualise the genomic location of the differentially methylated regions between radioresistant and parent cell lines. This demonstrates that different regions methylated in H460 and HT29 when comparing resistant to parental samples, and that the HT29 samples are predominantly hypo-methylated in contrast to the H460 cell lines.

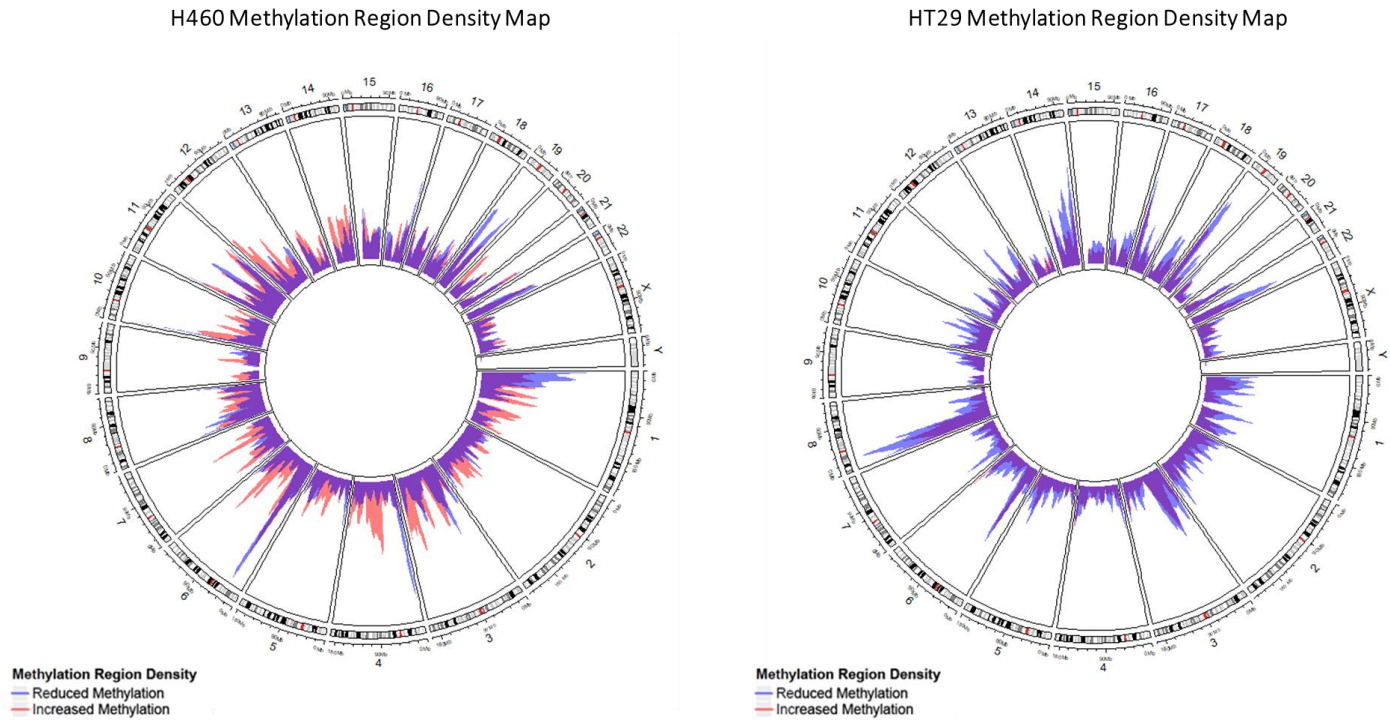


Figure 4.40 Methylation Region Density Maps comparing Resistant to Parent
Methylation Region Density Map, reduced methylation in blue, increased methylation in red, genomic location on circumferential plot. Left demonstrates methylation in H460 cell lines Right demonstrates HT29 lines.

The overlapping hyper and hypo-methylated probes in both HT29 and H460 were computed, the venn diagram in Figure 4.41A demonstrates that there are 529 probes that are commonly Hypermethylated (increased methylation) in the radioresistant vs parental samples from both cell lines. The venn diagram in Figure 4.41B demonstrates that there are 640 probes that are commonly Hypomethylated (reduced methylation) in the radioresistant vs parental samples from both cell lines. The relative methylation of probes is visualised in the methylation heatmaps (Figure 4.41, A,B bottom).

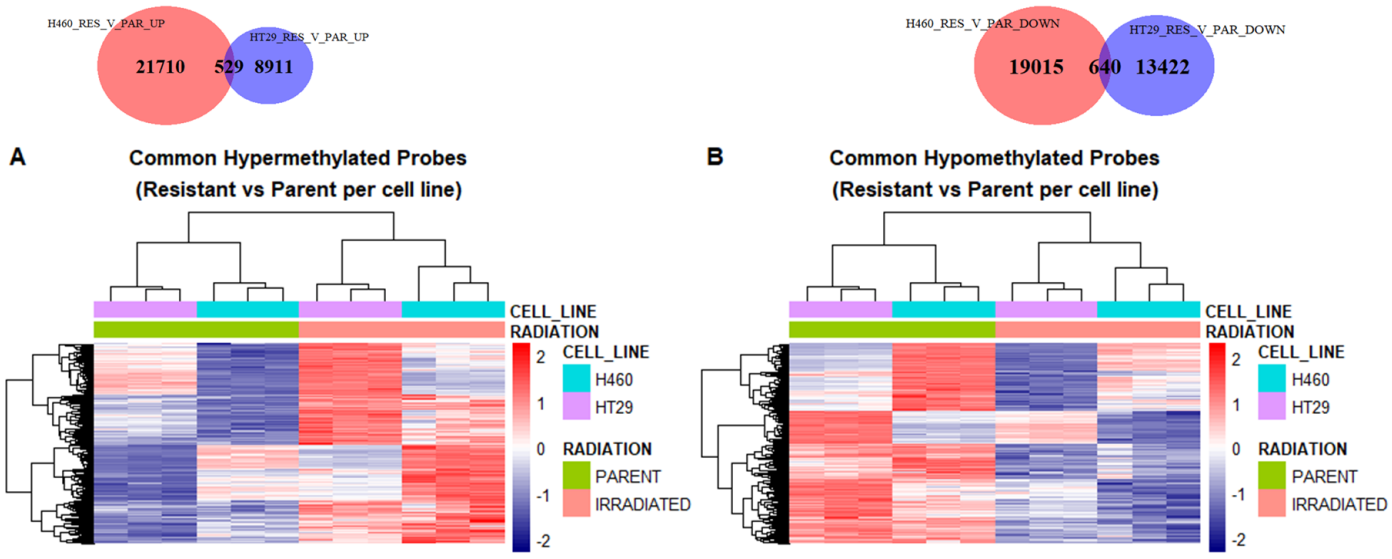


Figure 4.41 Commonly Differentially Methylated Probes
A demonstrates Common hypermethylated probes (padj <0.05), top demonstrates Venn Diagram of overlap between H460 and HT29, bottom heatmap of methylation
B demonstrates Common hypomethylated probes (padj <0.05), top demonstrates Venn Diagram of overlap between H460 and HT29, bottom heatmap of methylation

Of the overlapping probes that are hypomethylated in radioresistant cell lines, the top 3 hypomethylated genes were ADAMTS9 (5 probes), CSMD1 (4 probes) and TCF4 (3 probes).

4.28 Top differentially Methylated Region H460 - FAM83A

In order to further investigate the top Differentially Methylated Region (Figure 4.39) which is also one of the top differentially expressed genes (Figure 4.27) between H460 radioresistant and parental lines and identify potential pathways associated with FAM83A RNA gene expression a correlation analysis (Spearman) was performed between FAM83A and all other genes in the RNA dataset (Appendix 2.8). The list of genes was then ordered by correlation and GSEA performed on the ranked list which identified pathways including Cell Cycle, Cellular Senescence, and p53 pathway as being significantly enriched (padj <0.05) in the ranked gene list (Appendix 2.9/2.10)

4.29 External prognostic validation cell line results in TCGA Lung Adenocarcinoma dataset

To explore whether the identified geneset, increased in expression in radioresistant H460 cells, (Figure 4.26) is prognostic in Radiotherapy treated lung cancer patients, The Cancer Genome Atlas Lung Adenocarcinoma (TCGA LUAD) dataset was investigated. Patients who received radiotherapy were identified, clinical characteristics are demonstrated below in Figure 4.42.

	Overall (N=61)
Age (y)	
Mean (SD)	62.6 (11.3)
Median [Min, Max]	63.0 [39.0, 86.0]
Sex	
Male	26 (42.6%)
Female	35 (57.4%)
Stage	
STAGE I	1 (1.6%)
STAGE IA	9 (14.8%)
STAGE IB	8 (13.1%)
STAGE IIA	4 (6.6%)
STAGE IIB	8 (13.1%)
STAGE IIIA	22 (36.1%)
STAGE IIIB	3 (4.9%)
STAGE IV	6 (9.8%)
Radiation Therapy	
Yes	61 (100%)

Figure 4.42 Baseline Characteristics of Radiotherapy Treated Cohort in TCGA LUAD

The previously identified genes significantly (padj <0.01) increased genes between Radioresistant and Parent cell lines Figure 4.29 were applied to the TCGA dataset, of the 20 genes only 10 genes were expressed with non-zero values. The expression of these genes is demonstrated in Figure 4.43 with clinical Stage and Overall Survival Status (0 - alive, 1 - dead). The genes do not cluster the patients into separate groups and the expression of the individual genes is not highly correlated (Appendix 2.11

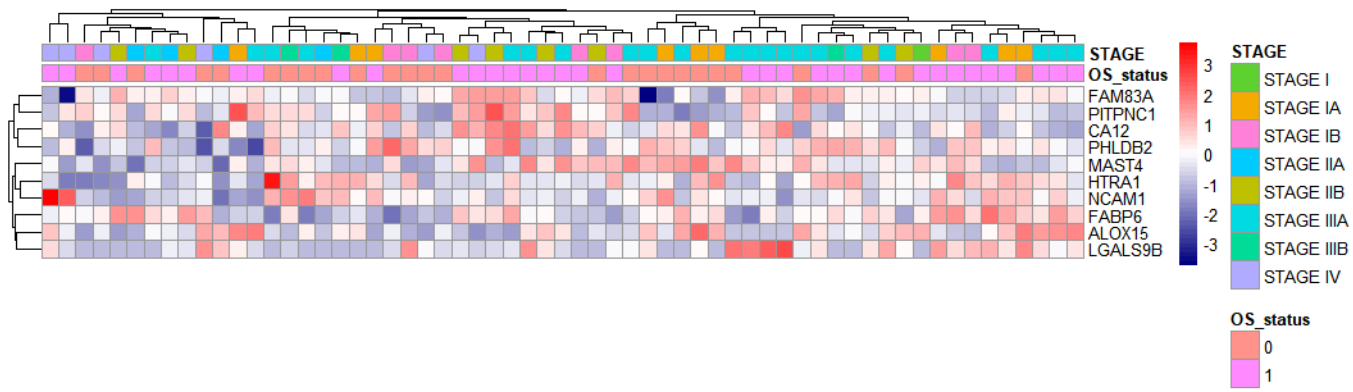


Figure 4.43 Heatmap of significantly (padj <0.01) differentially expressed genes between Radioresistant and Parent cell lines TCGA LUAD data
Genes are demonstrated on y axis, patients on x axis, non-heirarchical clustering demonstrates no notable clusters

The mean expression of these 10 genes was calculated per patient, denoted as Lung_Score_Scaled henceforth, Boxplot 2.46 demonstrates that the mean expression of the 10 genes is significantly different in the cohort of patients who are dead compared to those who are alive Wilcoxon p = 0.0018.

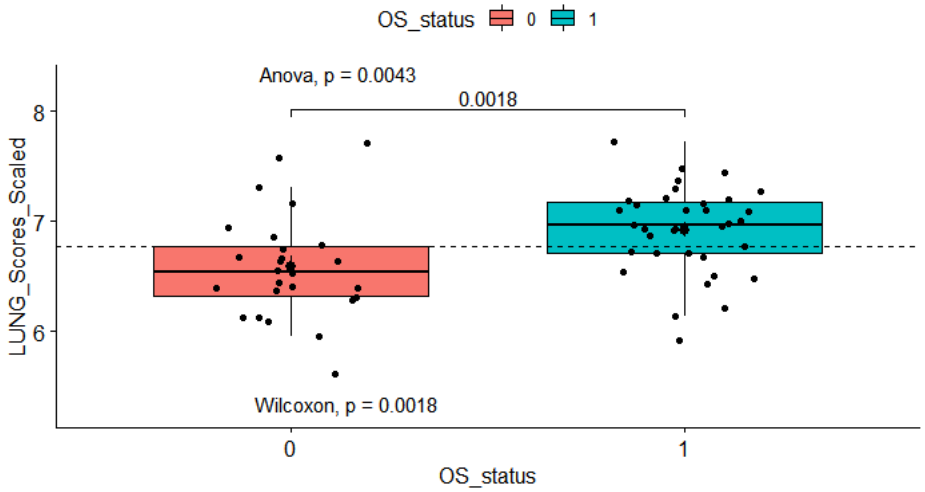


Figure 4.44 Boxplot of significantly (padj <0.01) differentially expressed genes between Radioresistant and Parent cell lines TCGA LUAD data (median cutoff)
Boxplot demonstrating mean expression of 10 genes significantly (padj <0.01) differentially expressed genes between Radioresistant and Parent cell lines, significantly (Wilcoxon p = 0.0018, Anova p = 0.0043) differs in expression between patients who are alive (red boxplot) compared to those who are dead (green boxplot) following radiotherapy.

In order to explore whether the expression of the 10 genes is associated with survival the mean expression of all 10 genes was calculated per patient and the cohort divided in half by

the median value. Kaplan Meier survival curves demonstrate a difference in overall survival between the radiotherapy treated patients with high expression vs low expression (median cutoff) of the genes significantly increased in radioresistant cell lines (Figure 4.45).

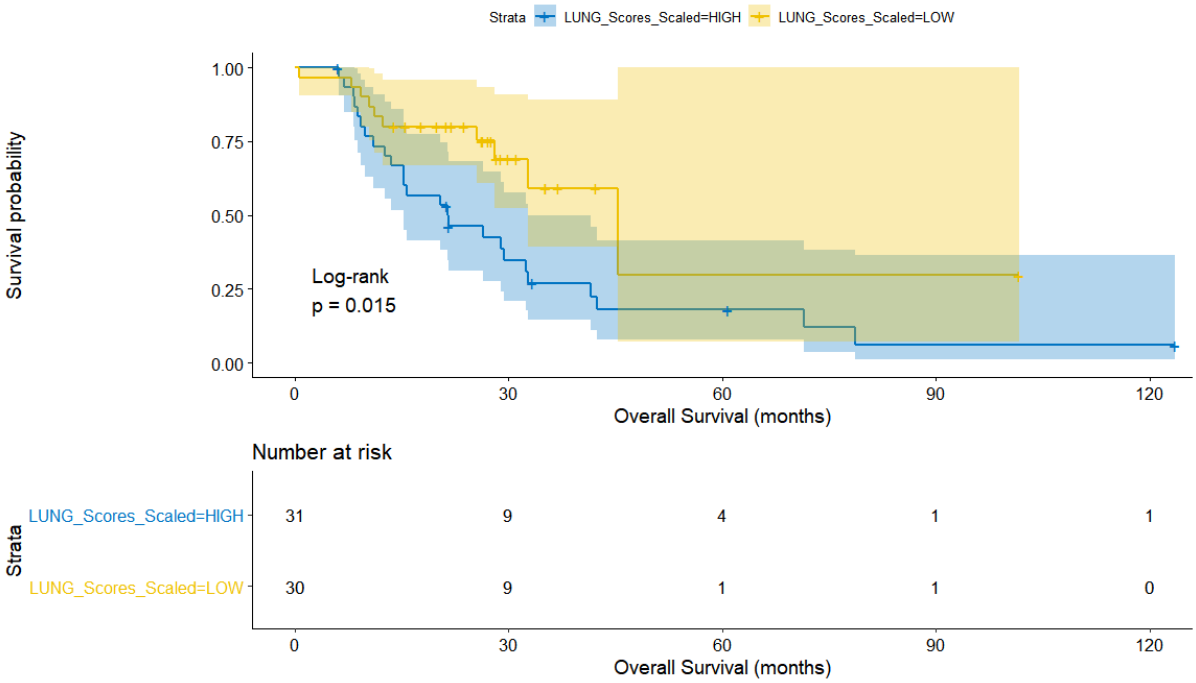


Figure 4.45 Survival Curves of differentially expressed genes between Radioresistant and Parent cell lines in Radiotherapy treated patients TCGA LUAD data (median cutoff)

Kaplan-Meier survival analysis comparing high vs low expression cohorts (median cutoff) of radiotherapy treated patients demonstrates significance p = 0.015 (log-rank).

A Cox Proportional hazards model was used including expression of the genes significantly increased in radioresistant cell lines (median cutoff) as well as clinical Stage. This demonstrated the cohort with lower expression of radioresistant genes had a better survival HR 0.45 (0.21-0.99) which was statistically significant at a pvalue of 0.048 (Figure 4.46).

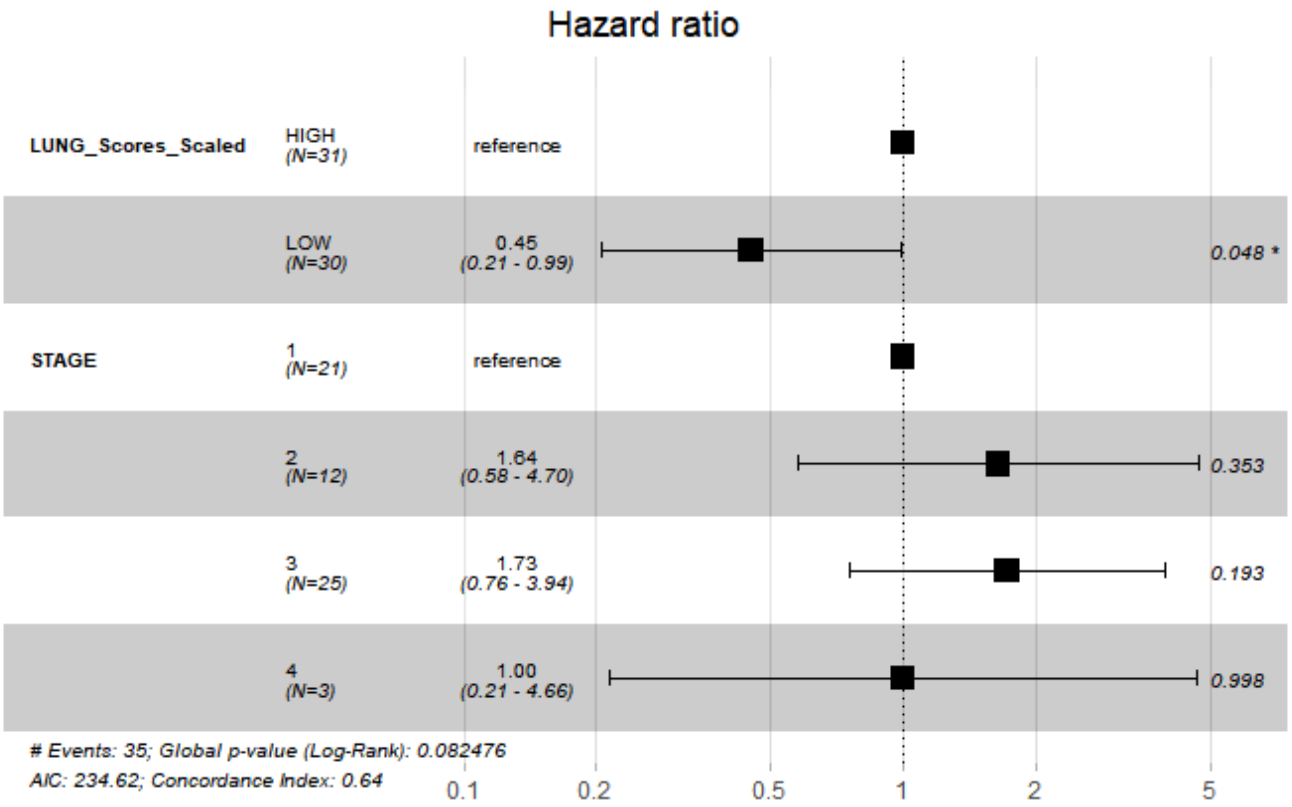


Figure 4.46 Forest Plot of Hazard ratios for Radiotherapy treated patients genes significantly increased in radioresistant cell lines (median cutoff) and clinical Stage. Demonstrates a Hazard Ratio of 0.45(0.21-0.99) low expression cohort (p = 0.048)

To further explore whether there were similar differences in outcome for non-radiotherapy

treated patients, the analysis was repeated in the non-radiotherapy treated cohort.

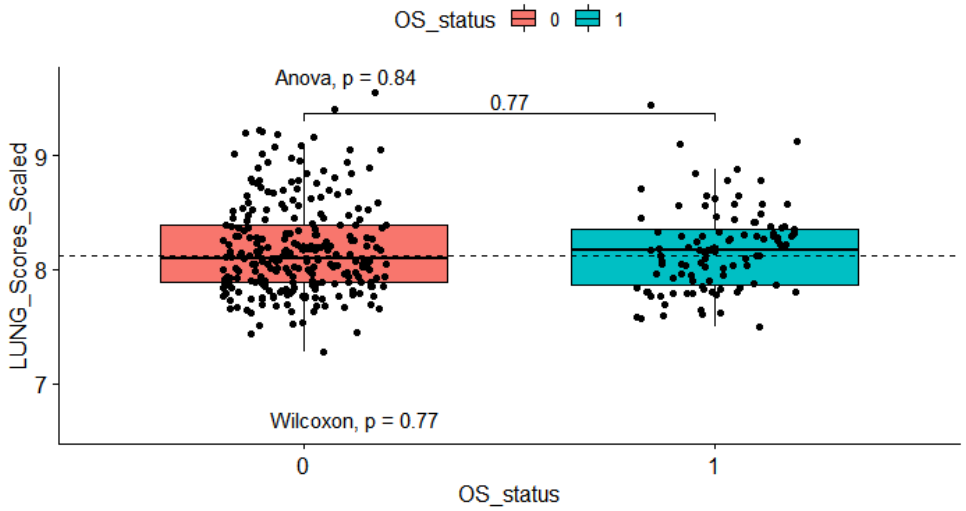


Figure 4.47 Boxplot of significantly (padj <0.01) differentially expressed genes between Radioresistant and Parent cell lines TCGA LUAD data (median cutoff) Mean expression of 10 genes significantly (padj <0.01) differentially expressed genes between Radioresistant and Parent cell lines, does not differ in expression between patients who are alive compared to those who are dead in cohorts not treated with radiotherapy.

There was no difference in the mean expression of the 10 genes between the cohort of patients who are dead compared to those who are alive in the non-radiotherapy treated cohort Figure 4.45. Kaplan Meier survival curves did not demonstrate a significant difference in overall survival (log-rank p = 0.22) in non-radiotherapy treated patients divided into high expression vs low expression cohorts (median cutoff) of the genes significantly increased in radioresistant cell lines (Figure 4.48). The survival curves can be seen to cross and although not significant the cohort with lower median expression appear to trend towards having worse overall survival, which is the converse to the radiotherapy treated patients, where high expression was associated with worse survival (Figure 4.45)

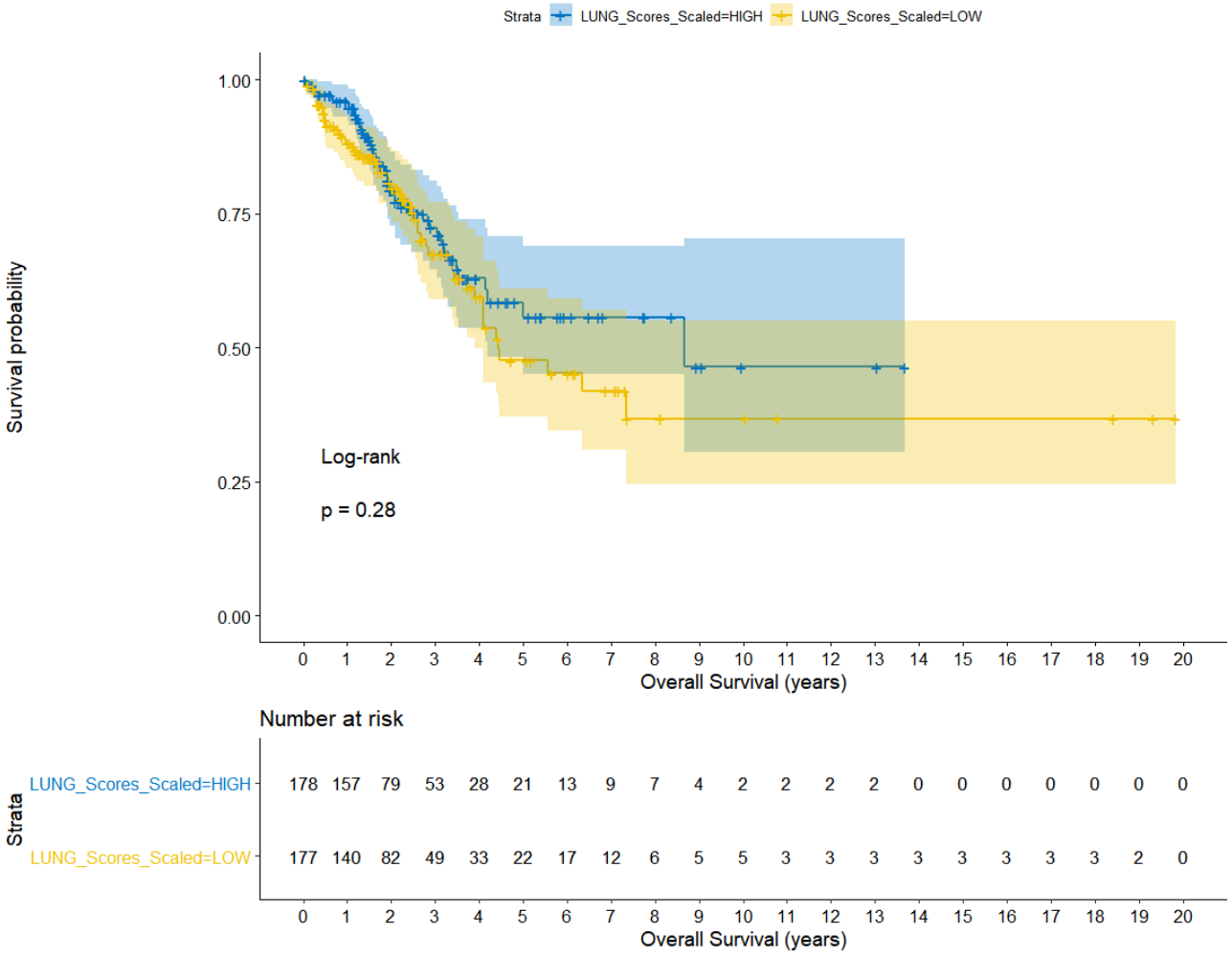


Figure 4.48 Survival Curves of differentially expressed genes between Radioresistant and Parent cell lines in Non-Radiotherapy treated patients TCGA LUAD data (median cutoff)

Kaplan-Meier survival analysis comparing high vs low expression cohorts (median cutoff) of patients not treated with radiotherapy demonstrates no difference in survival.

These findings demonstrate that genes significantly increased in radioresistant cell lines are prognostic in a radiotherapy treated cohort of patients with lung cancer. The cohort of patients who were dead had a statistically significant higher expression of these genes Wilcoxon p= 0.0018 (Figure 4.44) , this was not the case in the non-radiotherapy treated cohort, Wilcoxon p = 0.77 (Figure 4.47). There is a difference in survival as seen by survival curves between High and Low expression cohorts (median cutoff) in patients treated with radiotherapy, with higher expression of genes increased in radioresistant cell lines having worse overall survival log-rank p = 0.015 (Figure 4.45), this was not replicated in the non-radiotherapy treated cohort (p = 0.28) (Figure 4.58). Cox Proportional hazards model including Stage and expression of genes increased in radioresistant cell lines demonstrated that the cohort with lower expression of radioresistant genes had a better survival HR 0.45 (0.21-0.99), which was statistically significant at a pvalue of 0.048 (Figure 4.46).

4.30 External prognostic validation senescence signature in Stratification in Colorectal cancer (SCORT) dataset

The SCORT dataset includes patients with rectal adenocarcinoma, treated with neo-adjuvant chemoradiotherapy with data including each patient's clinical response to neo-adjuvant radiotherapy. The dataset contains transcriptomic data mRNA from micro-arrays, with clinical followup and parameters including response to treatment.

	Overall (N=223)
PRE_T_STAGE	
	44 (19.7%)
1	1 (0.4%)
2	19 (8.5%)
3	127 (57.0%)
4	11 (4.9%)
no MRI	3 (1.3%)
x	3 (1.3%)
Missing	15 (6.7%)
PRE_N_STAGE_ADJUSTED	
	44 (19.7%)
0	67 (30.0%)
1	67 (30.0%)
2	27 (12.1%)
x	2 (0.9%)
Missing	16 (7.2%)
Response.to.Treatment	
Complete	33 (14.8%)
Good Partial	77 (34.5%)
Minimal	7 (3.1%)
Partial	106 (47.5%)

Figure 4.49 Clinical Characteristics of SCORT Cohort
Patient characteristics in SCORT cohort, PRE_T_STAGE = pretreatment T stage, PRE_N_STAGE = pretreatment nodal stage, Response.to.Treatment = response to treatment

In order to assess the prognostic value of the Fridman Senescence UP signature, which was previously demonstrated to be significantly increased in radioresistant and parental cell lines 24 hours following radiation compared to baseline, ssGSEA scores were calculated for each patient.

Mean ssGSEA scores were calculated per response to treatment group (Complete, Good Partial, Partial and Minimal) and demonstrated to be significantly lower in Complete responders than good partial responders p = 0.0017, and lower in Complete responders than partial responders p = 0.00015 (Figure 4.50 Panel A). The density distribution for each group is seen in Appendix 1.13. The cohort was then divided into High and Low cohorts by the median ssGSEA score. There is a higher percentage of complete responders and lower percentage of partial responders in the cohort with lower expression of the senescence

pathway Figure 4.50 Panel B, which is demonstrated to be statistically significant, see contingency table Figure 4.50 Panel C with a Kruskal-Wallis rank sum test (pvalue 0.018).

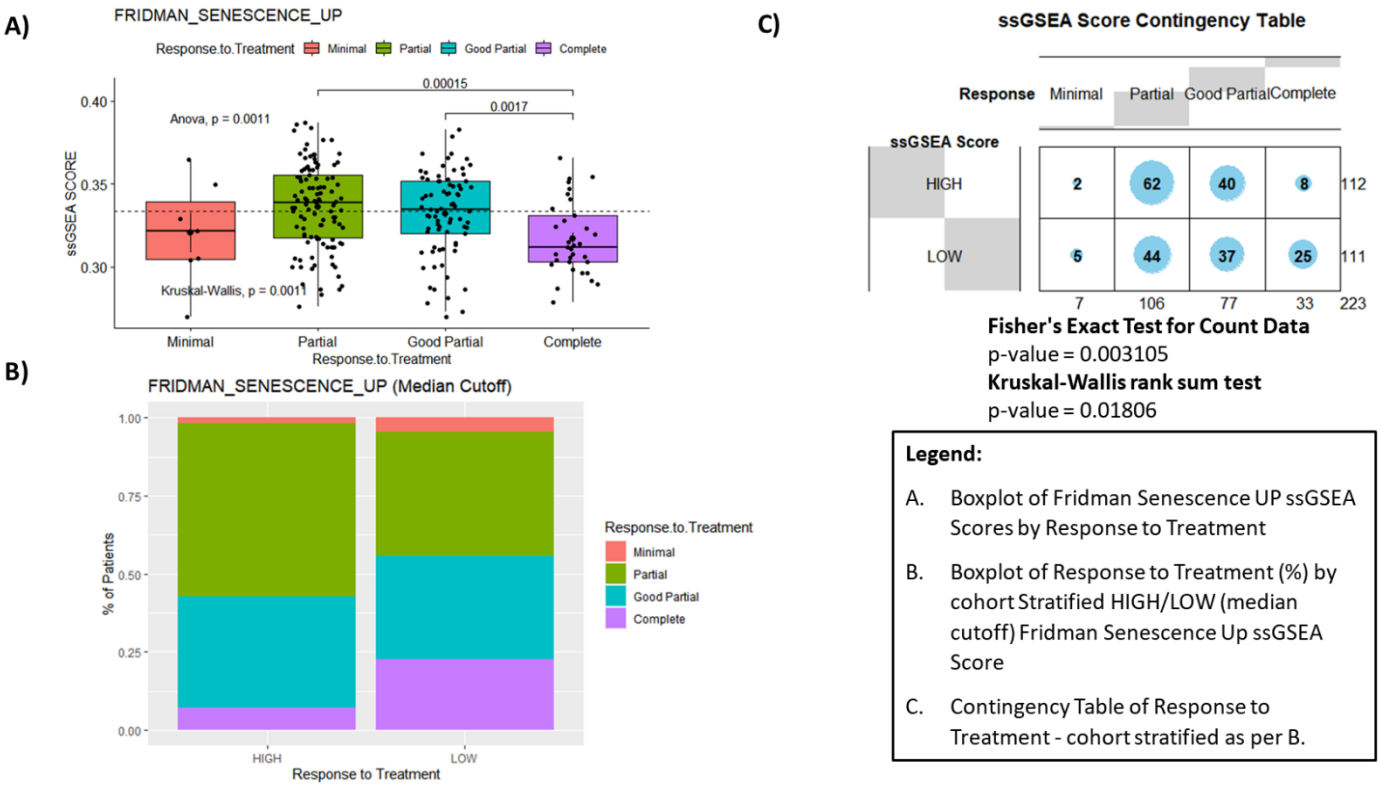


Figure 4.50 Clinical application of Fridman Senescence Up ssGSEA score in SCORT Cohort

Panel A demonstrates significant difference in ssGSEA scores between response to treatment categories by Anova (p = 0.0011) Panel B demonstrates % of response to treatment in high/low (median cutoff) senescence ssGSEA score. Panel C demonstrates contingency table between response and high/low senescence ssGSEA cohorts Kruksal-Wallis rank sum test (pvalue = 0.018)

To further explore the prognostic features of a senescence signature expression, the dataset was further reduced to only include patients with full clinical data. The patient characteristics are listed in Figure 4.51, the single T1 Stage patient has been incorporated into Stage 2. The response to treatment outcomes were binary categorised to incomplete or complete response and a regression model used with clinical characteristics; T stage, N Stage and the High/Low Fridman Senescence Up ssGSEA score expression cohort (median cuttoff).

	Overall (N=158)
T	
2	20 (12.7%)
3	127 (80.4%)
4	11 (7.0%)
NODE	
0	66 (41.8%)
1	66 (41.8%)
2	26 (16.5%)

Figure 4.51 Clinical Characteristics of Cohort with Full Clinical Data
T = T-Stage, Node = Nodal status

The odds ratios of having a high senescence ssGSEA score (median cutoff) are demonstrated in Figure 4.52, and demonstrate that patients with high senescence ssGSEA scores have statistically significantly (pvalue = 0.02) higher odds ratio of having an incomplete response to treatment at 3.02 (1.23-8.09).

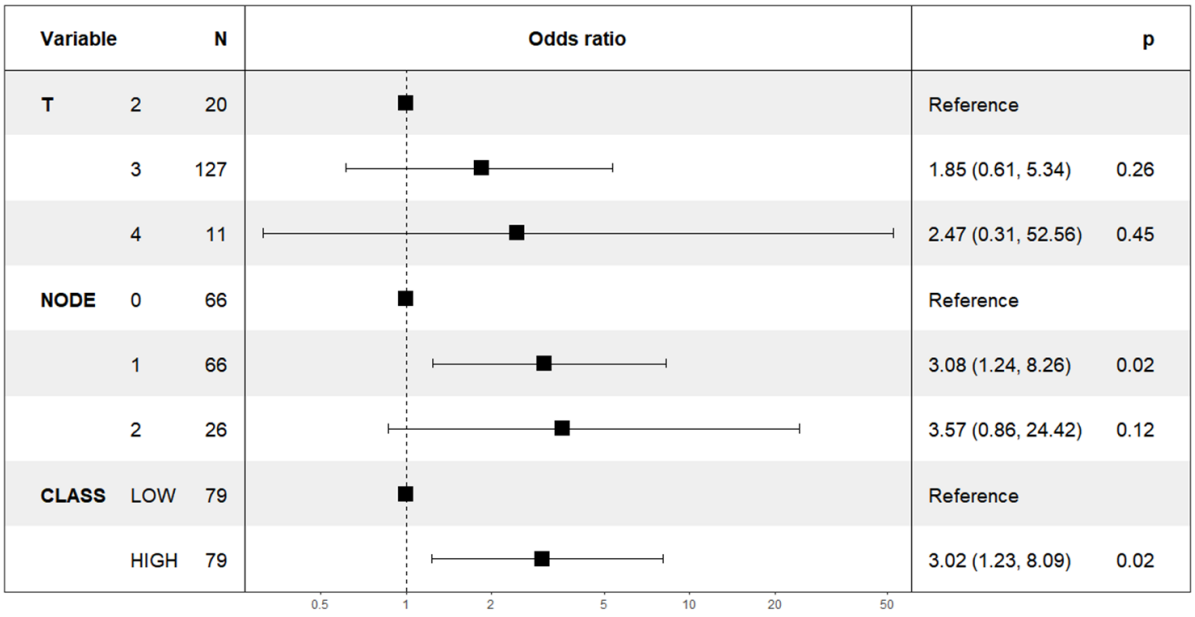


Figure 4.52 Odds Ratios of Incomplete response to treatment
T demonstrates T-Stage, NODE demonstrates nodal status, CLASS demonstrates high or low senescence ssGSEA score (median cutoff), High expression Odds Ratio=3 (p val 0.02)

To further identify whether there was prognostic benefit in other Radiotherapy treated cohorts the expression of the Fridman Senescence Up pathway was examined in a localised

Prostate Cancer Radiotherapy treated cohort and a further Rectal Cancer cohort treated with neo-adjuvant radiotherapy. The package SigCheck was used to generate survival curves for high and low expression cohorts (cohort divided into high and low by median expression of the pathway genes) in each dataset (Figure 4.53). The senescence up signature is prognostic for biochemical progression free survival in the Prostate cohort (p=0.021), with patients in the high expression cohort demonstrating reduced biochemical progression free survival, this was not translated into meaningful metastatic progression free survival. Within the rectal cancer cohort GSE87211, the higher expression group can be seen to have worse metastatic progression free survival, however this does not reach statistical significance. (p = 0.064). In both of these cohorts and the SCORT cohort, it can be noted that the cohorts with higher expression of the Fridman Senescence Up pathway, have worse clinical outcomes following radiotherapy.

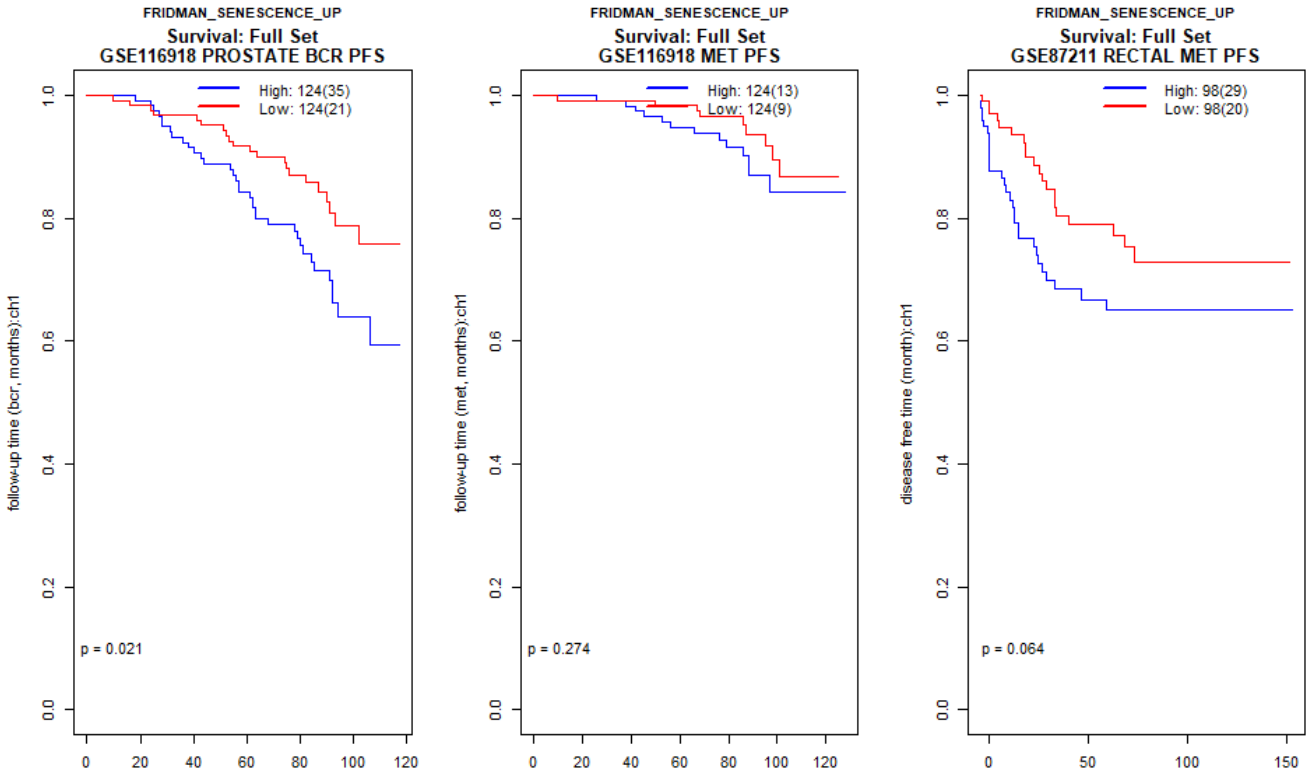


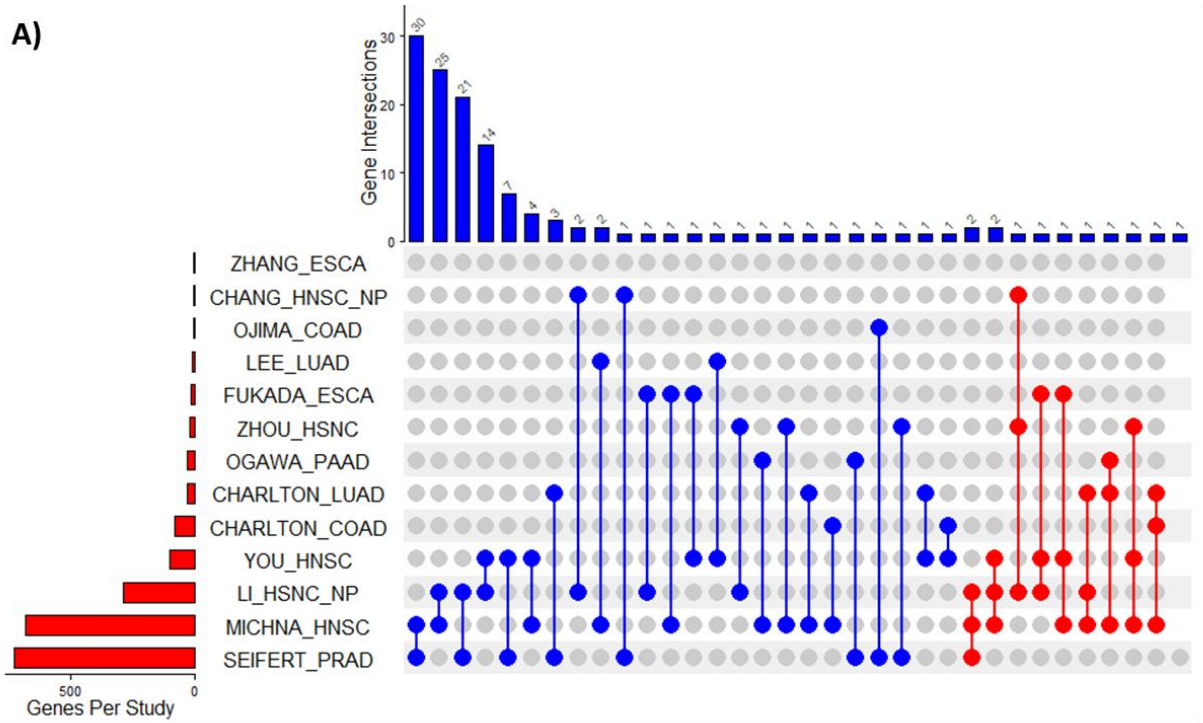
Figure 4.53 Survival Curves for Fridman Senescence Up pathway in external datasets (GSE116918 and GSE87211)

demonstrates Kaplan-Meier survival analysis of high or low senescence ssGSEA cohorts (median cutoff) in; Left Panel -Biochemical Recurrence progression free survival in GSE116918 (p = 0.021), Middle Panel -Metastatic progression free survival in GSE116918 (p = 0.274) , Right Panel- Metastatic progression free survival in GSE87211

4.31 Updated analysis of radioresistant cell line gene expression overlaps

The previously performed analysis to identify overlapping genes from experiments where radioresistant cell lines generated, was updated to include the performed cell line experimental results for HT29 and H460 cell lines.

A)



Histology Key:

COAD - Colorectal Adenocarcinoma

ESCA - Oesophageal Adenocarcinoma

HNSC - Head & Neck Squamous Cell Carcinoma

HNSC_NP - Nasopharyngeal Squamous Cell Carcinoma

LUAD - Non-small Cell Lung adenocarcinoma

PAAD - Pancreatic Adenocarcinoma

PRAD – Prostate Adenocarcinoma

B)

GENE	ENSEMBL ID	description
HLA-B	ENSG00000234745	major histocompatibility complex, class I, B [Source:HGNC Symbol;Acc:HGNC:4932]
FLNB	ENSG00000136068	filamin B [Source:HGNC Symbol;Acc:HGNC:3755]
AREG	ENSG00000109321	amphiregulin [Source:HGNC Symbol;Acc:HGNC:651]
PHLDB2	ENSG00000144824	pleckstrin homology like domain family B member 2 [Source:HGNC Symbol;Acc:HGNC:29573]
IL6	ENSG00000136244	interleukin 6 [Source:HGNC Symbol;Acc:HGNC:6018]
VEGFA	ENSG00000112715	vascular endothelial growth factor A [Source:HGNC Symbol;Acc:HGNC:12680]
ACTA2	ENSG00000107796	actin alpha 2, smooth muscle [Source:HGNC Symbol;Acc:HGNC:130]
PCSK5	ENSG00000099139	proprotein convertase subtilisin/kexin type 5 [Source:HGNC Symbol;Acc:HGNC:8747]
NT5E	ENSG00000135318	5'-nucleotidase ecto [Source:HGNC Symbol;Acc:HGNC:8021]
SGK1	ENSG00000118515	serum/glucocorticoid regulated kinase 1 [Source:HGNC Symbol;Acc:HGNC:10810]
GADD45A	ENSG00000116717	growth arrest and DNA damage inducible alpha [Source:HGNC Symbol;Acc:HGNC:4095]

Legend:

A. Upset Plot – Overlapping genes increased in generated Radioresistant lines at Baseline.

B. Genelist of Overlapping genes increased Radioresistant vs Parent in 3 Studies

Figure 4.54 Updated Upset Plot of Radioresistant Cell Line experiment overlaps

Upset Plot demonstrating overlap between gene expression studies with genes increased in radioresistant cells on the left and genes decreased on the right. The y axis barplot shows the number of genes per study. The x axis barplot shows the number of overlapping genes. The central panel demonstrates which studies contain overlapping genes overlapping 2 studies in blue and genes overlapping 3 studies in red

The 11 overlapping genes that are identified in 3 studies as being increased in radioresistant cells were explored in the SCORT dataset (Figure 4.54). The mean expression of these 11 genes was calculated per patient, denoted as META_scaled_score henceforth. The median score of patients per group was calculated and found to be lower in patients who had a complete response than those who had a good partial (p = 0.03) and patients who had a complete response and a partial response (p = 0.0017). When the cohort was divided into 2 by median expression there were a higher proportion of complete responders in the lower expression cohort than the higher expression cohort, 26 vs 7 respectively.

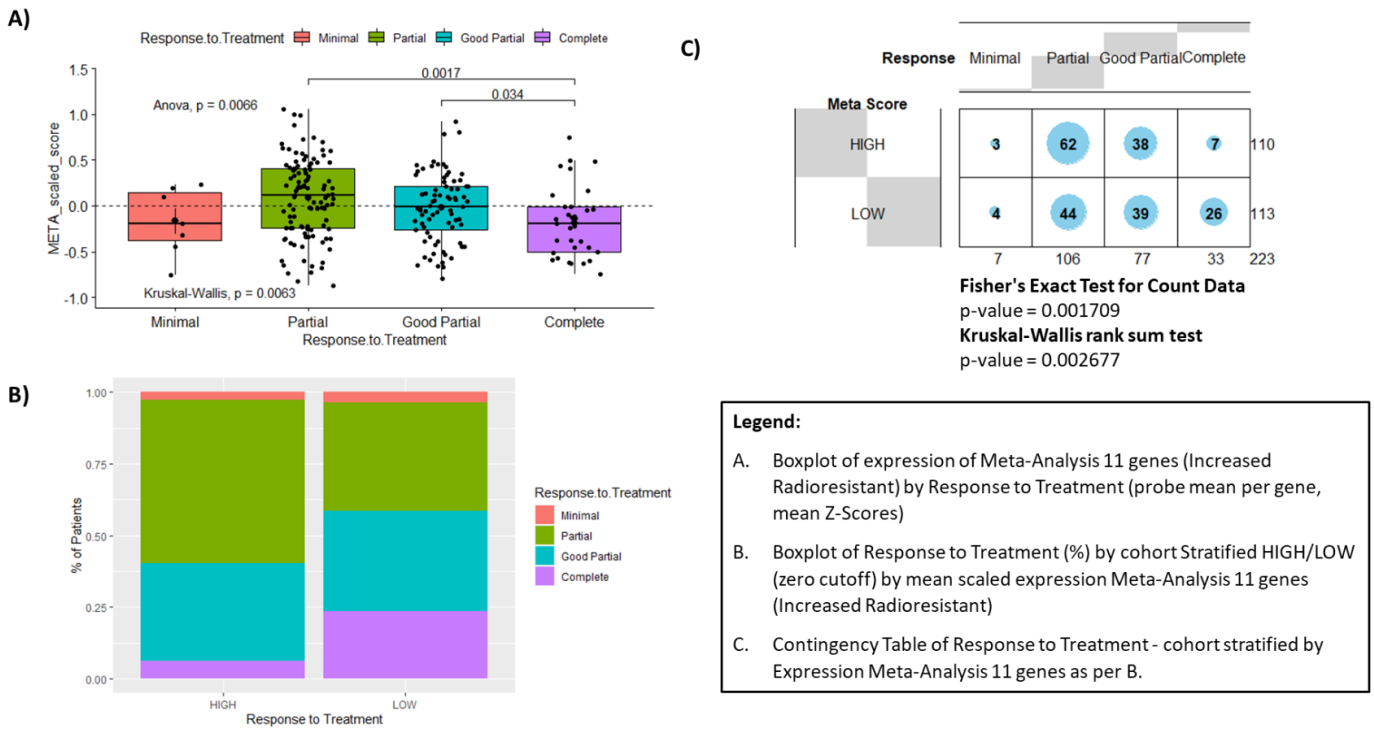


Figure 4.55 Clinical Application SCORT dataset:
Panel A demonstrates significant difference in mean expression of 11 genes between response to treatment categories by Anova (p = 0.0066) Panel B demonstrates % of response to treatment in high/low (median cutoff) senescence ssGSEA score. Panel C demonstrates contingency table between response and high/low senescence ssGSEA cohorts Kruksal-Wallis rank sum test (pvalue = 0.0017)

4.32 Chapter 4 Summary

The initial PCA comparison (Figure 4.1) demonstrates that radioresistant cell lines cluster together with parental cell lines across all timepoints, indicating that the most substantial differences between samples relate to their intrinsic cell of origin type. H460 cell lines separated into parental and resistant by the second Principal component, demonstrating that there are intrinsic differences in gene expression between parental and resistant cell lines (Figure 4.4). On differential expression analysis between resistant and parental line 55 genes were significantly decreased (log fold change < -1) in radioresistant samples and 257 genes significantly increased (log fold change >1) with an adjusted pvalue of <0.05. Figure 4.14 demonstrates a core group of 52 genes that were consistently significantly (padj <0.05) increased (LFC >1) in radioresistant samples compared to parental samples across all timepoints, 11 genes were found to be significantly decreased (LFC < -1). These genes were associated with pathways associated with catabolic pathways and g-protein receptor pathways on over-representation analysis (Figure 4.15). The 628 genes exclusively significantly increased in radioresistant cells 24 hours after radiation, identified pathways associated with response to Endoplasmic reticulum stress, Ras signalling transduction, macro-autophagy and autophagy as being significantly enriched, suggesting that these pathways are upregulated in the radioresistant cells following repeat irradiation (Figures 4.18, 4.19). The 339 genes decreased at 24 hours after radiation in radioresistant cells identifies pathways that are significantly enriched for including DNA replication, ribosome biogenesis, double strand break repair via homologous recombination and RNA splicing, suggesting these pathways are downregulated in resistant cells following radiation (Figure 4.19, 4.20). To account for differential expression over time, a time interaction model was utilised, 126 genes were increased (LFC>1) in radioresistant samples, were significant at padj <0.01 and 26 decreased (LFC <1) (Figure 4.21). Gene Set Enrichment analysis of these genes demonstrated Cell death signalling and p75 NTR receptor-mediated signalling

as activated and G2/M checkpoints, DNA replication and Mitotic G2-G2/M phases as suppressed (Figure 4.23). Geneset enrichment analysis comparing Time 24 to baseline identified pathways including Fridman Senescence UP and Yamashita Methylated in Prostate Cancer (Figure 4.27) as being significantly increased 24 hours after repeat radiation. To further investigate this a previously published geneset demonstrated to be reduced in expression in senescent cells is reduced in radioresistant cells on single sample geneset enrichment analysis (padj < 0.01) as demonstrated at 24 hours in the heatmap in Figure 4.28.

Methylation Principle Component Analysis demonstrated similar findings to RNA sequencing with clustering by cell type and H460 parental and resistant cell lines separated by the second Principle component (Figure 4.35). The top differentially methylated region (Figure 4.36, 4.37) corresponds to the gene FAM83A, which is also one of the top differentially expressed genes on RNA sequencing (Figure 4.27). Correlation analysis (Spearman) of FAM83A RNA gene expression was performed and GSEA identified pathways including Cell Cycle, Cellular Senescence, and p53 pathway as being significantly enriched (padj <0.05) in the ranked gene list (Appendix 2.9/2.10). Pathway enrichment of differentially methylated probes identified TP53 targets, SUZ12 targets, PRC2 targets, multiple pathways associated with H3K27ME3 as being both enriched in hypo- and hypermethylated probes. (Figure 4.38, 4.39).

Expression of genes increased in radioresistant lung cancer cell lines in patients treated with radiotherapy in the TCGA Lung Cancer dataset demonstrated that the cohort with lower expression of radioresistant genes had a better survival HR 0.45 (0.21-0.99), p value of 0.048, this was not replicated in non-radiotherapy treated patients. Senescence associated geneset Fridman Senescence UP, was demonstrated to be significantly differentially expressed on ssGSEA analysis 24 hours following radiation (Figure 4.27), and prognostic in the SCORT rectal dataset (Figure 4.52) with patients with high senescence ssGSEA scores have statistically significantly (pvalue = 0.02) higher odds ratio of having an incomplete response to radiotherapy treatment (3.02, 1.23-8.09).

4.33 Chapter 4 Discussion

The finding of the initial PCA comparison (Figure 4.1) demonstrating that radioresistant cell lines cluster together with parental cell lines, suggests that the main gene expression differences between (non-isogenic) cell lines of differing radioresistance are due to intrinsic gene expression differences in cell type. This finding provides evidence that previous gene expression comparisons in the literature where different cell types measured by SF2, are compared to infer genes related to radioresistance[64-98], are likely to provide minimally informative results. Whilst the differentially expressed genes identified in these studies may include genes of relevance to radioresistance, the majority will not be, otherwise the PCA findings would identify a clear separation between all radioresistant and parental cell lines. This finding further supports that the methodology of generating resistant cell lines and comparing to parental, utilised in Chapter 3 is the most likely to identify candidate radioresistance genes, and helps to explain the lack of coherence of published gene expression study findings in the literature (Table 1.2).

The finding that H460 cell lines separate into parental and resistant (Figure 4.4) by the second Principal component, demonstrates that there are intrinsic differences in gene expression between parental and resistant cell lines, it is interesting to note that this was not by the first Principal Component, suggesting that other differences between cell samples may have a greater impact on gene expression.

The finding that pathways associated with response to Endoplasmic reticulum stress, Ras signalling transduction, macroautophagy and autophagy pathways are enriched for in the set of genes exclusively upregulated by radioresistant cells 24 hours following repeat irradiation, (Figures 4.18, 4.19) suggests that these are responses unique to the radioresistant cells. Autophagy is a known biological response toward stress including ionizing radiation, and the successful execution of autophagy appears to facilitate cell survival[159]. Several studies have

demonstrated that autophagy inhibitors in addition to IR enhances radiosensitization in numerous cell lines and that conversely an autophagy inducer increased cell proliferation[160]. Enrichment analysis of genes exclusively decreased in radioresistant cells, identified pathways associated with DNA replication, ribosome biogenesis, double strand break repair via homologous recombination and RNA splicing (Figure 4.19, 4.20). Interestingly Figure 4.27 demonstrates via ssGSEA that pathways associated with IR response are similarly increased in both resistant and parental cell lines over time. The results in Chapter 2 of cell cycle analysis previously identified a greater mean percentage of cells were in the more radioresistant G2 phase in radioresistant lines than parental, 11.5% vs 9.3% (Table 3.6) at baseline (pre-radiation) and 24 hours post further 2Gy irradiation (21.2% and 24.4% vs 14.2%).

Single Sample Geneset enrichment analysis also demonstrated that a Senescence pathway from the MSigDB curated gene sets (Fridman Senescence Up) and another transcriptome signature of cellular senescence[161] is differentially expressed between the parental and radioresistant lines (padj = 0.01), clearly demonstrated at Time 24 (Figure 4.26, and heatmap Appendix 2.5). Senescence is known to be induced by ionising radiation and believed to confer resistance to therapy via a number of mechanisms including epigenetic modification via chromatin modulation and is associated with immune stimulation following DNA damage in the secretory phenotype[149-151]. Whilst clinically important the role of immune stimulation is complex and not reflected in cell lines experiments but it is notable that the class 1 Major Histocompatibility (MHC) presentation genes HLA-B, and HLA-C are some of the most significantly (padj < 0.01) differentially expressed genes that are increased in radioresistant cell lines (Figure 4.21).

The Senescence Associated Secretory phenotype (SASP) is understood to mediate many of the effects of senescence and plays a significant role[149-150]. One of the most overexpressed genes in senescent cells compared to parental cells was Amphiregulin (AREG), (Figure 4.29), known to be a low-affinity EGFR ligand. AREG is frequently secreted by senescent cells and has previously been implicated as having a role in the regulation of senescence in

melanocytes[156]. A recent quantitative proteomic study identifies AREG as one of the most abundant proteins in Oncogene induced Senescence and further demonstrates that AREG promotes in vitro reprogramming by accelerating proliferation and the mesenchymal-epithelial transition via EGFR signalling[162].

One of the recent key findings is that senescence, which was previously considered to be a permanent process, can be dynamic. The H3K9-active demethylases JMJD2C and LSD1 have been demonstrated to mediate escape from oncogene induced senescence in both fibroblast and melanocyte models[163], the H3K9 histone methyltransferase Suv39h1 or p53 is required for senescence, specifically for the maintenance of therapy-induced senescence (TIS) in various malignancies[155].

The top Differentially Methylated Region (Figure 4.37) between radioresistant and parental cell lines was FAM83A, also found to be one of the top differentially expressed genes (Figure 4.27) on RNAseq analysis. Previously demonstrated to be prognostic in lung cancer, FAM83A promotes the proliferative and invasive abilities of cancer cells via epithelial-mesenchymal transition and the Wnt signalling pathway[164]. FAM83A has recently been demonstrated to bind directly to β-catenin and inhibit its degradation. Furthermore, FAM83A has also been demonstrated to enhance oncogenic Wnt/β-catenin-mediated transcription promoting β-catenin-TCF4 interaction[165]. TCF4 was also noted to be one of the top hypomethylated genes between radioresistant and parental lines in common between both cell lines (H460 and HT29) used in the methylation analysis (Figure 4.41). Further exploration of pathways associated with FAM83A using per gene correlations as a ranking metric for the GSEA algorithm to identify significantly associated pathways identified pathways including Cell Cycle, Cellular Senescence, and p53 pathway, suggesting that FAM83A may be associated with multiple downstream effects (Appendix 2.9). However single gene correlative analysis provides only association not causation.

The most differentially methylated pathways between radioresistant and parental cell lines involved pathways associated with trimethylation of histone H3 on lysine 27 (H3K27me3), (Figures 4.38, 4.39). Trimethylation of histone H3 on lysine 27 (H3K27me3) is associated

with chromatin condensation, affecting Double Strand Break repair, and studies have demonstrated the effects of radiation on H3K27me3 levels in tumour and normal cell lines[166], identifying a rapid decrease following irradiation. Silencing the H3K27 demethylase UTX using an siRNA enhanced radiosensitivity as did treatment with a H3K27 demethylase inhibitor[166] suggesting that the differential methylation of H3K27me3 pathways may be contributing towards radioresistance.

Taken together the above results in Chapter 4 demonstrate that I have achieved the second aim of this thesis to identify molecular features associated with a radioresistant phenotype through gene expression and methylation analysis.

The refined geneset increased in expression in radioresistant H460 cells, (Figure 4.26) was demonstrated to be prognostic in Radiotherapy treated lung cancer patients from the TCGA. This demonstrated the cohort with lower expression of radioresistant genes had a better survival HR 0.45 (0.21-0.99) which was statistically significant at a pvalue of 0.048 (Figure 4.48) but this was not significant in the non-radiotherapy treated cohort (Figure 4.50). This suggests that there is potential for the differentially expressed genes in the analysis to predict for survival outcomes in lung cancer radiotherapy treated patients, however the size of the radiotherapy treated dataset (n=61) precludes any robust conclusions to be drawn and further studies in larger cohorts would be required. There was not a strong correlation between genes in the TCGA data suggesting that they are not a coherently expressed group of genes, which could limit potential application as a predictive signature. (Appendix 2.11) The Senescence associated geneset Fridman Senescence UP, was demonstrated to be significantly differentially expressed on ssGSEA analysis 24 hours following radiation (Figure 4.26). In the SCORT rectal dataset (Figure 4.53) with patients with high senescence ssGSEA scores had a statistically significantly (pvalue = 0.02) higher odds ratio of having an incomplete response to treatment (3.02, 1.23-8.09). This signature was also demonstrated to

be prognostic in a radiotherapy treated Prostate Cancer cohort (GSE11691) and a non-significant (0.064) difference in Metastatic Progression free survival in a further rectal radiotherapy treated cohort (Figure 4.24). These findings suggest that a transcriptomic signature of Senescence may have potential clinical utility in predicting response to radiotherapy across tumour types, which warrants further investigation in other large scale clinical datasets. The results contained in Chapter 4 demonstrate that I have achieved the third aim of this thesis, to identify the predictive value of features associated with radioresistant cell line phenotype in clinical cohorts.

Although the results of this Chapter show interesting results there are some clear limitations of the experiment. Notably the cell line experiment utilised a very small number of cell lines, with only 3 cell lines used to generate radioresistant cell lines, and only 2 cell lines contributing towards gene expression and methylation studies. This limitation could mean that the wider extrapolation of findings may be limited due to the small sample size. The cell line experiment as discussed in Chapter 3 utilised 2-dimensional monoculture which does not accurately replicate the complex 3-dimensional structure of tumours and potential for nutrient and oxygen gradients which could affect response to radiation.
Similarly, although the findings of gene expression studies were demonstrated to be prognostic in a cohort of radiotherapy treated patients from the TCGA Lung Adenocarcinoma dataset (Figure 4.48) and not in non-radiotherapy a treated cohort (Figure 4.50), these findings need careful consideration. The number of patients treated with radiation was small (n=61), and there are no provided details about the radiation schedule in terms of dose, fractionation and intent. For example, Figure 4.42 demonstrates that 6 of the patients treated with radiation had metastatic disease at outset (Stage IV) and radiotherapy in this situation would not usually be delivered with curative intent. A more formal evaluation of the prognostic power of the candidate genes in a well curated dataset, ideally from a retrospective or prospective trial would be necessary to evaluate the potential for clinical implementation.

The application of a Senescence Signature by Fridman[167] to the SCORT dataset does utilise well curated trial data for evaluation. One limitation of the type of analysis performed in this analysis is that it is reliant on categorisation of response to treatment, rather than more readily quantifiable endpoint such as overall survival. One potential critique of such an analysis could be that the categorisation of radiological response as an endpoint could be considered to be less robust than time dependent outcome data such as overall or progression free survival. To further explore the prognostic benefit of

In order to place the results of the radioresistant cell line gene expression experiments into the context of the established literature, the previous computation of gene overlaps performed in the Introduction (Figure 1.1) was repeated to include the results of the cell line experiments performed in Chapter 4 (Figure 4.54). As previously discussed, the PCA plot (Figure 4.1) demonstrates that cell type rather than radioresistant phenotype is the main driver of gene expression differences between cell lines so the repeat computation was again performed only with the studies which generated radioresistant cell lines to compare to parental. The result of this comparison has identified a set of 11 genes which are increased in generated radioresistant cell lines in at least 3 separate studies, and it can be noted from the upset plot (Figure 4.54) that for each of these overlaps they are from a range of histological cell types. Interestingly these results, despite being performed in cell-lines, demonstrate a number of genes which would likely have a wider role *in-vivo*. The class 1 Major Histocompatibility (MHC) presentation gene HLA-B displays peptides to the immune system. The gene product of NT5E can act as an inhibitory immune checkpoint molecule[169]. Interleukin-6 (IL6) is a pro-inflammatory cytokine with known roles in inflammation, autoimmunity and cancer[170]. Vascular endothelial growth factor A has an intrinsic role in angiogenesis, the receptor of which is targeted by the monoclonal antibody cancer treatment Bevacizumab[171]. As previously outlined AREG is associated with the SASP and promotes *in-*

vitro reprogramming by accelerating proliferation and the mesenchymal-epithelial transition via EGFR signalling[156].

These results suggest that accepting the limitation of this experiment being cell-line based when taken into context with previously published results in the literature, it may have the potential to identify mechanisms of radioresistance that warrant further exploration *in vivo*.

Within the literature there are a number of gene expression classifiers which have been applied to lung cancer, predominantly derived from resected lung cancer specimens to derive the benefit of adjuvant therapy[172]. Dehing-Oberije utilised blood biomarkers, including lactate dehydrogenase, C-reactive protein, osteopontin, carbonic anhydrase IX, interleukin (IL) 6, IL-8, carcinoembryonic antigen (CEA) to develop a multivariate model based on a large population of radiotherapy treated patients (n = 322)[173]. Utilising a training and validation approach they developed a model incorporating sex, performance status, forced expiratory volume, number of positive lymph node stations, tumour volume and two blood biomarkers (CEA and IL-6) which was demonstrated to be prognostic with an Area Under the Curve (AUC) of 0.81.

Peinado-Serrano identified candidate genes from a panel of lung cancer cell lines and performed univariate Cox regression to identify genes with an impact on survival in patients with either Adenocarcinoma or Squamous Cell carcinoma in the TCGA dataset[174]. They performed multiple Cox regression with cross-validation in which two thirds of the cohort were used for training and the remaining one third for validation. They identified six gene signatures predictive for overall survival and progression-free survival (APOBEC3B, GOLM1, FAM117A, KCNQ1OT1, PCDHB2 and USP43). There is no overlap between identified genes in this study and the genes utilised in the validation performed in this study (Figure 4.26). One of the potential confounders of this study is the use of two separate clinical entities Lung Squamous Cell carcinoma and Lung adenocarcinoma to generate a homogenous signature. Also given the same small dataset (n = 107) was utilised for both training and validation the applicability of these results to larger datasets would need to be

validated. One further critique would be that the authors did not investigate the signature in the non-radiotherapy treated cohort, as they may have identified a prognostic signature. The other main study by Scott was GARD (genome-based model for adjusting radiotherapy dose), which combines the information derived from the radiosensitivity index (RSI) and the linear quadratic model. Scott demonstrates variation in radiation response not only between different tumour types but also heterogeneity within tumour types with clinical data from 8,271 patients across 20 tumour sites[12]. A numerical value for GARD was calculated with greater therapeutic effect of radiotherapy treatment seen with higher GARD scores, however this was not predictive of survival. Yet there has been no prospective validation of GARD and a critique of the validity of this approach, which relies on the assumption of an interaction between RSI and dose, identifying that the addition of an interaction term between RSI and dose does not explain more of the outcome variance, as would be expected[56].

The significance of the work outlined in this chapter is that it outlines molecular alterations in radioresistant cell-lines following clinically relevant fractionated radiation. These changes are seen not only at baseline but are demonstrated to be dynamic in terms of differing responses following further irradiation, complementing the cell line characterisation performed in Chapter 3. These results could be further developed to explore predictive signatures of radiotherapy outcome in clinical cohorts either through the identified genes or senescence signature and may identify potential therapeutic targets for further investigation e.g. FAM83A in order to augment the response to radiation in an effort to mitigate the development of radioresistance.

CHAPTER 5

5.1 Prostate cancer radical radiotherapy clinical dataset

5.11 Chapter 5 Introduction and Aims:

Localised Prostate Cancer is treated with curative intent by radical surgery, or radical radiotherapy with androgen-deprivation therapy (ADT), with equivalent rates of post-treatment disease progression and overall survival (~97%) at 15-years clinical follow-up [175,176]. Radical radiotherapy is the preferred treatment options for the elderly, comorbid or those with more advanced disease. It is well recognized that with prolonged follow-up a sub-group of patients experience disease progression following radical radiotherapy[176]. High Risk groups for recurrence are defined by clinico-pathological features including, Prostate Specific Antigen (PSA) greater than 20 ng/ml, tumour grade (Gleason Grade Group from 1-5 with 5 having the most aggressive phenotype) and radiological staging (>T3 disease spread beyond the capsule of the prostate)[177,178]. In addition to clinical and pathological features, there is great potential for molecular features from diagnostic biopsies to offer additional information to identify patients who are at a higher risk of relapse following radical radiotherapy. Utilising molecular characterisation could facilitate more accurate pre-treatment risk-stratification, which in turn would allow more appropriate treatment selection through intensification of dose, more extensive treatment of pelvic nodal regions, prolongation of hormone therapy duration, or increased clinical follow-up.

Unlike breast cancer, no prognostic and predictive molecular classifiers are currently routinely used in the clinic in the radical radiotherapy setting in the UK, although several have been developed for Prostate Cancer. Classifiers include the Oncotype, Decipher and

Prolaris[179-182], and metastatic and hypoxia signatures[183,184]. The Decipher and cell cycle progression (CCP) signatures have demonstrated clinical utility in predicting disease recurrence after both radical surgery and radiotherapy [185-190]. These classifiers were originally developed from prostatectomy surgical specimens so they may not be the optimal classifier for patients treated with radical radiotherapy as they are derived from patients treated with a completely different treatment modality. A signature of hypoxia which is prostate cancer specific has been developed from a previous hypoxia metagene and demonstrated to predict biochemical recurrence in the salvage radical radiotherapy setting, where patients develop local disease recurrence following radical prostatectomy. This signature is also found to be independently prognostic for radical radiotherapy treated patients with localised Prostate Cancer[191]. Methylation-based molecular features of Prostate Cancer have also been found to be associated with clinical outcome after salvage radical radiotherapy[192-195].

The clinical utility of a metastatic signature which was also derived from surgical specimens has been demonstrated in a cohort of patients by Jain with localised Prostate Cancer treated with radical radiotherapy which used microarray sequencing technology from diagnostic prostate biopsy samples[196]. Using prostate cancer diagnostic biopsies samples for RNA sequencing is technically challenging given the limited and often very small volume of tissue available. The Formalin Fixation Paraffin Embedding process which is a part of routine clinical processing also causes degradation of the samples because of cross-linking of RNA, and fragmentation over time can make RNA sequencing challenging.

3’ RNA sequencing (3’RNAseq) is potentially well suited for sequencing degraded RNA as it only utilised the 3’ ends of the RNA fragments. The 3’ end which is adjacent to the poly-A tail is more stable and less likely to undergo degradation over time. The technology can identify and analyse the last 75 base pairs of the 3’ end of the RNA molecule, which is less likely to be fragmented and simplifies the analysis pathway by providing a single read per gene transcript. This provides an advantage over the standard RNA sequencing techniques which

look at reads throughout the RNA transcript, which are likely to have become fragmented. 3'RNAseq, like other RNAseq technologies, enables multiplexing of samples which provides an economy of sequencing at scale, which makes it more suitable for clinical applications.

NanoString analysis is an alternative technique for molecular analysis of clinical samples, this technique uses reporter probes to hybridise mRNA. The platform reports fewer genes (~600 for nanoString, versus ~ 20,000 for 3'RNAseq) but it significantly reduces computational resources and analysis time as it can be analysed on a standard computer. The proprietary software analysis pipeline allows researchers with no previous bioinformatic expertise to perform the analysis.

Illumina 850k methylation analysis is a recently updated technique which has greater power over previous methylation arrays due to the increased number of probes (850,000). The increased number of probes allows detection of differential methylation at an increased resolution to previous technology. Utilising these distinct molecular analysis technologies orthogonally increases the potential to identify biologically relevant genes compared to utilising a single genomic technique in isolation, as it allows for a degree of cross-validation of findings across technology platforms.

The Aim of this Chapter is to identify features of a radioresistant phenotype from a clinical cohort of patients with prostate cancer treated with radiotherapy and identify their predictive value in an external cohort of prostate cancer patients treated with radiotherapy.

5.12 Clinical Sample Selection

Clinical samples from 19 patients were obtained from a prior clinical study Prostate Cancer Mechanisms of Progression (PROMPT) which had full ethical approval. RNA samples from 18 of the 19 patients were included in this study (n=6 with Stable disease, n=5 with Progressive disease, and n=8 with *de novo* Metastatic, Figure 5.1). One sample failed the 3'RNA sequencing process and was excluded from further analysis.

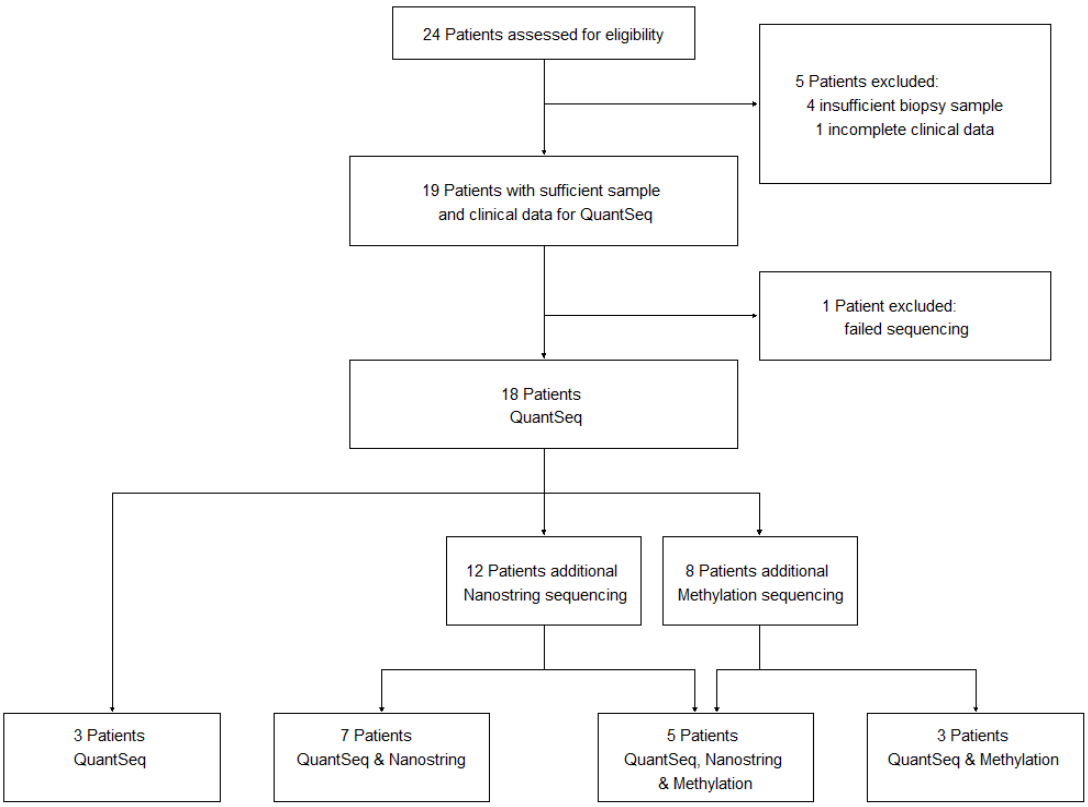


Figure 5.1: CONSORT flow diagram

The baseline clinical and pathological features of the two patient groups receiving radical radiotherapy (i.e. those with Stable disease, and those with Progressive disease) were similar, with a similar mean PSA, and with both groups containing 4 patients with higher-risk disease (Table 5.1). Patients who had progressed post-radiotherapy were noted to have a higher grade group, with 4 patients with grade group 3 and one with grade group 5 disease. Patients who were Stable post-radiotherapy comprised one patient with grade group 1, two with grade group 2, one with grade group 3 and one with grade group 5 disease.

	Stable (N=5)	Progressed (N=5)	Metastatic (N=8)	Overall (N=18)
T Stage				
<=T2	3 (60.0%)	4 (80.0%)	2 (25.0%)	9 (50.0%)
>=T3	2 (40.0%)	1 (20.0%)	6 (75.0%)	9 (50.0%)
N Stage				
N0	5 (100%)	5 (100%)	6 (75.0%)	16 (88.9%)
N1	0 (0%)	0 (0%)	2 (25.0%)	2 (11.1%)
M Stage				
M0	5 (100%)	5 (100%)	0 (0%)	10 (55.6%)
M1	0 (0%)	0 (0%)	8 (100%)	8 (44.4%)
Gleason Score (total)				
6	1 (20.0%)	0 (0%)	0 (0%)	1 (5.6%)
7	3 (60.0%)	4 (80.0%)	1 (12.5%)	8 (44.4%)
8	0 (0%)	0 (0%)	2 (25.0%)	2 (11.1%)
9	1 (20.0%)	1 (20.0%)	5 (62.5%)	7 (38.9%)
Grade Group				
1	1 (20.0%)	0 (0%)	0 (0%)	1 (5.6%)
2	2 (40.0%)	0 (0%)	1 (12.5%)	3 (16.7%)
3	1 (20.0%)	4 (80.0%)	0 (0%)	5 (27.8%)
4	0 (0%)	0 (0%)	2 (25.0%)	2 (11.1%)
5	1 (20.0%)	1 (20.0%)	5 (62.5%)	7 (38.9%)
Prostate Specific Antigen (ng/mL)				
Mean (SD)	21.7 (33.0)	21.9 (23.9)	2270 (5510)	1020 (3720)
Median [Min, Max]	8.70 [1.20, 80.3]	6.40 [2.50, 52.0]	101 [18.7, 15900]	25.3 [1.20, 15900]
Risk Group (D'Amico) / Metastatic				
High	4 (80.0%)	4 (80.0%)	0 (0%)	8 (44.4%)
Intermediate	0 (0%)	1 (20.0%)	0 (0%)	1 (5.6%)
Low	1 (20.0%)	0 (0%)	0 (0%)	1 (5.6%)
Metastatic	0 (0%)	0 (0%)	8 (100%)	8 (44.4%)

Table 5.1. Patient Characteristics Table

Eighteen patient samples underwent 3'RNAseq; 12 underwent nanoString and 8 Methylation. A summary of techniques used per patient, and their associated clinical groups, is included in Figure 5.2.

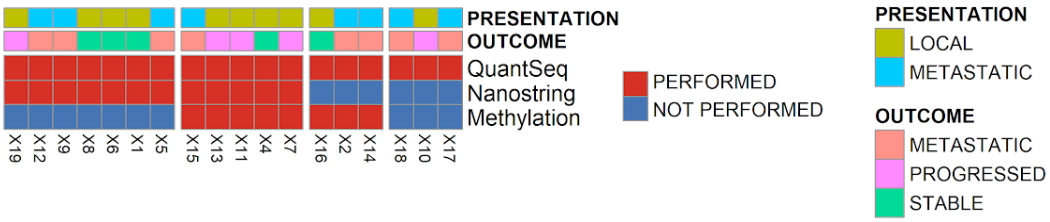


Figure 5.2: Summary of Sequencing Techniques

Heatmap demonstrating sequencing technique per patient, y axis demonstrates technique, x axis demonstrates individual patients, Quantseq = 3' RNA sequencing

Further details of case selection are provided in Appendix Results Table 3.1.

The sigQC tool[20] was used as a quality control tool, following sample sequencing as per Methods, to evaluate the performance of previously validated Prostate Cancer signatures

(prostate hypoxia, Decipher, Prolaris and Oncotype) in the generated 3'RNA sequencing dataset. Rather than comparing the performance of prospective signatures across datasets, the SigQC tool was used to compare the performance of the previously validated signatures in the experimental dataset to performance in a "gold standard" TCGA prostate (pan-cancer) dataset, and a previously published external dataset of patients with localised Prostate Cancer treated with radical radiotherapy by Jain (GSE116918) [97]. The Radar plot in Figure 5.3 of Quality Control metrics demonstrate that these signatures perform at a comparable level in the experimental dataset compared to previously published datasets, providing a validation of the utility of the dataset and indicating the general applicability of 3'RNAseq technology.

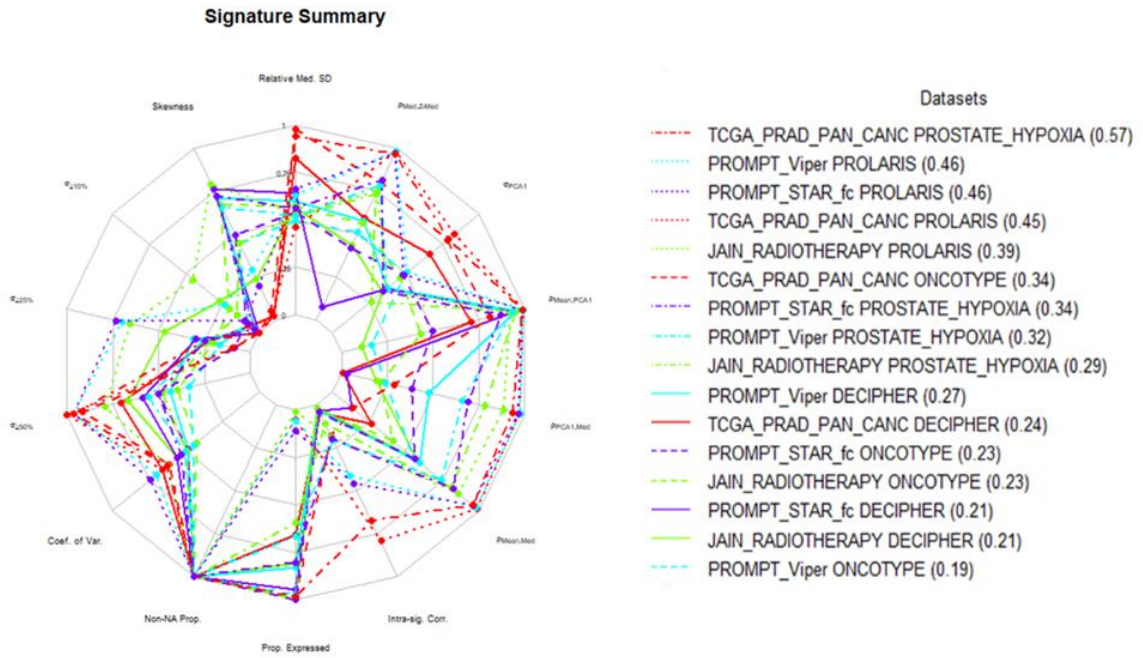


Figure 5.3 Radar Plot of Signature Summary scores in dataset

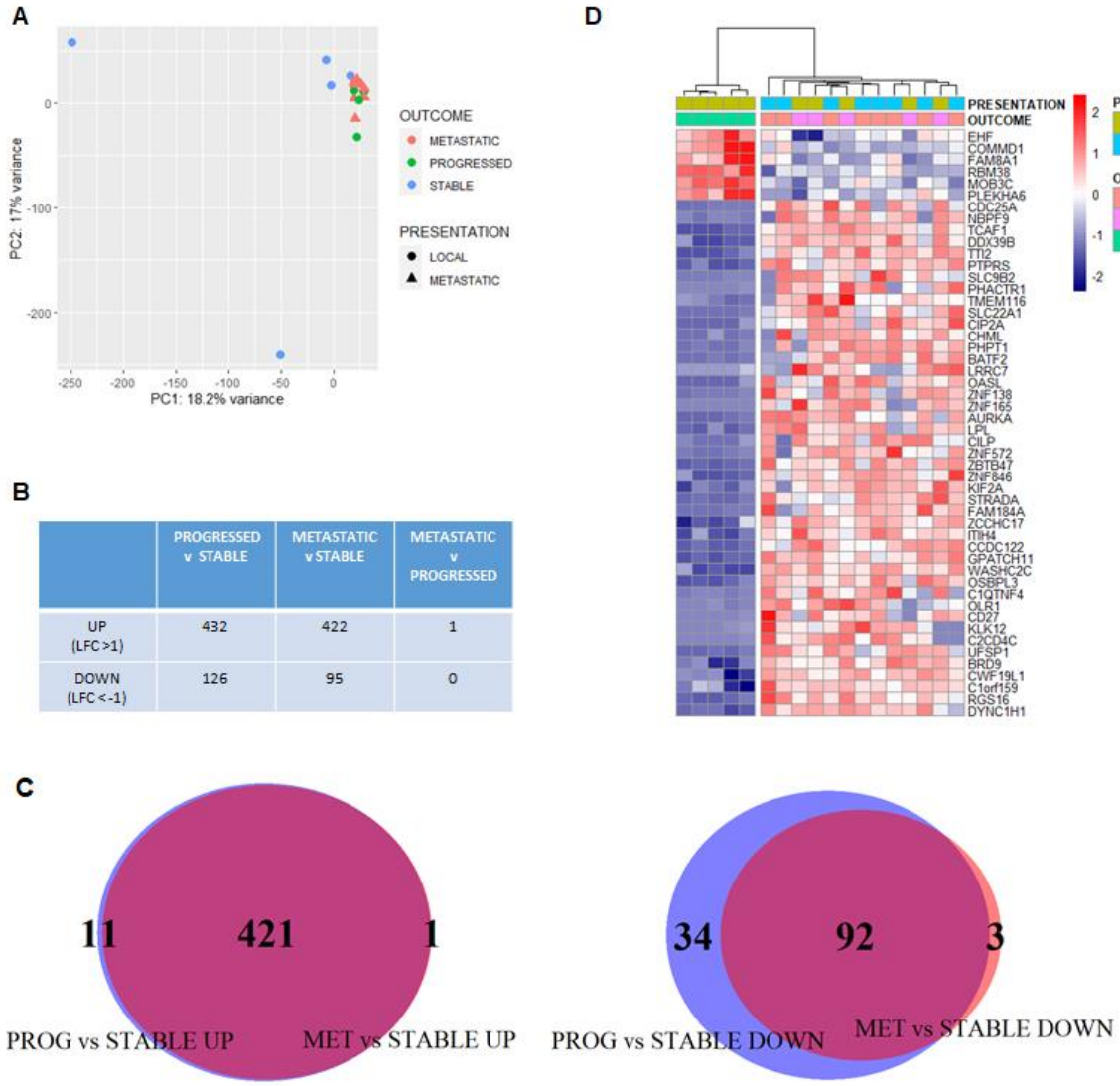
SigQC Signature Summary scores in radar plot for previously validated signatures (Prostate Hypoxia, Prolaris, Oncotype and Decipher) in prostate cancer datasets PROMPT, TCGA Prostate and Jain et al Cohort. Overall signature performance is comparable in the PROMPT dataset See Appendix 3.33 for further expression of signatures

5.13 Differential Gene Expression

Principal Component Analysis demonstrated some separation of the Progressive disease and *de novo* Metastatic cases from the cases with Stable disease post-radical radiotherapy (Figure 5.4A). Differential gene expression analysis to compare Progressive disease post-radiotherapy to Stable disease cases identified 558 significantly differentially expressed genes (adjusted p value (*padj)* < 0.05, Log fold change LFC >1 or < -1). Of these 432 had increased expression, and 126 decreased expression) (Figure 5.4B). A similar number of differentially expressed genes were observed between *de novo* Metastatic cases and those with Stable disease (n=422 increased expression, n=95 decreased expression) (Figure 5.4B).

Only one gene was significantly differentially expressed between Progressive disease and *de novo* Metastatic cases (Figure 5.4B). Most differentially expressed genes between the two comparisons (Stable disease versus Progressive disease and Stable disease versus *de novo* Metastatic) were not only found to be in common, but also differentially expressed in the same direction (n=421 increased expression, n=92 decreased expression) (Figure 5.4C).

Hierarchical clustering of the top 50 differentially expressed protein-coding genes between Stable disease and Progressed disease cases, demonstrates a similar expression pattern between the cases who progressed following radical radiotherapy and those with *de novo* Metastatic disease. The expression of these genes is clearly distinct from the expression in the cases with Stable disease (Figure 5.4D). This supports the findings of principal component analysis, that there are underlying differences in the overall gene expression between the groups.



	PROGRESSED v STABLE	METASTATIC v STABLE	METASTATIC v PROGRESSED
UP (LFC >1)	432	422	1
DOWN (LFC < -1)	126	95	0

Legend:

A) Principal Component Analysis (PCA) of normalised (rlog) gene expression data demonstrates some separation of stable and progressed /metastatic patients.

B) Table showing number of differentially expressed genes between comparisons of Metastatic, Progressed and Stable patients (adjusted pvalue <0.05, logfold change >1/< -1)

C) Venn Diagram showing overlap of differentially expressed genes as per B by direction, increased (left) between Progressed vs Stable patients (PROG_V_STABLE_UP) and Metastatic vs Stable (MET_V_STABLE_UP), and decreased (right) between Progressed vs Stable (PROG_V_STABLE_DOWN) and Metastatic vs Stable (MET_V_STABLE_DOWN) identifying common genes are differentially expressed between groups.

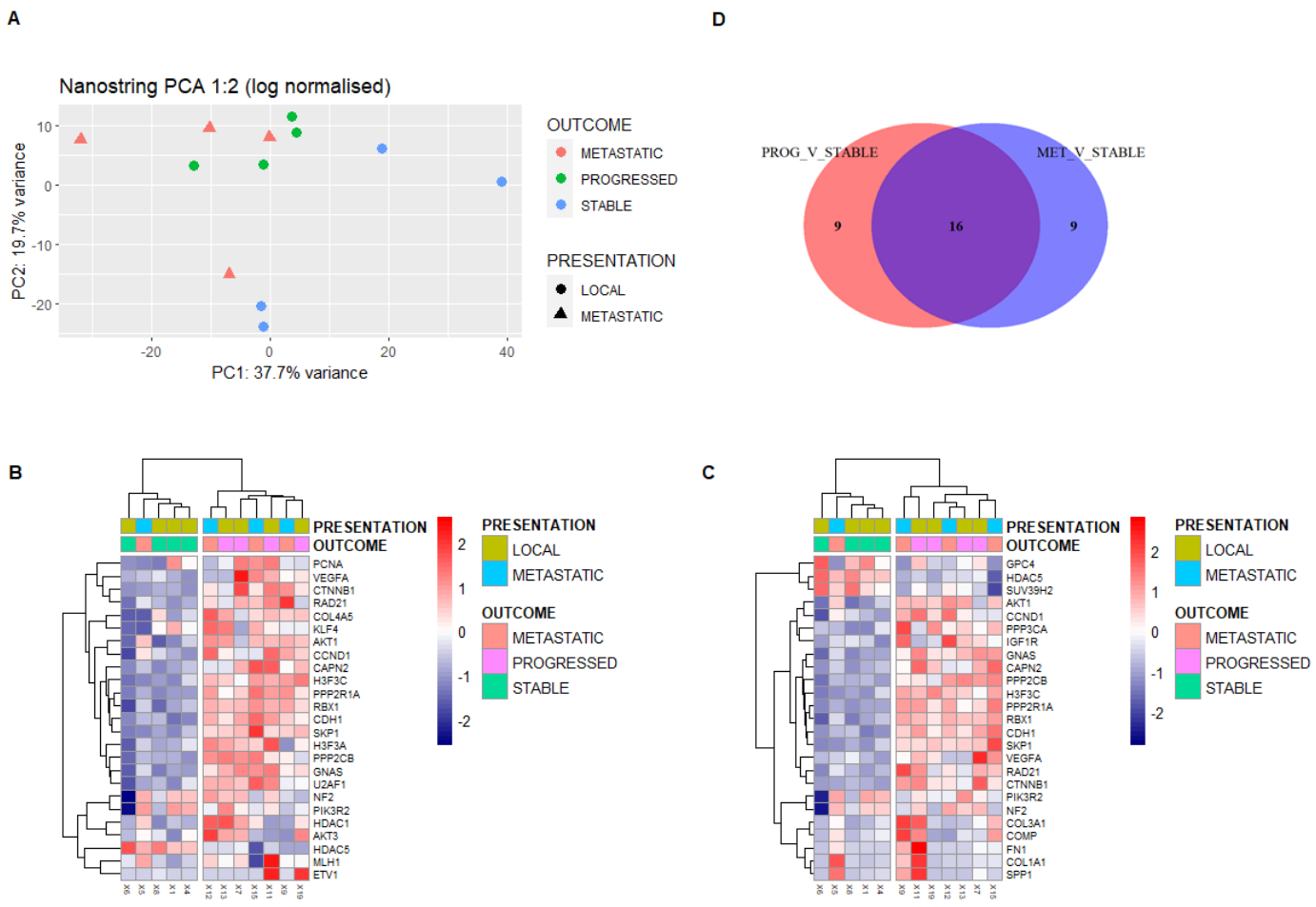
D) Heatmap of top 50 differentially expressed protein coding genes between stable and progressed patients (adjusted pvalue < 0.05). Non-hierachical clustering (ward.D2 method and euclidean distance) demonstrates clustering of progressed and metastatic patients with a similar expression pattern.

Figure 5.4 3'RNAseq differential gene expression analysis

Panel A demonstrates some separation of Stable samples from Progressed/Metastatic, Panel B shows there are minimal gene expression differences between progressed and metastatic samples. Panel C shows that the differences between Stable and Progressed/ Metastatic samples are coherent, Panel D demonstrates this visually.

Over-representation analysis using Gene Ontology (GO) Pathways of the differentially expressed genes increased in cases with Progressed vs Stable disease (padj <0.01), and *de novo* Metastatic disease vs Stable disease (padj <0.05), identified pathways associated with spindle pole and centrosome function respectively (Appendix Figures 3.4-3.5).

NanoString analysis was performed on twelve samples (4 from each of the three groups) and this demonstrated similar differences in gene expression patterns between Stable disease cases and those with either Progressed disease or *de novo* Metastatic disease. A Principal component analysis demonstrated separation of the Stable disease cases from the Progressed disease and *de novo* Metastatic cases (apart from one metastatic case). Similar expression patterns were seen in the top 25 differentially expressed genes using the NSolver analysis pipeline comparing Stable cases and Progressed cases at a pvalue threshold of $p<0.05$, and those identified between Progressed cases and *de novo* Metastatic cases, notably this result was not statistically significant on correction for multiple testing. Hierarchical clustering demonstrated that the Stable disease cases clustered along with one *de novo* Metastatic case using the log normalised nanoString expression data for these top differentially expressed genes (Figure 5.5B). The comparison of Stable disease cases versus *de novo* Metastatic cases identified a similar pattern (Figure 5.5C). The majority of the differentially expressed genes were overlapping with 16 of the top 25 differentially expressed genes (64%) in common between the two comparisons (Figure 5.5D).



Legend:

A) Principal Component Analysis (PC 1:2) of normalised (log) Nanostring gene expression data demonstrates some separation of stable and progressed /metastatic patients.

B) Heatmap of top 25 genes differentially expressed by Nanostring Nsolver analysis between Progressed and Stable patients ($p<0.05$), expression data is log normalised and scaled. Non-hierachical clustering (ward.D2 method and euclidean distance) demonstrates clustering of majority of progressed and metastatic patients with a similar expression pattern.

C) Heatmap of top 25 genes differentially expressed genes ($p<0.05$) between Metastatic and Stable patients, Heatmap and clustering (as per B) demonstrates similarity between Metastatic and Progressed patients.

Figure 5.5 nanoString differential gene expression analysis.

There was a good correlation between nanoString expression (log normalized) and 3'RNAseq expression (rlog normalized) with an overall correlation coefficient of 0.68 (Spearman correlation. An overlap of 4 differentially expressed genes between Stable disease and Progressed disease cases (increased genes: *GNAS*, *ETV1*, *COL2A1*; decreased gene: *HDAC5*) was identified between the 3'RNAseq differential expression analysis utilising less stringent cutoffs (*padj* < 0.2) and the nanoString analysis (*p* <0.05)

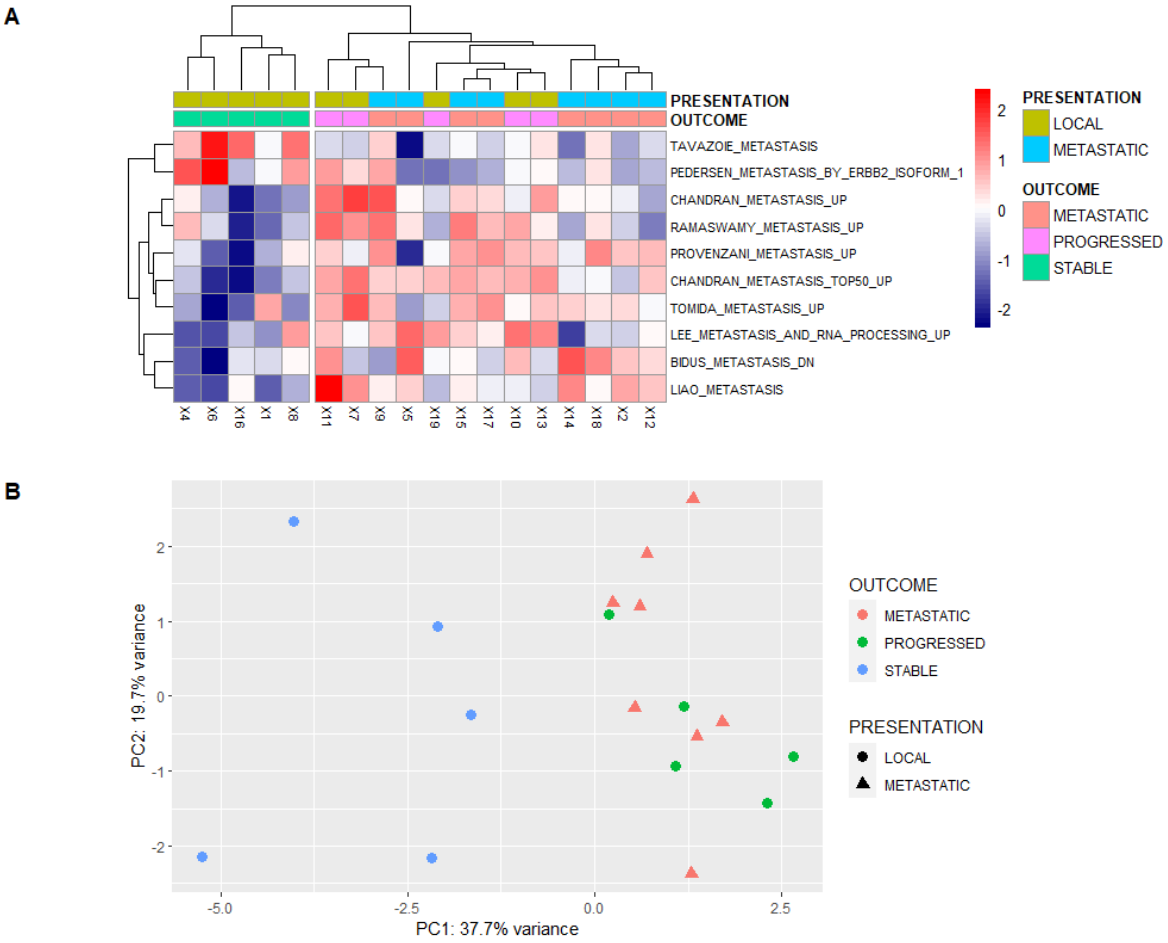
(Appendix Table 3.6). These genes are concordant in direction i.e. increased or decreased in Stable vs Progressed cases in both differential expression analysis performed with each platform.

Pathways associated with metastasis, centrosome and methylation were differentially expressed between Stable disease cases and the Progressed cases and *de novo* Metastatic cases using Single sample gene set expression analysis (ssGSEA) of the 3'RNAseq data (Appendix Figure 3.7). To further explore whether this was due to the same set of genes being present in the different pathways, the overlaps between pathways were computed. Only the centrosome pathways contain overlapping sets of genes (Appendix, Figure 3.8, Table S3).

To further explore differences between Progressed disease and *de novo* Metastatic cases compared to the Stable post-radiotherapy cases, single sample Gene Set Enrichment Analysis (ssGSEA) was performed restricted to a subset of metastatic pathways from the MSigDB database (a curated repository of published signatures) that had been previously validated.

Ten signatures were demonstrated to be statistically significant (*padj* <0.1) as visualised in the heatmap (Figure 5.6A). A clear separation of Stable disease cases compared to the Progressed disease and *de novo* Metastatic cases can be seen on Principal component analysis (PCA) of the ssGSEA scores. This demonstrates differences in the expression of genes between the cohorts of genes from previously identified metastatic signatures (Figure 5.6B).

5.14 Gene Set Enrichment Analysis



Legend:

A) Guided single sample gene set analysis (ssGSEA) identifies differences in metastasis pathways between stable and progressed/metastatic patients (Metastatic signatures subset of C2 MsigDB curated pathways, adjusted pvalue < 0.1, Quantseq rlog normalised data)

B) Principle component analysis of metastasis signatures ssGSEA scores as defined above (A) demonstrates separation of stable and progressed/metastatic groups

Figure 5.6 Guided ssGSEA of metastatic signatures

Panel A demonstrates heatmap of guided ssGSEA metastasis signatures, x axis demonstrates patients, y axis demonstrates significant pathways (p adj <0.1) Panel B demonstrates Principal Component Analysis of ssGSEA scores.

In order to explore the biological relevance and clinical applicability of the differentially expressed genes from the 3'RNAseq data, the top 50 increased protein coding genes in Progressed disease cases versus Stable disease, which were also increased in *de novo* Metastatic cases versus Stable cases (Figure 5.3A) were taken. The expression of these genes was explored in a large external radical radiotherapy dataset by Jain, which contains

micro-array RNA sequencing data for 248 Prostate Cancer patients treated with radical radiotherapy and their associated clinical follow-up. The probe sets corresponding to the top 50 increased protein-coding genes were selected (only 37 genes of the top 50 genes were presented in the sequencing data, accounting for 116 probes, as per Methods. The cohort was divided by the median value of summed expression of all 37 genes, into a high and low expression cohort. The clinical outcomes of the high and low expression cohort were analysed. Biochemical recurrence-free survival (p<0.001), and metastatic progression-free survival (p<0.001), were observed to be significantly different between the high- and low-expression cohorts (Figure 5.7).

5.15 Application of results to external radical radiotherapy dataset (Jain)

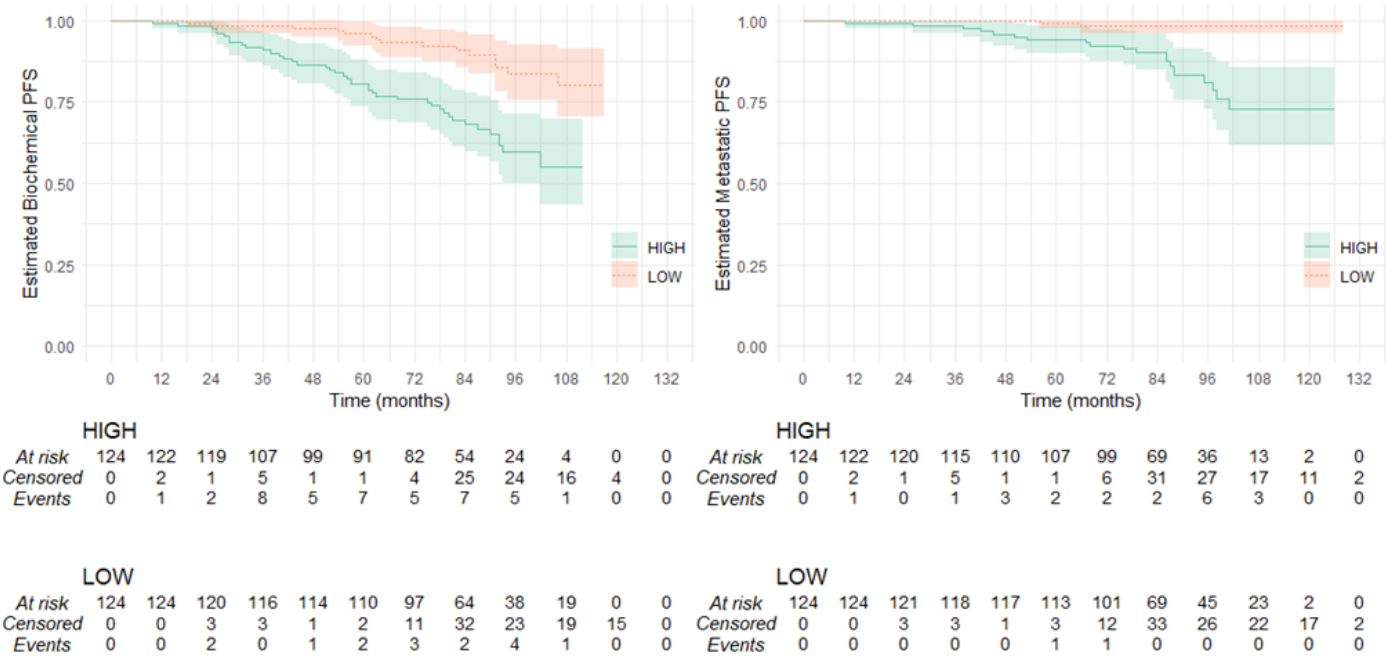


Figure 5.7 Survival plots of application to radical radiotherapy dataset

Left Panel demonstrates Survival plot (KMunicate format) of Biochemical Progression Free survival, of top 50 increased genes in Progressed vs Stable patients (Quantseq), in Jain radiotherapy cohort divided by median expression into HIGH and LOW cohorts. Y axis shows Estimated Survival Time and x axis in months. The Right Panel demonstrates as per above for Metastatic Progression Free Survival.

To further explore whether these results were due to chance, signature performance for biochemical recurrence-free survival and metastatic progression-free survival was significant compared to random signatures, known cancer signatures, and permutations of survival and feature data performed using the SigCheck package with 1000 iterations (Appendix Figures 3.8, 3.9, 3.10).

Using a per gene univariable cox proportional hazards model, four genes (*CDC25A*, *OLR1*, *CDON* and *DDX39B*) were found to be independently associated with biochemical progression-free survival (Hazard ratio lower limit >1) with a hazard ratio of 3.47 (1.79 – 6.7) for high compared to low expression cohorts (Appendix Table S4, Figure 3.11). These genes were demonstrated to predict for biochemical (AUC 76.8) and metastatic (AUC 82.9) progression-free survival, utilising Receiver Operator Curves with the combined gene expression of 4 genes as a continuous variable (Appendix Figure 3.12, Figure 5.13). The expression of these genes was demonstrated to outperform standard clinico-pathological features (Gleason score and PSA) used to predict likelihood of recurrence. To explore whether the expression of individual genes was associated with specific clinical characteristics a multivariate cox proportional hazard model was utilised, results are shown in Appendix Table 3.11.

Due to limited sample material Methylation analysis was only successfully performed on 8 samples (2 samples from patients with Stable disease, 3 from patients with Progressed disease, and 3 from patients with *de novo* Metastatic disease). Methylation analysis was performed on two illumina 850K EPIC arrays. Standard protocols were used for normalization and filtering of methylation probes as per Methods. Initial analysis with Principal component analysis identified similar findings to the 3'RNAseq data, with a degree of separation of the samples from patients with Stable disease samples from those with Progressed disease or *de novo* Metastatic (Figure 5.8A).

5.16 Methylation analysis

Legend:

A) Principle Component Analysis, post-normalisation and filtering demonstrates separation of Stable and Progressed/Metastatic samples by Principle Component 1.
B) Differential Methylation post-normalisation and filtering, identifies minimal differential methylation (adjusted pvalue < 0.05) between Metastatic and Progressed samples, and directionally coherent overlap between Progressed vs Stable and Metastatic vs Stable samples in differentially methylated probes (Mvals)
C) Heatmap demonstrating overlap of differentially methylated probes as per above (B), increased (LFC > 1) in Progressed and Metastatic vs Stable.
D) Heatmap demonstrating overlap of differentially methylated probes as per above (B), decreased (LFC < -1) in Progressed and Metastatic vs Stable.

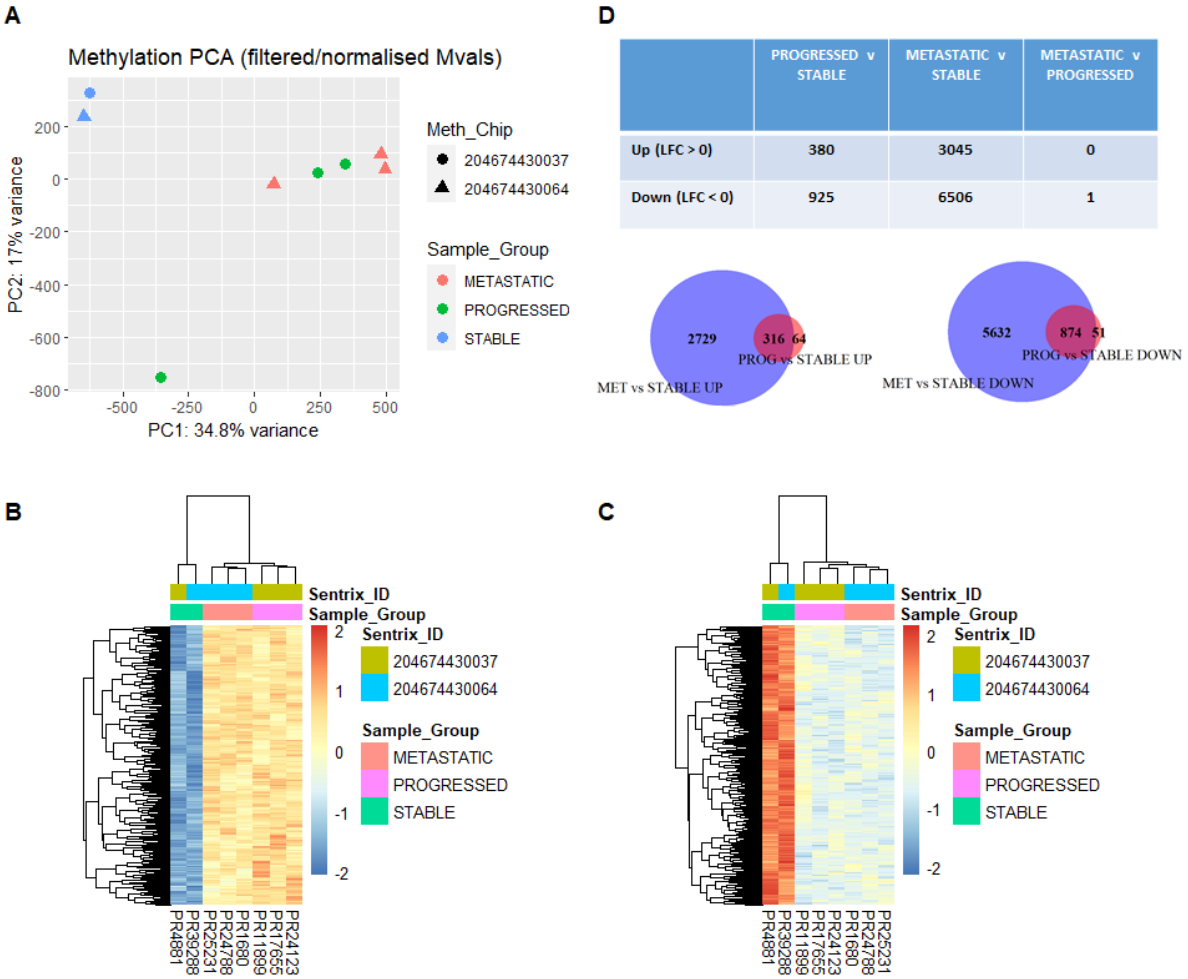


Figure 5.8: Methylation analysis

Panel A demonstrates PCA plots by category, Panel B demonstrates heatmap of differentially methylated probes (Mvals) patients on x axis, probes on y axis.

Between samples from patients with Stable disease compared to samples from patients with Progressed disease, 1305 probes were demonstrated to be differentially methylated using a strict cutoff of padj <0.05. Between samples from patients with Stable disease compared to

those with *de novo* Metastatic disease 9551 probes were identified as significantly differentially methylated again using a strict cutoff of padj <0.05.

A significant degree of overlap was seen between comparisons with the majority of probes (94.5%(874/925 probes) hypomethylated in Stable disease versus Progressed disease also demonstrated to be hypomethylated , comparing Stable disease versus *de novo* Metastatic samples (Figure 5.8D). A similar finding was also noted in the hypermethylated probes, and most probes (83.1% (316/380 probes) hypermethylated in Stable disease versus Progressed disease samples are also hypermethylated in Progressed disease versus *de novo* Metastatic samples (see Venn Diagram Figure 5.8D). This similarity is further demonstrated as only 1 probe was significantly differentially methylated between Progressed disease samples versus *de novo* Metastatic cases, suggesting that they are similar in terms of methylation patterns. Heatmaps of the coherently differentially methylated probes display the similar patterns of methylation seen between Metastatic and Progressed cases compared to stable cases (Figures 5.8B,3.88C).

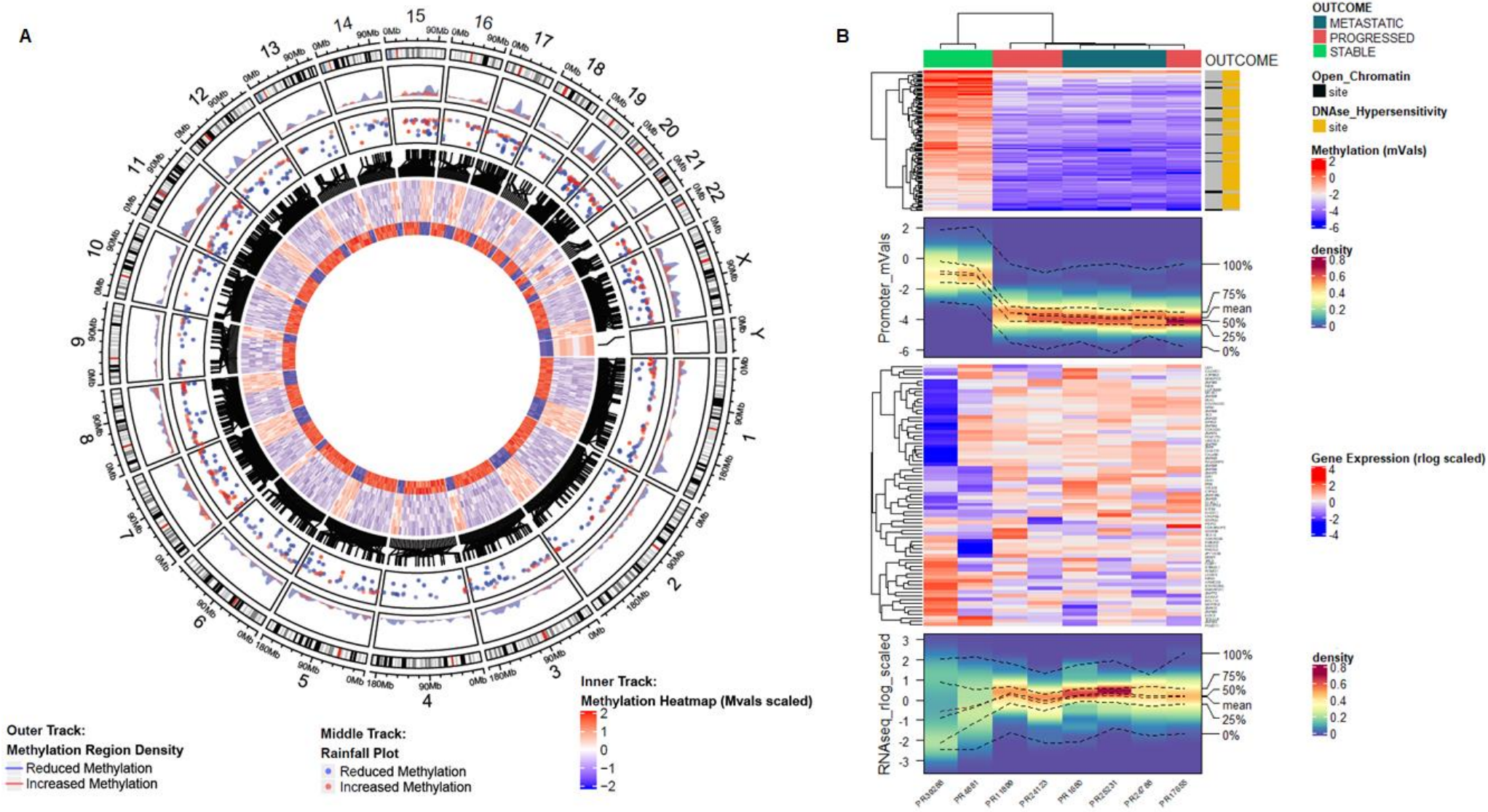
To investigate which genes are associated with the differentially methylated probes, Differentially Methylated Regions (DMRs) analysis was conducted between groups. More than 600 DMRs were identified between Stable disease cases compared to samples from cases which had Progressed following radical radiotherapy or patients with *de novo* Metastatic disease. These genes included known biologically important genes in Prostate Cancer such as *GNAS* and *AR*. The Methylation plots demonstrate similarities in mean methylation between Progressed disease and *de novo* Metastatic samples, which are distinct compared to the Stable disease samples (Appendix Figures 3.15).

To identify where the overlapping differentially methylated probes in Venn diagrams (Figure 5.8B) lie in genomic space they were mapped to their genomic location to visualise the Differentially Methylated Regions and allow visualisation of localisation to individual

chromosomes. Chromosome 19 is identified as having the largest peak; the outer track of Figure 5.9A demonstrates the chromosomal location of hypo-methylated regions in blue and hyper-methylated regions in red with a clear peak at chromosome 19 that is distinct from other regions. The genomic coordinates of each region are displayed in the Rainfall plot which lies in the middle track of Figure 5.9A. For the rainfall plot the y-axis demonstrates the minimum distance to neighbouring regions, and again there can be seen a cluster of differentially methylated probes which locate to Chromosome 19. An over-representation analysis was performed to identify if these regions correspond to specific pathways by first identifying the corresponding genes to the probes in this region and performing pathway analysis. Over-representation analysis identified significant pathways ($p$adj<$1x10^{-14}$) associated with RNA pol II cis-regulatory region (GO molecular function/gprofiler), sequence-specific DNA binding, DNA binding transcription factor activity (GO Biological Process) and Herpes Simplex 1 infection (KEGG) (Appendix Figures 15-17).

5.17 Integration of transcriptomic and methylation data

Figure 5.9 Methylation and gene expression analysis of coherently differentially methylated probes


A
Outer Track:
Methylation Region Density
Reduced Methylation
Increased Methylation
Middle Track:
Rainfall Plot
Reduced Methylation
Increased Methylation
Inner Track:
Methylation Heatmap (Mvals scaled)
B
OUTCOME
METASTATIC
PROGRESSED
STABLE
Open_Chromatin
site
DNAse_Hypersensitivity
site
Methylation (mVals)
density
Promoter_mVals
100%
75%
mean
50%
25%
0%
Gene Expression (rlog scaled)
RNAseq_rlog_scaled

Legend:

A) Circular plot of coherently differentially methylated regions between Progressed/Metastatic and Stable patients
    a. Outer Track : Density plot showing fraction of genomic window covered by differentially methylated regions in Progressed and Metastatic patients versus Stable patients, hypo-methylated regions shown in blue and hyper-methylated regions in red.
    b. Middle Track: Rainfall plot showing genomic coordinates of each region, y-axis corresponds to minimum distance to neighbouring regions
    c. Inner Track: Circular heatmap of coherently differentially methylated probes (Mvals scaled). Outer 3 lanes correspond to metastatic patients, middle 3 lanes progressed, and inner 2 lanes stable patients.

B) Combined heatmaps of promoter methylation, gene expression and density plots.
    a. Top Panel - Methylation heatmap of promoter associated probes hypomethylated in progressed post-RT/de novo metastatic compared to stable post-RT patients. Probes are normalised/filtered and differential methylation performed as per methods. Hierarchical clustering is via "euclidean" distance and "ward.D2" method. Open chromatin regions (OCRs) and DNase I hypersensitivity sites (DHSs) are annotated demonstrating promoter associated cpg sites are either DHSs or OCRs.
    b. Second Panel - Methylation density heatmap visualizing the distribution of mVals as per a). Columns are clustered as per a)
    c. Third Panel - Gene Expression (Quantseq) of genes corresponding to cpgs in a) & b). Expression is normalised (rlog - DESeq2) and scaled. Rows are clustered by expression and columns as per a).
    d. Bottom Panel - Gene expression density heatmap displayed as density of Quantseq expression values as per c). Columns are clustered as per a)

To demonstrate the relationship between differential methylation and 3'RNAseq gene expression, the coherently hypo-methylated probes between Progressed disease compared to Stable disease cases and *de novo* Metastatic cases compared to Stable disease cases were identified (Figure 5.9B). This list of probes was subset to probes that are associated with gene promoters. Integration of the methylation data with 3'RNAseq (rlog normalized) expression data was conducted for the corresponding 84 genes which were commonly hypomethylated in promoter associated regions. This integrated dataset was used to visualise methylation and expression heatmaps as well as density plots on the same graphic. Performing unsupervised hierarchical clustering of these differentially methylated promoter associated probes identifies that samples from patients with Stable disease post-radiotherapy cluster together, distinct from samples from patients with Progressed disease post-radiotherapy or *de novo* Metastatic cases (Figure 5.9B Top Panel). This demonstrates similarity in the methylation profiles of these gene promoter associated probes in Stable samples that is distinct from those with progressive or metastatic disease. Additional annotation identifies that most probes relate to open chromatin regions or DNase I hypersensitivity sites (Figure 5.9B Top Panel).

To further explore similarities in methylation profiles a methylation density heatmap was constructed to display the distribution of methylation values (mVals) (Figure 5.9B Second Panel). The methylation density is noticeably compact and concordant between Progressed disease and *de novo* Metastatic cases and again distinct from Stable cases. Patients with Stable disease appear to demonstrate more variation in the corresponding 3'RNAseq gene expression of the corresponding genes to the hypo-methylated promoter-associated probes (Figure 5.9B Third Panel) than that seen in the progressive disease and metastatic patient samples. Density plots of gene expression demonstrate that the expression of genes corresponding to hypo-methylated promoters is more closely distributed in the Progressed disease and *de novo* Metastatic samples, whereas gene expression in the Stable disease samples (where the promoter-associated probes were comparatively hyper-methylated)

shows a greater degree of variation (Figure3. 9B Bottom Panel). Similar findings are seen in the RNA expression of the same set of differentially methylated genes in the full 3' RNAseq cohort (Appendix Figure 5.20).

Using two separate platforms (enrichr and gprofiler), over-representation analysis was performed of the hypo- and hyper-methylated genes when comparing the Stable disease cases versus Progressed disease and *de novo* Metastatic cases. This pathway analysis identifies histone H3 lysine 4 trimethylation (H3K4me3), histone H3 lysine 27 trimethylation (H3K27me3), and Polycomb Repressive Complex 2 (PRC) pathways as being significantly enriched for in the results of the differential methylation analysis (Appendix Figures 3.16-3.20).

5.18 Further exploration in TCGA Prostate Adenocarcinoma data

To further explore the expression of genes identified as increased in Progressed vs Stable (Figure 5.4D) in a larger dataset, the TCGA PRAD dataset containing 550 samples was used, with a strict significance cutoff ($p<0.01$). Pearson correlation analysis identified 4 clusters of correlated genes (Figure 5.10), with pathway analysis identifying the distinct gene clusters as being most enriched for Microtubule motor activity (Green), Regulation of T-Cell Apoptotic Process (Blue), Fibroblast Growth Factor-activated receptor activity, and Histone Serine Kinase activity).

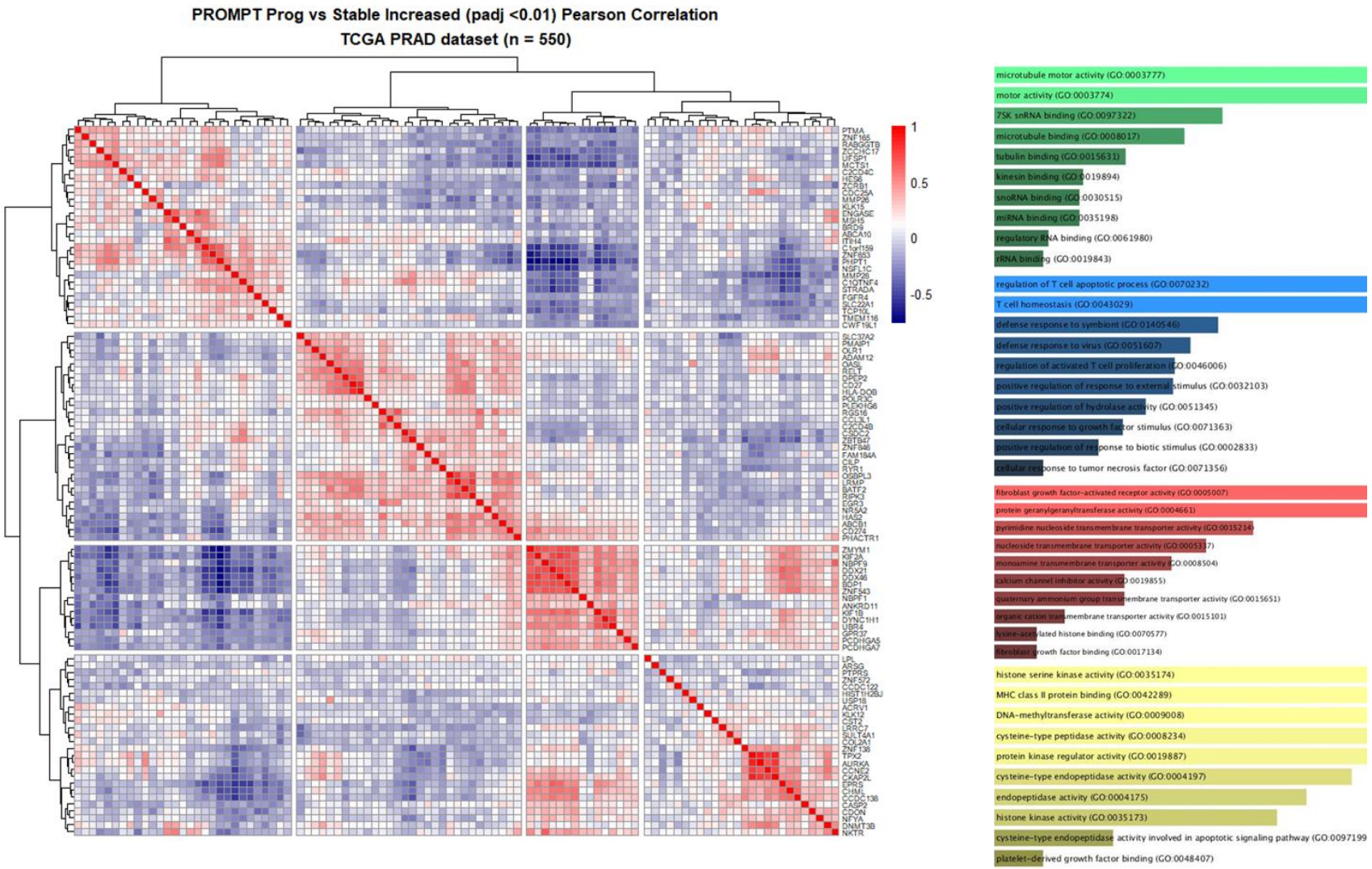


Figure 5.10. TCGA Prostate Adenocarcinoma data expression of genes identified as increased in Progressed vs Stable (p<0.01) in TCGA PRAD dataset containing 550 samples was used, with a strict significance cutoff (p<0.01). Hierarchical clustering of Pearson correlation identifies clusters of correlated genes. Pathways associated with each cluster identified via enrichment analysis are shown in the table right.

This demonstrates that there are coherently expressed sets of genes which have distinct biological functions that are differentially expressed between the cases that have Progressed following radical radiotherapy and those that are stable. Further analyses of these specific pathways could be further investigated to potentially improve clinical outcomes.

5.19 Chapter Summary and Discussion

Transcriptional signatures prognostic of metastatic recurrence have been described in Prostate Cancer, and some of these have been validated in patients treated with radical radiotherapy[197-202]. However, all these studies, apart from the 'hypoxia signature', describe signatures derived from radical prostatectomy specimens rather than from pre-treatment prostate biopsies from patients undergoing radical radiotherapy. DNA methylation is altered in Prostate Cancer development and progression[195], however whilst prognostic biomarkers have been developed[192] this approach has not been specifically used for radical radiotherapy-treated patients.

One of the clinical samples failed 3'RNAseq analysis but all other technical analyses produced utilisable data. Of note one clinical sample achieved fewer counts than others (500,000 versus ~8 million), for several reasons this sample was retained within the overall analysis. 3'RNAseq only measures transcripts from the 3' end adjacent to the poly-A tail, from a single read a single count is produced for the whole transcript as each mRNA molecule has a single poly-A tail. Standard RNAseq sequencing measures multiple fragments often across the same gene transcript, so conventional minimum count sizes established for standard RNA sequencing technologies are likely to be overestimates of minimum sample size required. Differential expression analysis pipelines can take account of differing library sizes per sample and have statistical methods to make allowances for these differences between samples. The analysis was also performed with this sample excluded and compared to the results with the presented results with inclusion with minimal change to the results. Additionally, the analysis was repeated utilising Subread aligner

versus STAR with this sample both included and excluded, again achieving similar results, to explore the impact of mapping techniques (Supplementary Data).

The 3'RNAseq differential gene expression analysis demonstrated several similarities in gene expression profiles between Prostate Cancer samples from patients with Progressive disease post-radiotherapy and those with *de novo* Metastatic disease, that were distinct from the group of patients with stable disease post radiotherapy. These similarities are seen in principal component analysis, the number of overlapping differentially expressed genes between the different comparisons, that these overlapping genes were differentially expressed in the same direction between groups i.e. increased in both comparisons and is clearly visualised with sample clustering with hierarchical clustering analysis.

These gene expression similarities may suggest a shared underlying biological phenotype between the samples from patients who progressed following radiotherapy and the *de novo* Metastatic samples, which could potentially be utilised in the future as an additional source of prognostic information. Given the historical nature of samples utilised in this study one explanation could be that historical radiological staging techniques were unable to identify patients who had metastatic disease at the outset. Modern staging techniques such as PSMA-PET can identify low volume metastatic disease that was previously occult on bone scans. Another potential is that the patients had micro-metastatic disease which was below the threshold of any radiological staging method, if this were the case the ability to identify those at higher risk from diagnostic biopsy material could potentially alter management. A further intriguing possibility is that the patients were adequately staged but the molecular changes that define metastatic prostate cancer also confer a degree of relative radioresistance.

The similarities in RNA expression between groups observed in the 3'RNAseq analysis was further identified in the nanoString analysis, specifically in the Principal Component Analysis and differential expression analysis which mirrored the results of the 3' RNAseq. There is a notable incomplete separation of the patients with Stable disease from those with

progressed and metastatic disease, with one *de novo* Metastatic sample clustering with the stable samples on hierarchical clustering. One reason for this is likely to be the more limited scope of Nanostring relying on only ~800 reporter genes compared to ~20,000 for 3'RNAseq, which significantly reduces the power of the technique for this sort of investigational approach, rather than a validation type scenario.

Another possibility is the underlying differences in technique with nanoString utilizing reporter probes to hybridize mRNA whereas 3'RNAseq sequences the 3' end of transcripts adjacent to the poly-A tail. Overall, the correlation between 3'RNAseq (rlog normalized) expression and nanoString (log normalized) expression taking only the genes present in both datasets demonstrated good concordance with a correlation coefficient of 0.68 using Spearman. Per-sample correlation when different normalization techniques were used for the 3'RNAseq were also broadly similar.

Differentially expressed genes from the 3'RNAseq analysis were not statistically significant on correction for multiple testing in the Nanostring dataset, using the NSolver analysis software, the results of both techniques are supporting. These two orthogonal sequencing platforms both clearly demonstrate similarities between Prostate Cancer samples from patients with Progressive disease post-radiotherapy and those with *de novo* Metastatic disease, that were distinct from the group of patients with stable disease post radiotherapy.

Over-representation analysis of the differentially expressed genes identified Gene Ontology pathways associated with spindle pole and centrosome as increased in samples with Progressive disease post-radiotherapy and those with *de novo* Metastatic disease, stable disease post radiotherapy. Single Sample Geneset Enrichment also demonstrated pathways associated with Centrosomes as increased, possibly because of a higher rate of mitosis in those samples. A previously validated measure of Cell Cycle Progression has demonstrated clinical utility in predicting metastatic progression post radiotherapy, although there is minimal overlap of genes between these genesets [91,104].

The increased expression of spindle pole genes in patients who progress following radical radiotherapy could be due to loss of other cell cycle checkpoints and increased reliance on spindle assembly checkpoint [105]. In Prostate Cancer cell lines centrosome amplification is demonstrable in lines with increased metastatic behaviour, the increased expression of centrosome pathway genes could be due to centrosome amplification in the samples of patients who progress after radiotherapy or with metastatic disease at presentation, although studies have demonstrated *in situ* centrosome loss in primary Prostate Cancer samples [106]. These findings could be connected to increased chromosomal instability and aneuploidy.

Across both platforms 3'RNAseq and nanoString there were four genes (*GNAS*, *ETV1*, *COL2A1* and *HDAC5*) that were consistently differentially expressed comparing patients who were Stable post-radiotherapy with those who progressed post radiotherapy, using a stringent 3'RNAseq cutoff (*p*adj<0.05) and a less stringent Nanostring cutoff (p<0.2). When a cox proportional hazards model was used on univariable analysis only COL2A1 and ETV1 were significantly prognostic.

GNAS is known to be one of the most altered genes in Prostate cancer with a large-scale integrative clinical genomic study confirming that in advanced Prostate Cancer GNAS is one of the most common mutated genes[203].

Frequently overexpressed in aggressive Prostate Cancer via a chromosomal translocation with androgen-responsive promoters, ETV1 is a transcription factor of the ETS family that modulates cell growth and migration[204]. In cooperation with the JMJD2 histone demethylase, ETV1 can initiate the formation of Prostate Cancer[205]. COL2A1 has previously been identified as a potential Prostate Cancer risk gene [206].

HDAC5 is a histone deacetylase, which is frequently downregulated or deleted in Prostate Cancer. This loss results in an increase in H3K27 acetylation and subsequent impairment of RetinoBlastoma mediated cell cycle pro-oncogene repression[207].

The expression of HDAC5 was found to be reduced in both Nanostring and 3-RNAseq analysis in Progressed post-radiotherapy patients compared to Stable post-radiotherapy patients. To investigate the prognostic power of these four genes (GNAS, ETV1, COL2A1 and HDAC*5*) they were explored in an external dataset of radical radiotherapy treated patients with prostate cancer by Jain [196] and found to be prognostic for biochemical progression-free survival, though not for metastatic progression-free survival (Supplementary Figure 5.11).

Single Sample gene expression analysis, limited to a subset of previously described metastatic signatures, identifies that metastatic pathways are differentially expressed between the Stable post-radiotherapy and progressed post-radiotherapy patients/*de novo* metastatic patients (Figure 5.5A,5.5B). Principal component analysis demonstrates a clear split between these groups, and similarities between the Progressed post-radiotherapy patients and *de novo* metastatic patients who cluster together separately from the Stable patients.

The Top 50 protein coding genes differentially expressed between Progressed post-radiotherapy patients and Stable post-radiotherapy patients in the clinical dataset generated and applied to a large external clinical dataset of radiotherapy treatment patients with prostate cancer (Figure 5.7). Of note the identified genes do not overlap at all with previously validated prognostic signatures in prostate cancer Prolaris, Oncotype, Decipher, Metastatic Assay and Prostate Hypoxia gene lists (Supplementary Figure S29). When the Jain cohort is separated into high and low expression cohorts of the 37 represented genes, there is a clear separation in survival curves, in the expected direction, and the geneset is prognostic for both biochemical and metastatic progression-free survival (Figure 5.7C,5.7D). This was further explored and demonstrated to be more prognostic than expected by chance selection of a random set of genes, and more prognostic than a set of selected cancer signatures. On univariable analysis four genes (CDC25A, OLR1, CDON and DDX39B) were identified as being independently prognostic, these four genes have previously in the literature been

reported to be associated with Prostate Cancer progression and/or metastasis [207-210]. Although not adequately powered for biomarker selection the results of this experiment do demonstrate that there is great potential to identify prognostic information from pre-treatment diagnostic biopsies which could be leveraged for clinical benefit.

Extensive DNA hypo-methylation patterns have been previously identified in the context of Prostate Cancer. Recent investigations into advanced stages of the disease have identified an increase in specific regions exhibiting greater levels of hypo-methylation as the transition occurs from normal tissue to localised Prostate Cancer and eventually to metastatic disease[211]. The experimental methylation observations align with this finding, with comparable methylation patterns seen in both Progressive post radiotherapy cases and *de novo* Metastatic cases. There is a greater degree of methylation in *de novo* Metastatic cases compared to Stable Prostate Cancer cases, than in the comparison between Progressive Prostate Cancer cases and Stable Prostate Cancer cases (Figure 5.8B, 5.8C).

Methylation analysis demonstrated separation of Stable-Prostate Cancer cases and a clear separation between patients with Stable Prostate Cancer and those with Progressive Prostate Cancer/*de novo* Metastatic, seen in both principal component analysis and differential methylation analysis. A significant proportion of differentially methylated genes exhibited consistent changes in the same direction (Figure 6 A-D). This study is constrained by a relatively small sample size with small sample numbers per group particularly for methylation analysis. Previous research has shown similar methylation patterns in primary and metastatic tumour samples derived from the same patient[212], reinforcing the concept that methylation characteristics can be detected in primary tumour biopsies that are indicative of metastatic disease. These DNA methylation findings offer further evidence, beyond the 3' RNAseq and nanoString data, of biological similarities observed in baseline biopsy samples from patients with Progressive Prostate Cancer post radiotherapy and *de*

novo Metastatic conditions at both the methylomic and transcriptomic levels, that is distinct from patients who have stable disease post radiotherapy.

The methylation technology used in this study, utilising 850K methylation probes, can identify subtle differences in methylation patterns within individual genes. This study has identified specific differentially methylated DNA regions that correspond to genes with established biological and prognostic significance in Prostate Cancer, including Androgen Receptor (AR) and GNAS.

Androgen Receptor hypomethylation has previously been observed in metastatic Prostate Cancer [213,214]. Within this Chapter I have demonstrated that GNAS exhibited hypomethylation and increased expression in samples of patients who progressed post radical radiotherapy or had *de novo* Metastatic samples when compared to samples from patients who had stable disease post radiotherapy. The relationship between DNA methylation and gene expression is complex, some studies have linked increased gene expression with hypermethylation [215]. In this analysis of differential methylation, samples from patients with Progressive Prostate Cancer post-radiotherapy and *de novo* Metastatic samples exhibited hypomethylation in the initial regions of genes and hypermethylation in other regions when contrasted with Stable Prostate Cancer samples. Variations in methylation within the same gene were also identified as exemplified by the case of GNAS further illustrating the complexity of methylation and difficulties with interpretation[215].

The peak genomic region of DNA hypomethylation comparing patients with Progressive Prostate Cancer post-radiotherapy or *de novo* Metastatic disease to those stable post-radiotherapy was on chromosome 19. Pathway enrichment analysis of genes encoded at the site of peak differential methylation density on chromosome 19 demonstrates significant over-representation of pathways associated with RNA pol II cis-regulatory region. RNA Pol II–associated chromatin interactions have been previously identified as key determinants of transcriptional regulation in Prostate Cancer [216].

RNA Pol II chromatin interactions have previously been demonstrated to be orchestrated by H3K4m3 and/or H3K27 acetylation marks [216]. Over-representation analysis of differentially methylated probes between progressed/metastatic cases and stable cases identified pathways associated with H3K4m3, H3K27m3, and PRC2 as significantly differentially methylated pathways (Supplementary Figures S25, S26). The fact that these pathways appear in pathway analyses of both hyper- and hypo-methylated genes, again demonstrates the complexity of methylation events in comparison to gene expression, where different areas of the same genes and hence pathways can be either hyper- or hypo- methylated between conditions. The Polycomb Group Protein Enhancer of Zeste 2 (EZH2) function is known to be with the Polycomb Repressor Complex 2 (PRC2) catalysing H3K27m3 on target gene promoters [217,218] and has been assocaiated with metastatic progression post-radical radiotherapy and metastatic Prostate Cancer progression[219]. Genes which are previously identified as downstream targets of EZH2 expression are demonstrated to be significantly differentially expressed between progressed/metastatic and stable groups (Supplementary Figure S24), suggesting that this pathway may be involved.

The results in Chapter 5 demonstrate that I have achieved the fourth aim of this thesis - to identify features of a radioresistant phenotype from a clinical cohort of patients with prostate cancer treated with radiotherapy and identify their predictive value in an external cohort of prostate cancer patients treated with radiotherapy.

In the cell line experiments, which utilised non-prostate cancer cell lines, the most differentially methylated pathways between radioresistant and parental cell lines involved pathways associated with trimethylation of histone H3 on lysine 27 (H3K27me3), (Figure 2.40). This suggests a common pathway of altered H3K27me3 in radioresistant cells and is supported by the analysis of clinical prostate samples, where pre-treatment samples with these features are associated with progression post-radiotherapy. These altered methylation profiles may lead to altered transcriptional profiles with increased expression of known poor prognostic genes e.g. FAM83A in cell line analysis, ETV1 in prostate cancer

clinical samples. This is supported by studies in other tumour types where loss of H3K27 Trimethylation promotes radiotherapy resistance in medulloblastoma [220].

The significantly ($p<0.01$) increased genes in Progressed and *de Novo* Metastatic patients when applied to the Prostate Cancer TCGA dataset cluster into 4 groups by correlation (Figure 5.10); Microtubule motor activity (Green), Regulation of T-Cell Apoptotic Process (Blue), Fibroblast Growth Factor-activated receptor activity, and Histone Serine Kinase activity. These suggest that there may be additional factors which may be influential in radiotherapy outcomes that are not considered by standard radiobiology including immune cell activation and growth factor activation. Although there is no direct overlap there are similar findings from the result of the update computation of overlap from cell line studies Figure 5.54 which identified the inflammatory cytokine IL-6 and the inhibitory immune checkpoint molecule NT5E as upregulated in radioresistant cell lines.

The results of this Chapter demonstrate similar gene expression patterns between *de novo* metastatic and progressed patients compared to patients who were stable following radiotherapy. One slight inconsistency was the minimal overlap between the findings of NanoString and RNAseq findings with only 4 genes differentially expressed in both platforms despite a relatively good correlation between expression levels (correlation co-efficient 0.68). The most likely reason for this finding is the limited number of genes that are available via the Nanostring platform ~200 compared to 30,000 by RNAseq, which limits the power of Nanostring to assess differential expression.

The small number of patient samples in this study is a large limitation, however the findings were applicable to a much larger number of patients when validated in the external dataset. The consistent findings of similar expression/methylation patterns across platforms does suggest that there are underlying molecular similarities between patients who progress following radiotherapy and those with *de novo* metastatic disease that are distinct from patients with stable disease following radiotherapy.

In the context of the existing literature as previously outlined the research in this Chapter is but one of numerous transcriptomic predictors of outcome in Prostate cancer however it does have some key advantages compared to other studies. This is the first study to derive a signature from radiotherapy treated patients, using 3’ RNA sequencing of historic pre-treatment diagnostic biopsies. The advantage of this is that it is possible to integrate into the standard of care where most biopsies are FFPE, also it demonstrates that valuable information can be identified from such studies, which were previously considered to be too degraded for successful RNA sequencing, unlocking the potential from historical samples.

The aim of this Chapter to identify features of a radioresistant phenotype from a clinical cohort of patients with prostate cancer treated with radiotherapy and identify their predictive value in an external cohort of prostate cancer patients treated with radiotherapy has been achieved.

Chapter 6: Discussion

6.11 Summary:

The aim of this thesis was to utilise a systems biology and functional genomics approach to identify mechanisms of resistance to radiation and develop a predictive classifier. Review of the previous literature identified that the most coherent results of gene expression studies investigating radiosensitivity was obtained when comparing radioresistant cell lines generated through repeat irradiation to parent cell lines.

In Chapter 3 radioresistant cell lines were successfully established through fractionated daily irradiation comparable to clinical schedules (60Gy in 2Gy per treatment over 6 weeks) and through clonogenic assays demonstrated to be more resistant to radiation than parental cell lines (Figures 3.14-3.16). Irradiated cell lines were characterised and a higher percentage of cells in G2 was demonstrated 24 hours following further 2 Gy irradiation ($p= 0.016$) across cell lines, a greater proportion of senescent cells was noted in radioresistant cells following radiation, and radioresistant cells had slower growth rates. Utilising competitive ecological modelling and the experimentally derived growth rates a model was generated to explore the interaction of a mixed population of less prevalent, slow growing, radioresistant sub-population and a more prevalent, faster growing and more radiosensitive population which outcompeted the radioresistant population in normal growth conditions. The output of this model demonstrated that the radiosensitive population remained the dominant population under normal growth but under the pressure of daily fractionated radiation, despite the total cell number reducing with radiation, the proportion and number of radioresistant cells increased.

In Chapter 4 a gene expression experiment was performed to characterise differential gene expression between parental and radioresistant cell-lines at baseline and assess dynamic

changes following repeat irradiation with 2 Gy. Principal Component Analysis demonstrated H460 resistant and parental cell lines separate by the second Principal Component (Figure 4.3). Utilising a time interaction model a compact set of genes differentially expressed in parental lines across all timepoints was identified (Figure 4.22). Geneset Enrichment Analysis identified pathways such as Mitotic Spindle Checkpoint and G2/M transition as suppressed in radioresistant cell lines (Figure 4.23). Further investigation with single sample geneset enrichment analysis of the curated genesets (C2 MSigDB) identified pathways including Senescence (Fridman Senescence Up) and Methylation (Yamashita Methylated in Prostate Cancer) to be significantly increased 24 hours following repeat irradiation (Figure 4.27). A further Senescence signature from the literature of genes decreased in senescence, was demonstrated to be significantly (padj = 0.0101) lower in terms of ssGSEA score in radioresistant than parental lines. Utilising a shrinkage method a compact set of genes increased in radioresistant H460 cells was derived.

Methylation analysis identified similar pattern of separation of resistant and parental lines on PCA to gene expression (Figure 4.35), and the top differentially methylated region FAM83A was also noted to be one of the most significantly differentially expressed genes that was increased in radioresistant cell-lines. Further investigation of FAM83A gene expression, through gene correlation and pathway analysis, identified significant enrichment of Cell Cycle, Cellular Senescence, and p53 pathways (Appendix 2.9/2.10). Over-representation analysis of differential methylation identified multiple pathways associated with H3K27ME3 associated with both increased and decreased methylated probes (Figure 4.38,4.39).

The compact set of genes increased in radioresistant H460 cells was applied to a clinical dataset of radiotherapy treated patients from the TCGA and the cohort with lower expression of radioresistant genes had a better survival HR 0.45 (0.21-0.99), which was statistically significant at a pvalue of 0.048 (Figure 4.46). The previously identified Senescence signature was demonstrated to be predictive of outcome in the SCORT dataset with high senescence ssGSEA scores having a statistically significantly (pvalue = 0.02) higher odds ratio of having an incomplete response to treatment at 3.02 (1.23-8.09). The senescence signature was

also associated with worse outcomes in two other clinical cohorts one prostate (p = 0.021) and one rectal (p = 0.064) (Figure 4.53).

The results of the cell line experiments and clinical samples when taken together with the ecological modelling suggest a potential common mechanism of resistance to radiation. Radiation induces cellular changes including senescence and altered methylation including trimethylation of histone H3 on lysine 27 (H3K27me3). Senescence and methylation changes including Trimethylation of histone H3 on lysine 27 (H3K27me3) induce altered transcriptional profiles that promote radioresistance. Radioresistant sub-populations under the evolutionary pressure of repeat irradiation can become the dominant population, leading to failure of treatment.

In Chapter 5 a small cohort of prostate cancer patients' pre-treatment diagnostic biopsies underwent multi-omics sequencing. Coherent gene expression and methylation patterns were demonstrated in pre-treatment biopsies of patients progressing following radical radiotherapy, similar to patients who presented with metastatic disease and distinct from patients who had stable disease following treatment (Figures 5.4, 5.5, 5.8). The significantly differentially expressed genes in patients who had progressed following radiotherapy compared to those who were stable demonstrated the potential to predict for treatment failure in a large external dataset of radiotherapy treated patients (Figure 5.7).

Challenges of deriving clinically informative signatures for radiotherapy:

As outlined from the literature review in the introduction one of the key limitations of any attempt to aggregate previous studies is the wide variation in methodologies utilised. A significantly more robust methodological framework to generate such studies as described in the Introduction alongside key reporting frameworks such as RIPOSTE may help to improve the transition from laboratory experiments to clinical application. Another key limitation is the

dynamic nature of the mapping of the human genome and the architecture associated with analysis. This includes the constantly updated and evolving gene nomenclature but also the methods of analysis, with the evolution of RNAseq technologies largely replacing the use of microarray. This lack of homogeneity can make the use of historical datasets for validation more complex, requiring an in-depth understanding of both analysis techniques to ensure that the appropriate technical translation of findings is performed when trying to validate results. Another barrier to large scale development of radiotherapy predictive signatures is the lack of well curated clinical datasets of radiotherapy treated patients available for analysis. This significantly limits the application of newer machine-learning techniques, which require vast datasets to establish meaningful findings, but also potentially biases any findings of validation in only one or two large scale datasets, where relatively small sample sizes mean that application to wider clinical cohorts is almost impossible to recommend. The establishment of high quality, well curated clinical datasets of radiotherapy patients is the first step towards developing clinically informative signatures for radiotherapy, as without them any laboratory-based investigation is unable to progress to clinical application. One key finding of this thesis is that 3'RNA sequencing produces sequencing data from historical diagnostic samples that is suitable for the task and could help to generate such cohorts. This technique could unlock the potential to utilise large historical cohorts of currently unused FFPE samples.

Future Perspectives

The potential for systems biology techniques to predict radiotherapy outcomes holds great potential for future application to the clinic, however rigorous prospective validation trials would be required to determine their clinical efficacy. One key practical consideration to their implementation is the integration to standard of care clinical pathways. Following validation, for such tests to have any clinical utility they would need to be accessible and implemented within standard clinical care pathways. Although this has already been implemented in some tumour types e.g. Breast Cancer and Oncotype Dx and the decision for adjuvant

chemotherapy, in the case of radiation therapy if these tests are being used to guide treatment, they would need to be accessible in a timely manner. The limitations on time taken for RNA extraction sequencing and analysis mean that this may be challenging to achieve at the scale and cost required for clinical implementation. One emerging technology that has developed due to the recent advances in Artificial Intelligence is the potential for digital pathology to be analysed to provide additional information on patient outcomes. One such example of this application that has been developed is digital pathology analysis to predict benefit of hormone therapy in men being treated with radiation for prostate cancer [221]. Another recent study has utilised machine learning of gene expression from large scale datasets to infer gene expression from whole slide images of histology [222].

Conclusion:

This study has demonstrated the potential for systems biology approaches to investigate mechanisms of radioresistance. Investigations have identified cellular senescence as a potential mechanism for radioresistance during clinically relevant fractionated irradiation in cell lines and demonstrated a transcriptomic signature of cellular senescence to be prognostic in a clinical dataset of radiotherapy treated prostate cancer patients. The top differentially expressed genes increased in radioresistant cell lines are also demonstrated to be clinically prognostic in a cohort of lung cancer patients treated with radiotherapy and non-prognostic in non-radiotherapy treated cohort and include a key gene (AREG) associated with cellular senescence and a known poorly prognostic gene (FAM83A).

Utilising ecological modelling and experimentally derived parameters including growth rate and radiosensitivity a potential mechanism for failure of radiotherapy treatment has been identified, whereby slow-growing radioresistant sub-populations which occupy a small proportion of the total tumour population during normal growth can become the dominant population during fractionated radiotherapy treatment, leading to treatment failure.

Gene expression similarities between clinical samples from patients who have progressed post-radical radiotherapy to the prostate or have de novo metastatic disease compared to patients who are stable, are demonstrated to be predictive of outcome in a large radiotherapy treated dataset.

There are similarities in pathways that are differentially methylated between the clinical prostate cancer samples that progressed or were stable following radical radiotherapy and the cell line experiments between radioresistant and parental cell with numerous pathways involved with H3K27m3. This suggests that this may be a potential target for further investigation to improve radiotherapy outcomes.

There is clear potential for systems biology and functional genomics approaches to identify mechanisms of resistance to radiation and develop predictive classifiers, however much further work is required before they are ready for implementation in radiation oncology clinical practice.

REFERENCES

1. Royal College of Radiologists. RCR - Clinical Oncology Workforce the Case for Expansion Dec 2014 accessed 10/11/23 https://www.rcr.ac.uk/media/ljpicmjs/rcr-publications_clinical-oncology-workforce-the-case-for-expansion_december-2014.pdf
2. Sandhu S, Sharpe M, Findlay Ú, Roe C, Broggio J, Spencer K, Thackray K. Cohort profile: radiotherapy dataset (RTDS) in England. BMJ Open. 2023 Jun 20;13(6):e070699. doi: 10.1136/bmjopen-2022-070699. PMID: 37339842; PMCID: PMC10314480.
3. Hall E, Giaccia A, Radiobiology for the Radiologist, (2012) Wolters Kluwer Health/Lippincott Williams & Wilkins, ISBN 1608311937, 9781608311934
4. Barnett GC, West CM, Dunning AM, Elliott RM, Coles CE, Pharoah PD, Burnet NG. Normal tissue reactions to radiotherapy: towards tailoring treatment dose by genotype. Nat Rev Cancer. 2009 Feb;9(2):134-42. doi: 10.1038/nrc2587. Epub 2009 Jan 16. PMID: 19148183; PMCID: PMC2670578.
5. Chang D.S., Lasley F.D., Das I.J., Mendonca M.S., Dynlacht J.R. (2014) Cell and Tissue Kinetics. In: Basic Radiotherapy Physics and Biology. Springer, Cham, https://doi.org/10.1007/978-3-319-06841-1_24
6. Good JS, Harrington KJ. The hallmarks of cancer and the radiation oncologist: updating the 5Rs of radiobiology. Clin Oncol (R Coll Radiol). 2013 Oct;25(10):569-77. doi: 10.1016/j.clon.2013.06.009. Epub 2013 Jul 10. PMID: 23850153.
7. Pajonk F, Vlashi E, McBride WH. Radiation resistance of cancer stem cells: the 4 R's of radiobiology revisited. Stem Cells. 2010 Apr;28(4):639-48. doi: 10.1002/stem.318. PMID: 20135685; PMCID: PMC2940232.
8. Tabasso AFS, Jones DJL, Jones GDD, Macip S. Radiotherapy-Induced Senescence and its Effects on Responses to Treatment. Clin Oncol (R Coll Radiol). 2019 May;31(5):283-289. doi: 10.1016/j.clon.2019.02.003. Epub 2019 Feb 27. PMID: 30826201

9. Boustani J, Grapin M, Laurent PA, Apetoh L, Mirjolet C. The 6th R of Radiobiology: Reactivation of Anti-Tumor Immune Response. Cancers (Basel). 2019 Jun 20;11(6):860. doi: 10.3390/cancers11060860. PMID: 31226866; PMCID: PMC6627091.

10. Eschrich SA, Pramana J, Zhang H, Zhao H, Boulware D, Lee JH, Bloom G, Rocha-Lima C, Kelley S, Calvin DP, Yeatman TJ, Begg AC, Torres-Roca JF. A gene expression model of intrinsic tumor radiosensitivity: prediction of response and prognosis after chemoradiation. Int J Radiat Oncol Biol Phys. 2009 Oct 1;75(2):489-96. doi: 10.1016/j.ijrobp.2009.06.014. PMID: 19735873; PMCID: PMC3038688.

11. Sparano JA, Gray RJ, Makower DF, Pritchard KI, Albain KS, Hayes DF, Geyer CE Jr, Dees EC, Goetz MP, Olson JA Jr, Lively T, Badve SS, Saphner TJ, Wagner LI, Whelan TJ, Ellis MJ, Paik S, Wood WC, Ravdin PM, Keane MM, Gomez Moreno HL, Reddy PS, Goggins TF, Mayer IA, Brufsky AM, Toppmeyer DL, Kaklamani VG, Berenberg JL, Abrams J, Sledge GW Jr. Adjuvant Chemotherapy Guided by a 21-Gene Expression Assay in Breast Cancer. N Engl J Med. 2018 Jul 12;379(2):111-121. doi: 10.1056/NEJMoa1804710. Epub 2018 Jun 3. PMID: 29860917; PMCID: PMC6172658.

12. Scott JG, Berglund A, Schell MJ, Mihaylov I, Fulp WJ, Yue B, Welsh E, Caudell JJ, Ahmed K, Strom TS, Mellon E, Venkat P, Johnstone P, Foekens J, Lee J, Moros E, Dalton WS, Eschrich SA, McLeod H, Harrison LB, Torres-Roca JF. A genome-based model for adjusting radiotherapy dose (GARD): a retrospective, cohort-based study. Lancet Oncol. 2017 Feb;18(2):202-211. doi: 10.1016/S1470-2045(16)30648-9. Epub 2016 Dec 18. Erratum in: Lancet Oncol. 2017 Feb;18(2):e65. doi: 10.1016/S1470-2045(16)30686-6. PMID: 27993569; PMCID: PMC7771305.

13. Yard BD, Adams DJ, Chie EK, Tamayo P, Battaglia JS, Gopal P, Rogacki K, Pearson BE, Phillips J, Raymond DP, Pennell NA, Almeida F, Cheah JH, Clemons PA, Shamji A, Peacock CD, Schreiber SL, Hammerman PS, Abazeed ME. A genetic basis for

the variation in the vulnerability of cancer to DNA damage. Nat Commun. 2016 Apr 25;7:11428. doi: 10.1038/ncomms11428. PMID: 27109210; PMCID: PMC4848553.

14. Schwartz JL, Murnane J, Weichselbaum RR. The contribution of DNA ploidy to radiation sensitivity in human tumour cell lines. Br J Cancer. 1999 Feb;79(5-6):744-7. doi: 10.1038/sj.bjc.6690119. PMID: 10070863; PMCID: PMC2362676.
15. Lynch TJ, Bell DW, Sordella R, Gurubhagavatula S, Okimoto RA, Brannigan BW, Harris PL, Haserlat SM, Supko JG, Haluska FG, Louis DN, Christiani DC, Settleman J, Haber DA. Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib. N Engl J Med. 2004 May 20;350(21):2129-39. doi: 10.1056/NEJMoa040938. Epub 2004 Apr 29. PMID: 15118073.
16. Druker BJ, Talpaz M, Resta DJ, Peng B, Buchdunger E, Ford JM, Lydon NB, Kantarjian H, Capdeville R, Ohno-Jones S, Sawyers CL. Efficacy and safety of a specific inhibitor of the BCR-ABL tyrosine kinase in chronic myeloid leukemia. N Engl J Med. 2001 Apr 5;344(14):1031-7. doi: 10.1056/NEJM200104053441401. PMID: 11287972
17. Nishiwaki Y, Vansteenkiste J, Kudoh S, Rischin D, Eek R, Horai T, Noda K, Takata I, Smit E, Averbuch S, Macleod A, Feyereislova A, Dong RP, Baselga J. Multi-Institutional Randomized Phase II Trial of Gefitinib for Previously Treated Patients With Advanced Non-Small-Cell Lung Cancer. J Clin Oncol. 2023 Feb 20;41(6):1162-1171. doi: 10.1200/JCO.22.02499. PMID: 36791474
18. Ward RA, Anderton MJ, Ashton S, Bethel PA, Box M, Butterworth S, Colclough N, Chorley CG, Chuaqui C, Cross DA, Dakin LA, Debreczeni JÉ, Eberlein C, Finlay MR, Hill GB, Grist M, Klinowska TC, Lane C, Martin S, Orme JP, Smith P, Wang F, Waring MJ. Structure- and reactivity-based development of covalent inhibitors of the activating and gatekeeper mutant forms of the epidermal growth factor receptor (EGFR). J Med Chem. 2013 Sep 12;56(17):7025-48. doi: 10.1021/jm400822z. Epub 2013 Aug 30. PMID: 23930994.

19. Yver A. Osimertinib (AZD9291)-a science-driven, collaborative approach to rapid drug design and development. Ann Oncol. 2016 Jun;27(6):1165-1170. doi: 10.1093/annonc/mdw129. Epub 2016 Mar 8. PMID: 26961148.
20. Verweij J, Casali PG, Zalcberg J, LeCesne A, Reichardt P, Blay JY, Issels R, van Oosterom A, Hogendoorn PC, Van Glabbeke M, Bertulli R, Judson I. Progression-free survival in gastrointestinal stromal tumours with high-dose imatinib: randomised trial. Lancet. 2004 Sep 25-Oct 1;364(9440):1127-34. doi: 10.1016/S0140-6736(04)17098-0. PMID: 15451219.
21. Karapetis CS, Khambata-Ford S, Jonker DJ, O'Callaghan CJ, Tu D, Tebbutt NC, Simes RJ, Chalchal H, Shapiro JD, Robitaille S, Price TJ, Shepherd L, Au HJ, Langer C, Moore MJ, Zalcberg JR. K-ras mutations and benefit from cetuximab in advanced colorectal cancer. N Engl J Med. 2008 Oct 23;359(17):1757-65. doi: 10.1056/NEJMoa0804385. PMID: 18946061.
22. De Roock W, Claes B, Bernasconi D, De Schutter J, Biesmans B, Fountzilas G, Kalogeras KT, Kotoula V, Papamichael D, Laurent-Puig P, Penault-Llorca F, Rougier P, Vincenzi B, Santini D, Tonini G, Cappuzzo F, Frattini M, Molinari F, Saletti P, De Dosso S, Martini M, Bardelli A, Siena S, Sartore-Bianchi A, Tabernero J, Macarulla T, Di Fiore F, Gangloff AO, Ciardiello F, Pfeiffer P, Qvortrup C, Hansen TP, Van Cutsem E, Piessevaux H, Lambrechts D, Delorenzi M, Tejpar S. Effects of KRAS, BRAF, NRAS, and PIK3CA mutations on the efficacy of cetuximab plus chemotherapy in chemotherapy-refractory metastatic colorectal cancer: a retrospective consortium analysis. Lancet Oncol. 2010 Aug;11(8):753-62. doi: 10.1016/S1470-2045(10)70130-3. Epub 2010 Jul 8. PMID: 20619739.
23. Harries M, Smith I. The development and clinical use of trastuzumab (Herceptin). Endocr Relat Cancer. 2002 Jun;9(2):75-85. doi: 10.1677/erc.0.0090075. PMID: 12121832.
24. Winter G, Milstein C. Man-made antibodies. Nature. 1991 Jan 24;349(6307):293-9. doi: 10.1038/349293a0. PMID: 1987490.

25. Hurt CN, Nixon LS, Griffiths GO, Al-Mokhtar R, Gollins S, Staffurth JN, Phillips CJ, Blazeby JM, Crosby TD. SCOPE1: a randomised phase II/III multicentre clinical trial of definitive chemoradiation, with or without cetuximab, in carcinoma of the oesophagus. BMC Cancer. 2011 Oct 28;11:466. doi: 10.1186/1471-2407-11-466. PMID: 22035459; PMCID: PMC3212828.

26. Rodriguez-Berriguete G, Ranzani M, Prevo R, Puliyadi R, Machado N, Bolland HR, Millar V, Ebner D, Boursier M, Cerutti A, Cicconi A, Galbiati A, Grande D, Grinkevich V, Majithiya JB, Piscitello D, Rajendra E, Stockley ML, Boulton SJ, Hammond EM, Heald RA, Smith GCM, Robinson HMR, Higgins GS. Small-Molecule Polθ Inhibitors Provide Safe and Effective Tumor Radiosensitization in Preclinical Models. Clin Cancer Res. 2023 Apr 14;29(8):1631-1642. doi: 10.1158/1078-0432.CCR-22-2977. PMID: 36689546; PMCID: PMC10102842.

27. Skwarski M, McGowan DR, Belcher E, Di Chiara F, Stavroulias D, McCole M, Derham JL, Chu KY, Teoh E, Chauhan J, O'Reilly D, Harris BHL, Macklin PS, Bull JA, Green M, Rodriguez-Berriguete G, Prevo R, Folkes LK, Campo L, Ferencz P, Croal PL, Flight H, Qi C, Holmes J, O'Connor JPB, Gleeson FV, McKenna WG, Harris AL, Bulte D, Buffa FM, Macpherson RE, Higgins GS. Mitochondrial Inhibitor Atovaquone Increases Tumor Oxygenation and Inhibits Hypoxic Gene Expression in Patients with Non-Small Cell Lung Cancer. Clin Cancer Res. 2021 May 1;27(9):2459-2469. doi: 10.1158/1078-0432.CCR-20-4128. Epub 2021 Feb 17. PMID: 33597271; PMCID: PMC7611473.

28. Buffa FM, Camps C, Winchester L, Snell CE, Gee HE, Sheldon H, Taylor M, Harris AL, Ragoussis J. microRNA-associated progression pathways and potential therapeutic targets identified by integrated mRNA and microRNA expression profiling in breast cancer. Cancer Res. 2011 Sep 1;71(17):5635-45. doi: 10.1158/0008-5472.CAN-11-0489. Epub 2011 Jul 7. PMID: 21737487.

29. Dhawan A, Barberis A, Cheng WC, Domingo E, West C, Maughan T, Scott JG, Harris AL, Buffa FM. Guidelines for using sigQC for systematic evaluation of gene signatures. Nat Protoc. 2019 May;14(5):1377-1400. doi: 10.1038/s41596-019-0136-8. Epub 2019 Apr 10. PMID: 30971781.
30. Masca NG, Hensor EM, Cornelius VR, Buffa FM, Marriott HM, Eales JM, Messenger MP, Anderson AE, Boot C, Bunce C, Goldin RD, Harris J, Hinchliffe RF, Junaid H, Kingston S, Martin-Ruiz C, Nelson CP, Peacock J, Seed PT, Shinkins B, Staples KJ, Toombs J, Wright AK, Teare MD. RIPOSTE: a framework for improving the design and analysis of laboratory-based research. Elife. 2015 May 7;4:e05519. doi: 10.7554/eLife.05519. PMID: 25951517; PMCID: PMC4461852.
31. Perou CM, Sørlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, Pollack JR, Ross DT, Johnsen H, Akslen LA, Fluge O, Pergamenschikov A, Williams C, Zhu SX, Lønning PE, Børresen-Dale AL, Brown PO, Botstein D. Molecular portraits of human breast tumours. Nature. 2000 Aug 17;406(6797):747-52. doi: 10.1038/35021093. PMID: 10963602.
32. Sorlie T, Tibshirani R, Parker J, Hastie T, Marron JS, Nobel A, Deng S, Johnsen H, Pesich R, Geisler S, Demeter J, Perou CM, Lønning PE, Brown PO, Børresen-Dale AL, Botstein D. Repeated observation of breast tumor subtypes in independent gene expression data sets. Proc Natl Acad Sci U S A. 2003 Jul 8;100(14):8418-23. doi: 10.1073/pnas.0932692100. Epub 2003 Jun 26. PMID: 12829800; PMCID: PMC166244.
33. de Ronde J, Wessels L, Wesseling J. Molecular subtyping of breast cancer: ready to use? Lancet Oncol. 2010 Apr;11(4):306-7. doi: 10.1016/S1470-2045(10)70036-X. PMID: 20359657.
34. Weigelt B, Mackay A, A'hern R, Natrajan R, Tan DS, Dowsett M, Ashworth A, Reis-Filho JS. Breast cancer molecular profiling with single sample predictors: a retrospective analysis. Lancet Oncol. 2010 Apr;11(4):339-49. doi: 10.1016/S1470-2045(10)70008-5. Epub 2010 Feb 22. PMID: 20181526.

35. Raj-Kumar PK, Liu J, Hooke JA, Kovatich AJ, Kvecher L, Shriver CD, Hu H. PCA-PAM50 improves consistency between breast cancer intrinsic and clinical subtyping reclassifying a subset of luminal A tumors as luminal B. Sci Rep. 2019 May 28;9(1):7956. doi: 10.1038/s41598-019-44339-4. PMID: 31138829; PMCID: PMC6538748.
36. Guinney J, Dienstmann R, Wang X, de Reyniès A, Schlicker A, Soneson C, Marisa L, Roepman P, Nyamundanda G, Angelino P, Bot BM, Morris JS, Simon IM, Gerster S, Fessler E, De Sousa E Melo F, Missiaglia E, Ramay H, Barras D, Homicsko K, Maru D, Manyam GC, Broom B, Boige V, Perez-Villamil B, Laderas T, Salazar R, Gray JW, Hanahan D, Tabernero J, Bernards R, Friend SH, Laurent-Puig P, Medema JP, Sadanandam A, Wessels L, Delorenzi M, Kopetz S, Vermeulen L, Tejpar S. The consensus molecular subtypes of colorectal cancer. Nat Med. 2015 Nov;21(11):1350-6. doi: 10.1038/nm.3967. Epub 2015 Oct 12. PMID: 26457759; PMCID: PMC4636487.
37. Lenz HJ, Ou FS, Venook AP, Hochster HS, Niedzwiecki D, Goldberg RM, Mayer RJ, Bertagnolli MM, Blanke CD, Zemla T, Qu X, Wirapati P, Tejpar S, Innocenti F, Kabbarah O. Impact of Consensus Molecular Subtype on Survival in Patients With Metastatic Colorectal Cancer: Results From CALGB/SWOG 80405 (Alliance). J Clin Oncol. 2019 Aug 1;37(22):1876-1885. doi: 10.1200/JCO.18.02258. Epub 2019 May 1. Erratum in: J Clin Oncol. 2020 Feb 20;38(6):656. doi: 10.1200/JCO.20.00020. PMID: 31042420; PMCID: PMC6675593.
38. Fontana E, Eason K, Cervantes A, Salazar R, Sadanandam A. Context matters-consensus molecular subtypes of colorectal cancer as biomarkers for clinical trials. Ann Oncol. 2019 Apr 1;30(4):520-527. doi: 10.1093/annonc/mdz052. PMID: 30796810; PMCID: PMC6503627.
39. Alderdice M, Richman SD, Gollins S, Stewart JP, Hurt C, Adams R, McCorry AM, Roddy AC, Vimalachandran D, Isella C, Medico E, Maughan T, McArt DG, Lawler M, Dunne PD. Prospective patient stratification into robust cancer-cell intrinsic subtypes

from colorectal cancer biopsies. J Pathol. 2018 May;245(1):19-28. doi: 10.1002/path.5051. Epub 2018 Mar 25. PMID: 29412457; PMCID: PMC5947827.

40. Harris BH, Barberis A, West CM, Buffa FM. Gene Expression Signatures as Biomarkers of Tumour Hypoxia. Clin Oncol (R Coll Radiol). 2015 Oct;27(10):547-60. doi: 10.1016/j.clon.2015.07.004. Epub 2015 Aug 14. PMID: 26282471.
41. Favaro E, Lord S, Harris AL, Buffa FM. Gene expression and hypoxia in breast cancer. Genome Med. 2011 Aug 26;3(8):55. doi: 10.1186/gm271. PMID: 21875443; PMCID: PMC3238181.
42. Bhandari V, Hoey C, Liu LY, Lalonde E, Ray J, Livingstone J, Lesurf R, Shiah YJ, Vujcic T, Huang X, Espiritu SMG, Heisler LE, Yousif F, Huang V, Yamaguchi TN, Yao CQ, Sabelnykova VY, Fraser M, Chua MLK, van der Kwast T, Liu SK, Boutros PC, Bristow RG. Molecular landmarks of tumor hypoxia across cancer types. Nat Genet. 2019 Feb;51(2):308-318. doi: 10.1038/s41588-018-0318-2. Epub 2019 Jan 14. PMID: 30643250.
43. Buffa FM, Harris AL, West CM, Miller CJ. Large meta-analysis of multiple cancers reveals a common, compact and highly prognostic hypoxia metagene. Br J Cancer. 2010 Jan 19;102(2):428-35. doi: 10.1038/sj.bjc.6605450. Erratum in: Br J Cancer. 2010 Sep 7;103(6):929. PMID: 20087356; PMCID: PMC2816644.
44. Piskol R, Huw L, Sergin I, Kljin C, Modrusan Z, Kim D, Kljavin N, Tam R, Patel R, Burton J, Penuel E, Qu X, Koeppen H, Sumiyoshi T, de Sauvage F, Lackner MR, de Sousa E Melo F, Kabbarah O. A Clinically Applicable Gene-Expression Classifier Reveals Intrinsic and Extrinsic Contributions to Consensus Molecular Subtypes in Primary and Metastatic Colon Cancer. Clin Cancer Res. 2019 Jul 15;25(14):4431-4442. doi: 10.1158/1078-0432.CCR-18-3032. Epub 2019 Apr 19. PMID: 31004000.
45. National Institute for Clinical Excellence, (2018) Tumour profiling tests to guide adjuvant chemotherapy decisions in early breast cancer Diagnostics guidance [DG34] Department of Health, Published date: December 2018.

46. Sparano JA, Gray RJ, Makower DF, Albain KS, Saphner TJ, Badve SS, Wagner LI, Kaklamani VG, Keane MM, Gomez HL, Reddy PS, Goggins TF, Mayer IA, Toppmeyer DL, Brufsky AM, Goetz MP, Berenberg JL, Mahalcioiu C, Desbiens C, Hayes DF, Dees EC, Geyer CE Jr, Olson JA Jr, Wood WC, Lively T, Paik S, Ellis MJ, Abrams J, Sledge GW Jr. Clinical Outcomes in Early Breast Cancer With a High 21-Gene Recurrence Score of 26 to 100 Assigned to Adjuvant Chemotherapy Plus Endocrine Therapy: A Secondary Analysis of the TAILORx Randomized Clinical Trial. JAMA Oncol. 2020 Mar 1;6(3):367-374. doi: 10.1001/jamaoncol.2019.4794. PMID: 31566680; PMCID: PMC6777230.
47. Woodward WA, Barlow WE, Jagsi R, Buchholz TA, Shak S, Baehner F, Whelan TJ, Davidson NE, Ingle JN, King TA, Ravdin PM, Osborne CK, Tripathy D, Livingston RB, Gralow JR, Hortobagyi GN, Hayes DF, Albain KS. Association Between 21-Gene Assay Recurrence Score and Locoregional Recurrence Rates in Patients With Node-Positive Breast Cancer. JAMA Oncol. 2020 Apr 1;6(4):505-511. doi: 10.1001/jamaoncol.2019.5559. PMID: 31917424; PMCID: PMC6990911.
48. Eschrich SA, Pramana J, Zhang H, Zhao H, Boulware D, Lee JH, Bloom G, Rocha-Lima C, Kelley S, Calvin DP, Yeatman TJ, Begg AC, Torres-Roca JF. A gene expression model of intrinsic tumor radiosensitivity: prediction of response and prognosis after chemoradiation. Int J Radiat Oncol Biol Phys. 2009 Oct 1;75(2):489-96. doi: 10.1016/j.ijrobp.2009.06.014. PMID: 19735873; PMCID: PMC3038688.
49. Scott JG, Berglund A, Schell MJ, Mihaylov I, Fulp WJ, Yue B, Welsh E, Caudell JJ, Ahmed K, Strom TS, Mellon E, Venkat P, Johnstone P, Foekens J, Lee J, Moros E, Dalton WS, Eschrich SA, McLeod H, Harrison LB, Torres-Roca JF. A genome-based model for adjusting radiotherapy dose (GARD): a retrospective, cohort-based study. Lancet Oncol. 2017 Feb;18(2):202-211. doi: 10.1016/S1470-2045(16)30648-9. Epub 2016 Dec 18. Erratum in: Lancet Oncol. 2017 Feb;18(2):e65. doi: 10.1016/S1470-2045(16)30686-6. PMID: 27993569; PMCID: PMC7771305.

50. Yang L, Taylor J, Eustace A, Irlam JJ, Denley H, Hoskin PJ, Alsner J, Buffa FM, Harris AL, Choudhury A, West CML. A Gene Signature for Selecting Benefit from Hypoxia Modification of Radiotherapy for High-Risk Bladder Cancer Patients. Clin Cancer Res. 2017 Aug 15;23(16):4761-4768. doi: 10.1158/1078-0432.CCR-17-0038. Epub 2017 Apr 11. PMID: 28400426.
51. Brooks JM, Menezes AN, Ibrahim M, Archer L, Lal N, Bagnall CJ, von Zeidler SV, Valentine HR, Spruce RJ, Batis N, Bryant JL, Hartley M, Kaul B, Ryan GB, Bao R, Khattri A, Lee SP, Ogbureke KUE, Middleton G, Tennant DA, Beggs AD, Deeks J, West CML, Cazier JB, Willcox BE, Seiwert TY, Mehanna H. Development and Validation of a Combined Hypoxia and Immune Prognostic Classifier for Head and Neck Cancer. Clin Cancer Res. 2019 Sep 1;25(17):5315-5328. doi: 10.1158/1078-0432.CCR-18-3314. Epub 2019 Jun 10. PMID: 31182433.
52. Sjöström M, Chang SL, Fishbane N, Davicioni E, Zhao SG, Hartman L, Holmberg E, Feng FY, Speers CW, Pierce LJ, Malmström P, Fernö M, Karlsson P. Clinicogenomic Radiotherapy Classifier Predicting the Need for Intensified Locoregional Treatment After Breast-Conserving Surgery for Early-Stage Breast Cancer. J Clin Oncol. 2019 Dec 10;37(35):3340-3349. doi: 10.1200/JCO.19.00761. Epub 2019 Oct 16. PMID: 31618132; PMCID: PMC6901281.
53. Marascio J, Spratt DE, Zhang J, Trabulsi EJ, Le T, Sedzorme WS, Beeler WH, Davicioni E, Dabbas B, Lin DW, Gore JL, Bloom M, Mann M, Mark JR, Calvaresi A, Godwin JL, McCue P, Hurwitz MD, Kelly WK, Lallas CD, Knudsen KE, Gomella LG, Dicker AP, Den RB. Prospective study to define the clinical utility and benefit of Decipher testing in men following prostatectomy. Prostate Cancer Prostatic Dis. 2020 Jun;23(2):295-302. doi: 10.1038/s41391-019-0185-7. Epub 2019 Nov 12. PMID: 31719663; PMCID: PMC7237345.
54. Van den Broeck T, Moris L, Gevaert T, Tosco L, Smeets E, Fishbane N, Liu Y, Helsen C, Margrave J, Buerki C, Davicioni E, Van Poppel H, Everaerts W, Weinmann S, Den R, Davis J, Schaeffer E, Karnes RJ, Claessens F, Joniau S. Validation of the Decipher Test for Predicting Distant Metastatic Recurrence in Men with High-risk

Nonmetastatic Prostate Cancer 10 Years After Surgery. Eur Urol Oncol. 2019 Sep;2(5):589-596. doi: 10.1016/j.euo.2018.12.007. Epub 2019 Jan 11. PMID: 31411980.

55. Zhao SG, Chang SL, Spratt DE, Erho N, Yu M, Ashab HA, Alshalalfa M, Speers C, Tomlins SA, Davicioni E, Dicker AP, Carroll PR, Cooperberg MR, Freedland SJ, Karnes RJ, Ross AE, Schaeffer EM, Den RB, Nguyen PL, Feng FY. Development and validation of a 24-gene predictor of response to postoperative radiotherapy in prostate cancer: a matched, retrospective analysis. Lancet Oncol. 2016 Nov;17(11):1612-1620. doi: 10.1016/S1470-2045(16)30491-0. Epub 2016 Oct 12. PMID: 27743920.

56. Mistry HB. Radiosensitivity Index is Not Fit to be Used for Dose Adjustments: A Pan-Cancer Analysis. Clin Oncol (R Coll Radiol). 2023 Sep;35(9):565-570. doi: 10.1016/j.clon.2023.02.018. Epub 2023 Mar 8. PMID: 36922240.

57. Alfonso JCL, Berk L. Modeling the effect of intratumoral heterogeneity of radiosensitivity on tumor response over the course of fractionated radiation therapy. Radiat Oncol. 2019 May 30;14(1):88. doi: 10.1186/s13014-019-1288-y. PMID: 31146751; PMCID: PMC6543639.

58. Jamal-Hanjani M, Wilson GA, McGranahan N, Birkbak NJ, Watkins TBK, Veeriah S, Shafi S, Johnson DH, Mitter R, Rosenthal R, Salm M, Horswell S, Escudero M, Matthews N, Rowan A, Chambers T, Moore DA, Turajlic S, Xu H, Lee SM, Forster MD, Ahmad T, Hiley CT, Abbosh C, Falzon M, Borg E, Marafioti T, Lawrence D, Hayward M, Kolvekar S, Panagiotopoulos N, Janes SM, Thakrar R, Ahmed A, Blackhall F, Summers Y, Shah R, Joseph L, Quinn AM, Crosbie PA, Naidu B, Middleton G, Langman G, Trotter S, Nicolson M, Remmen H, Kerr K, Chetty M, Gomersall L, Fennell DA, Nakas A, Rathinam S, Anand G, Khan S, Russell P, Ezhil V, Ismail B, Irvin-Sellers M, Prakash V, Lester JF, Kornaszewska M, Attanoos R, Adams H, Davies H, Dentro S, Taniere P, O'Sullivan B, Lowe HL, Hartley JA, Iles N,

Bell H, Ngai Y, Shaw JA, Herrero J, Szallasi Z, Schwarz RF, Stewart A, Quezada SA, Le Quesne J, Van Loo P, Dive C, Hackshaw A, Swanton C; TRACERx Consortium. Tracking the Evolution of Non-Small-Cell Lung Cancer. N Engl J Med. 2017 Jun 1;376(22):2109-2121. doi: 10.1056/NEJMoa1616288. Epub 2017 Apr 26. PMID: 28445112.

59. Li BT, Janku F, Jung B, Hou C, Madwani K, Alden R, Razavi P, Reis-Filho JS, Shen R, Isbell JM, Blocker AW, Eattock N, Gnerre S, Satya RV, Xu H, Zhao C, Hall MP, Hu Y, Sehnert AJ, Brown D, Ladanyi M, Rudin CM, Hunkapiller N, Feeney N, Mills GB, Paweletz CP, Janne PA, Solit DB, Riely GJ, Aravanis A, Oxnard GR. Ultra-deep next-generation sequencing of plasma cell-free DNA in patients with advanced lung cancers: results from the Actionable Genome Consortium. Ann Oncol. 2019 Apr 1;30(4):597-603. doi: 10.1093/annonc/mdz046. PMID: 30891595; PMCID: PMC6503621..

60. McDonald BR, Contente-Cuomo T, Sammut SJ, Odenheimer-Bergman A, Ernst B, Perdigones N, Chin SF, Farooq M, Mejia R, Cronin PA, Anderson KS, Kosiorek HE, Northfelt DW, McCullough AE, Patel BK, Weitzel JN, Slavin TP, Caldas C, Pockaj BA, Murtaza M. Personalized circulating tumor DNA analysis to detect residual disease after neoadjuvant therapy in breast cancer. Sci Transl Med. 2019 Aug 7;11(504):eaax7392. doi: 10.1126/scitranslmed.aax7392. PMID: 31391323; PMCID: PMC7236617.

61. Razavi P, Li BT, Brown DN, Jung B, Hubbell E, Shen R, Abida W, Juluru K, De Bruijn I, Hou C, Venn O, Lim R, Anand A, Maddala T, Gnerre S, Vijaya Satya R, Liu Q, Shen L, Eattock N, Yue J, Blocker AW, Lee M, Sehnert A, Xu H, Hall MP, Santiago-Zayas A, Novotny WF, Isbell JM, Rusch VW, Plitas G, Heerdt AS, Ladanyi M, Hyman DM, Jones DR, Morrow M, Riely GJ, Scher HI, Rudin CM, Robson ME, Diaz LA Jr, Solit DB, Aravanis AM, Reis-Filho JS. High-intensity sequencing reveals the sources of plasma circulating cell-free DNA variants. Nat Med. 2019 Dec;25(12):1928-1937.

doi: 10.1038/s41591-019-0652-7. Epub 2019 Nov 25. PMID: 31768066; PMCID: PMC7061455.

62. Rostami A, Lambie M, Yu CW, Stambolic V, Waldron JN, Bratman SV. Senescence, Necrosis, and Apoptosis Govern Circulating Cell-free DNA Release Kinetics. Cell Rep. 2020 Jun 30;31(13):107830. doi: 10.1016/j.celrep.2020.107830. PMID: 32610131.
63. Jackson HW, Fischer JR, Zanotelli VRT, Ali HR, Mechera R, Soysal SD, Moch H, Muenst S, Varga Z, Weber WP, Bodenmiller B. The single-cell pathology landscape of breast cancer. Nature. 2020 Feb;578(7796):615-620. doi: 10.1038/s41586-019-1876-x. Epub 2020 Jan 20. PMID: 31959985.
64. Ahn SJ, Choi C, Choi YD, Kim YC, Kim KS, Oh IJ, Ban HJ, Yoon MS, Nam TK, Jeong JU, Song JY, Chung WK. Microarray analysis of gene expression in lung cancer cell lines treated by fractionated irradiation. Anticancer Res. 2014 Sep;34(9):4939-48. PMID: 25202076.
65. Chen X, Zheng J, Zhuo ML, Zhang A, You Z. A six-gene-based signature for breast cancer radiotherapy sensitivity estimation. Biosci Rep. 2020 Dec 23;40(12):BSR20202376. doi: 10.1042/BSR20202376. PMID: 33179733; PMCID: PMC7711058.
66. Li Z, Ye L, Wang L, Quan R, Zhou Y, Li X. Identification of miRNA signatures in serum exosomes as a potential biomarker after radiotherapy treatment in glioma patients. Ann Diagn Pathol. 2020 Feb;44:151436. doi: 10.1016/j.anndiagpath.2019.151436. Epub 2019 Dec 11. PMID: 31865249.
67. Bosma SCJ, Hoogstraat M, van der Leij F, de Maaker M, Wesseling J, Lips E, Loo CE, Rutgers EJ, Elkhuizen PHM, Bartelink H, van de Vijver MJ. Response to Preoperative Radiation Therapy in Relation to Gene Expression Patterns in Breast Cancer Patients. Int J Radiat Oncol Biol Phys. 2020 Jan 1;106(1):174-181. doi: 10.1016/j.ijrobp.2019.09.002. Epub 2019 Sep 13. PMID: 31525407.

68. Liu J, Han M, Yue Z, Dong C, Wen P, Zhao G, Wu L, Xia J, Bin Y. Prediction of Radiosensitivity in Head and Neck Squamous Cell Carcinoma Based on Multiple Omics Data. Front Genet. 2020 Aug 18;11:960. doi: 10.3389/fgene.2020.00960. PMID: 33014019; PMCID: PMC7461877.
69. Ahmad P, Slavik M, Trachtova K, Gablo NA, Kazda T, Gurin D, Smilek P, Horakova Z, Gal B, Hermanova M, Slampa P, Sana J, Slaby O. Salivary microRNAs identified by small RNA sequencing as potential predictors of response to intensity-modulated radiotherapy in head and neck cancer patients. Cell Oncol (Dordr). 2020 Jun;43(3):505-511. doi: 10.1007/s13402-020-00507-7. Epub 2020 Apr 7. PMID: 32266559.
70. Knoops L, Haas R, de Kemp S, Majoor D, Broeks A, Eldering E, de Boer JP, Verheij M, van Ostrom C, de Vries A, van't Veer L, de Jong D. In vivo p53 response and immune reaction underlie highly effective low-dose radiotherapy in follicular lymphoma. Blood. 2007 Aug 15;110(4):1116-22. doi: 10.1182/blood-2007-01-067579. Epub 2007 May 4. PMID: 17483295.
71. Liu T, Du X, Sheng X. Genetic alterations following ionizing radiation in human ovarian cancer-derived endothelial cells. Mol Med Rep. 2014 Jun;9(6):2257-64. doi: 10.3892/mmr.2014.2096. Epub 2014 Mar 31. PMID: 24691555.
72. Smirnov DA, Brady L, Halasa K, Morley M, Solomon S, Cheung VG. Genetic variation in radiation-induced cell death. Genome Res. 2012 Feb;22(2):332-9. doi: 10.1101/gr.122044.111. Epub 2011 Aug 15. PMID: 21844125; PMCID: PMC3266040.
73. Supiot S, Gouraud W, Campion L, Jezéquel P, Buecher B, Charrier J, Heymann MF, Mahé MA, Rio E, Chérel M. Early dynamic transcriptomic changes during preoperative radiotherapy in patients with rectal cancer: a feasibility study. World J Gastroenterol. 2013 Jun 7;19(21):3249-54. doi: 10.3748/wjg.v19.i21.3249. PMID: 23745026; PMCID: PMC3671076.

74. Tsujiguchi T, Hirouchi T, Monzen S, Tabuchi Y, Takasaki I, Kondo T, Kashiwakura I. Expression analysis of radiation-responsive genes in human hematopoietic stem/progenitor cells. J Radiat Res. 2016 Jan;57(1):35-43. doi: 10.1093/jrr/rrv071. Epub 2015 Dec 9. PMID: 26661850; PMCID: PMC4708922.
75. Weidhaas JB, Li SX, Winter K, Ryu J, Jhingran A, Miller B, Dicker AP, Gaffney D. Changes in gene expression predicting local control in cervical cancer: results from Radiation Therapy Oncology Group 0128. Clin Cancer Res. 2009 Jun 15;15(12):4199-206. doi: 10.1158/1078-0432.CCR-08-2257. Epub 2009 Jun 9. PMID: 19509178; PMCID: PMC2758917.
76. Zempolich K, Fuhrman C, Milash B, Flinner R, Greven K, Ryu J, Forbes A, Kerlin K, Nichols RC, Gaffney DK. Changes in gene expression induced by chemoradiation in advanced cervical carcinoma: a microarray study of RTOG C-0128. Gynecol Oncol. 2008 May;109(2):275-9. doi: 10.1016/j.ygyno.2008.01.027. Epub 2008 Mar 4. PMID: 18299147.
77. Zhou J, Chen C, Li HF, Hu YJ, Xie HL. Revealing radiotherapy- and chemoradiation-induced pathway dynamics in glioblastoma by analyzing multiple differential networks. Mol Med Rep. 2017 Jul;16(1):696-702. doi: 10.3892/mmr.2017.6641. Epub 2017 May 29. PMID: 28560382; PMCID: PMC5482131.
78. Chen L, Ren P, Zhang Y, Gong B, Yu D, Sun X. Long non-coding RNA GAS5 increases the radiosensitivity of A549 cells through interaction with the miR-21/PTEN/Akt axis. Oncol Rep. 2020 Mar;43(3):897-907. doi: 10.3892/or.2020.7467. Epub 2020 Jan 15. PMID: 32020207; PMCID: PMC7041238.
79. Kim SI, Kang JW, Noh JK, Jung HR, Lee YC, Lee JW, Kong M, Eun YG. Gene signature for prediction of radiosensitivity in human papillomavirus-negative head and neck squamous cell carcinoma. Radiat Oncol J. 2020 Jun;38(2):99-108. doi: 10.3857/roj.2020.00136. Epub 2020 Jun 9. PMID: 33012153; PMCID: PMC7533413.

80. Khan MT, Yang L, More E, Irlam-Jones JJ, Valentine HR, Hoskin P, Choudhury A, West CML. Developing Tumor Radiosensitivity Signatures Using LncRNAs. Radiat Res. 2021 Apr 1;195(4):324-333. doi: 10.1667/RADE-20-00157.1. PMID: 33577642.
81. Liu W, Yu Z, Tang H, Wang X, Zhang B, Zhao J, Liu X, Zhang J, Wei M. Silencing KIF18B enhances radiosensitivity: identification of a promising therapeutic target in sarcoma. EBioMedicine. 2020 Nov;61:103056. doi: 10.1016/j.ebiom.2020.103056. Epub 2020 Oct 7. Erratum in: EBioMedicine. 2021 Dec;74:103710. doi: 10.1016/j.ebiom.2021.103710. PMID: 33038765; PMCID: PMC7648128.
82. Ahmed FE, Vos PW, Jeffries C, Wiley JE, Weidner DA, Mota H, Bonnerup C, Sibata C, Allison RR. Differences in mRNA and microRNA microarray expression profiles in human colon adenocarcinoma HT-29 cells treated with either Intensity-modulated Radiation Therapy (IMRT), or Conventional Radiation Therapy (RT). Cancer Genomics Proteomics. 2009 Mar-Apr;6(2):109-27. PMID: 19451095.
83. Aryankalayil MJ, Makinde AY, Gameiro SR, Hodge JW, Rivera-Solis PP, Palayoor ST, Ahmed MM, Coleman CN. Defining molecular signature of pro-immunogenic radiotherapy targets in human prostate cancer cells. Radiat Res. 2014 Aug;182(2):139-48. doi: 10.1667/RR13731.1. Epub 2014 Jul 8. PMID: 25003313; PMCID: PMC4216662.
84. Bassi C, Mello SS, Cardoso RS, Godoy PD, Fachin AL, Junta CM, Sandrin-Garcia P, Carlotti CG, Falcão RP, Donadi EA, Passos GA, Sakamoto-Hojo ET. Transcriptional changes in U343 MG-a glioblastoma cell line exposed to ionizing radiation. Hum Exp Toxicol. 2008 Dec;27(12):919-29. doi: 10.1177/0960327108102045. PMID: 19273547.
85. Ding LH, Shingyoji M, Chen F, Hwang JJ, Burma S, Lee C, Cheng JF, Chen DJ. Gene expression profiles of normal human fibroblasts after exposure to ionizing radiation: a comparative study of low and high doses. Radiat Res. 2005 Jul;164(1):17-26. doi: 10.1667/rr3354. PMID: 15966761.

86. Kis E, Szatmári T, Keszei M, Farkas R, Esik O, Lumniczky K, Falus A, Sáfrány G. Microarray analysis of radiation response genes in primary human fibroblasts. Int J Radiat Oncol Biol Phys. 2006 Dec 1;66(5):1506-14. doi: 10.1016/j.ijrobp.2006.08.004. Epub 2006 Oct 25. PMID: 17069989.
87. Ma H, Rao L, Wang HL, Mao ZW, Lei RH, Yang ZY, Qing H, Deng YL. Transcriptome analysis of glioma cells for the dynamic response to γ-irradiation and dual regulation of apoptosis genes: a new insight into radiotherapy for glioblastomas. Cell Death Dis. 2013 Oct 31;4(10):e895. doi: 10.1038/cddis.2013.412. PMID: 24176853; PMCID: PMC3920930.
88. Park WY, Hwang CI, Im CN, Kang MJ, Woo JH, Kim JH, Kim YS, Kim JH, Kim H, Kim KA, Yu HJ, Lee SJ, Lee YS, Seo JS. Identification of radiation-specific responses from gene expression profile. Oncogene. 2002 Dec 5;21(55):8521-8. doi: 10.1038/sj.onc.1205977. PMID: 12466973.
89. Otomo T, Hishii M, Arai H, Sato K, Sasai K. Microarray analysis of temporal gene responses to ionizing radiation in two glioblastoma cell lines: up-regulation of DNA repair genes. J Radiat Res. 2004 Mar;45(1):53-60. doi: 10.1269/jrr.45.53. PMID: 15133290.
90. Kim HS, Kim SC, Kim SJ, Park CH, Jeung HC, Kim YB, Ahn JB, Chung HC, Rha SY. Identification of a radiosensitivity signature using integrative metaanalysis of published microarray data for NCI-60 cancer cells. BMC Genomics. 2012 Jul 30;13:348. doi: 10.1186/1471-2164-13-348. PMID: 22846430; PMCID: PMC3472294.
91. Speers C, Zhao S, Liu M, Bartelink H, Pierce LJ, Feng FY. Development and Validation of a Novel Radiosensitivity Signature in Human Breast Cancer. Clin Cancer Res. 2015 Aug 15;21(16):3667-77. doi: 10.1158/1078-0432.CCR-14-2898. Epub 2015 Apr 22. PMID: 25904749.
92. Spitzner M, Emons G, Kramer F, Gaedcke J, Rave-Fränk M, Scharf JG, Burfeind P, Becker H, Beissbarth T, Ghadimi BM, Ried T, Grade M. A gene expression signature

for chemoradiosensitivity of colorectal cancer cells. Int J Radiat Oncol Biol Phys. 2010 Nov 15;78(4):1184-92. doi: 10.1016/j.ijrobp.2010.06.023. PMID: 20970032; PMCID: PMC7339824.

93. Torres-Roca JF, Eschrich S, Zhao H, Bloom G, Sung J, McCarthy S, Cantor AB, Scuto A, Li C, Zhang S, Jove R, Yeatman T. Prediction of radiation sensitivity using a gene expression classifier. Cancer Res. 2005 Aug 15;65(16):7169-76. doi: 10.1158/0008-5472.CAN-05-0656. PMID: 16103067.
94. Yard BD, Adams DJ, Chie EK, Tamayo P, Battaglia JS, Gopal P, Rogacki K, Pearson BE, Phillips J, Raymond DP, Pennell NA, Almeida F, Cheah JH, Clemons PA, Shamji A, Peacock CD, Schreiber SL, Hammerman PS, Abazeed ME. A genetic basis for the variation in the vulnerability of cancer to DNA damage. Nat Commun. 2016 Apr 25;7:11428. doi: 10.1038/ncomms11428. PMID: 27109210; PMCID: PMC4848553.
95. Zhang C, Girard L, Das A, Chen S, Zheng G, Song K. Nonlinear quantitative radiation sensitivity prediction model based on NCI-60 cancer cell lines. ScientificWorldJournal. 2014;2014:903602. doi: 10.1155/2014/903602. Epub 2014 Jun 17. PMID: 25032244; PMCID: PMC4083270.
96. Ishigami T, Uzawa K, Higo M, Nomura H, Saito K, Kato Y, Nakashima D, Shiiba M, Bukawa H, Yokoe H, Kawata T, Ito H, Tanzawa H. Genes and molecular pathways related to radioresistance of oral squamous cell carcinoma cells. Int J Cancer. 2007 May 15;120(10):2262-70. doi: 10.1002/ijc.22561. PMID: 17290400.
97. Chaudhry MA, Chodosh LA, McKenna WG, Muschel RJ. Gene expression profile of human cells irradiated in G1 and G2 phases of cell cycle. Cancer Lett. 2003 Jun 10;195(2):221-33. doi: 10.1016/s0304-3835(03)00154-x. PMID: 12767531.
98. Du XL, Jiang T, Wen ZQ, Li QS, Gao R, Wang F. Differential expression profiling of gene response to ionizing radiation in two endometrial cancer cell lines with distinct radiosensitivities. Oncol Rep. 2009 Mar;21(3):625-34. PMID: 19212620.
99. Zhao SG, Yu M, Spratt DE, Chang SL, Feng FY, Kim MM, Speers CW, Carlson BL, Mladek AC, Lawrence TS, Sarkaria JN, Wahl DR. Xenograft-based, platform-

independent gene signatures to predict response to alkylating chemotherapy, radiation, and combination therapy for glioblastoma. Neuro Oncol. 2019 Sep 6;21(9):1141-1149. doi: 10.1093/neuonc/noz090. PMID: 31121035; PMCID: PMC6736132.

100. Liu G, Zeng X, Wu B, Zhao J, Pan Y. RNA-Seq analysis of peripheral blood mononuclear cells reveals unique transcriptional signatures associated with radiotherapy response of nasopharyngeal carcinoma and prognosis of head and neck cancer. Cancer Biol Ther. 2020;21(2):139-146. doi: 10.1080/15384047.2019.1670521. Epub 2019 Nov 7. PMID: 31698994; PMCID: PMC7012055.

101. Sjöström M, Chang SL, Fishbane N, Davicioni E, Zhao SG, Hartman L, Holmberg E, Feng FY, Speers CW, Pierce LJ, Malmström P, Fernö M, Karlsson P. Clinicogenomic Radiotherapy Classifier Predicting the Need for Intensified Locoregional Treatment After Breast-Conserving Surgery for Early-Stage Breast Cancer. J Clin Oncol. 2019 Dec 10;37(35):3340-3349. doi: 10.1200/JCO.19.00761. Epub 2019 Oct 16. PMID: 31618132; PMCID: PMC6901281.

102. Li K, Zhu X, Li L, Ning R, Liang Z, Zeng F, Su F, Huang S, Yang X, Qu S. Identification of non-invasive biomarkers for predicting the radiosensitivity of nasopharyngeal carcinoma from serum microRNAs. Sci Rep. 2020 Mar 20;10(1):5161. doi: 10.1038/s41598-020-61958-4. PMID: 32198434; PMCID: PMC7083955.

103. Zhang Q, Bing Z, Tian J, Wang X, Liu R, Li Y, Kong Y, Yang Y. Integrating radiosensitive genes improves prediction of radiosensitivity or radioresistance in patients with oesophageal cancer. Oncol Lett. 2019 Jun;17(6):5377-5388. doi: 10.3892/ol.2019.10240. Epub 2019 Apr 10. PMID: 31186755; PMCID: PMC6507505.

104. Naghavi AO, Kim Y, Yang GQ, Ahmed KA, Caudell JJ. Alterations in genetic pathways following radiotherapy for head and neck cancer. Head Neck. 2020

Feb;42(2):312-320. doi: 10.1002/hed.26004. Epub 2019 Dec 13. PMID: 31833149; PMCID: PMC7771332.

105. Amundson SA, Grace MB, McLeland CB, Epperly MW, Yeager A, Zhan Q, Greenberger JS, Fornace AJ Jr. Human in vivo radiation-induced biomarkers: gene expression changes in radiotherapy patients. Cancer Res. 2004 Sep 15;64(18):6368-71. doi: 10.1158/0008-5472.CAN-04-1883. PMID: 15374940.

106. Bo H, Ghazizadeh M, Shimizu H, Kurihara Y, Egawa S, Moriyama Y, Tajiri T, Kawanami O. Effect of ionizing irradiation on human esophageal cancer cell lines by cDNA microarray gene expression analysis. J Nippon Med Sch. 2004 Jun;71(3):172-80. doi: 10.1272/jnms.71.172. PMID: 15226608.

107. Fu ZC, Wang FM, Cai JM. Gene expression changes in residual advanced cervical cancer after radiotherapy: indicators of poor prognosis and radioresistance? Med Sci Monit. 2015 May 5;21:1276-87. doi: 10.12659/MSM.893689. PMID: 25940978; PMCID: PMC4432617.

108. Klopp AH, Jhingran A, Ramdas L, Story MD, Broadus RR, Lu KH, Eifel PJ, Buchholz TA. Gene expression changes in cervical squamous cell carcinoma after initiation of chemoradiation and correlation with clinical outcome. Int J Radiat Oncol Biol Phys. 2008 May 1;71(1):226-36. doi: 10.1016/j.ijrobp.2007.10.068. PMID: 18406887.

109. Piening BD, Wang P, Subramanian A, Paulovich AG. A radiation-derived gene expression signature predicts clinical outcome for breast cancer patients. Radiat Res. 2009 Feb;171(2):141-54. doi: 10.1667/RR1223.1. PMID: 19267539; PMCID: PMC2662705.

110. Wang HP, Long XH, Sun ZZ, Rigaud O, Xu QZ, Huang YC, Sui JL, Bai B, Zhou PK. Identification of differentially transcribed genes in human lymphoblastoid cells irradiated with 0.5 Gy of gamma-ray and the involvement of low dose radiation inducible CHD6 gene in cell proliferation and radiosensitivity. Int J Radiat Biol. 2006 Mar;82(3):181-90. doi: 10.1080/09553000600632261. PMID: 16638715.

111. Foy JP, Bazire L, Ortiz-Cuaran S, Deneuve S, Kielbassa J, Thomas E, Viari A, Puisieux A, Goudot P, Bertolus C, Foray N, Kirova Y, Verrelle P, Saintigny P. A 13-gene expression-based radioresistance score highlights the heterogeneity in the response to radiation therapy across HPV-negative HNSCC molecular subtypes. BMC Med. 2017 Sep 1;15(1):165. doi: 10.1186/s12916-017-0929-y. PMID: 28859688; PMCID: PMC5580222.

112. Young A, Berry R, Holloway AF, Blackburn NB, Dickinson JL, Skala M, Phillips JL, Brettingham-Moore KH. RNA-seq profiling of a radiation resistant and radiation sensitive prostate cancer cell line highlights opposing regulation of DNA repair and targets for radiosensitization. BMC Cancer. 2014 Nov 4;14:808. doi: 10.1186/1471-2407-14-808. PMID: 25369795; PMCID: PMC4233036.

113. Ha Thi HT, Kim HY, Kim YM, Hong S. MicroRNA-130a modulates a radiosensitivity of rectal cancer by targeting SOX4. Neoplasia. 2019 Sep;21(9):882-892. doi: 10.1016/j.neo.2019.07.005. Epub 2019 Aug 3. PMID: 31387015; PMCID: PMC6690642.

114. He QE, Tong YF, Ye Z, Gao LX, Zhang YZ, Wang L, Song K. A multiple genomic data fused SF2 prediction model, signature identification, and gene regulatory network inference for personalized radiotherapy. Technol Cancer Res Treat. 2020 Jan-Dec;19:1533033820909112. doi: 10.1177/1533033820909112. PMID: 32329416; PMCID: PMC7225787.

115. Akerman GS, Rosenzweig BA, Domon OE, Tsai CA, Bishop ME, McGarrity LJ, Macgregor JT, Sistare FD, Chen JJ, Morris SM. Alterations in gene expression profiles and the DNA-damage response in ionizing radiation-exposed TK6 cells. Environ Mol Mutagen. 2005 Mar-Apr;45(2-3):188-205. doi: 10.1002/em.20091. PMID: 15657912.

116. Amundson SA, Lee RA, Koch-Paiz CA, Bittner ML, Meltzer P, Trent JM, Fornace AJ Jr. Differential responses of stress genes to low dose-rate gamma irradiation. Mol Cancer Res. 2003 Apr;1(6):445-52. PMID: 12692264.

117. Klopp AH, Jhingran A, Ramdas L, Story MD, Broadus RR, Lu KH, Eifel PJ, Buchholz TA. Gene expression changes in cervical squamous cell carcinoma after initiation of chemoradiation and correlation with clinical outcome. Int J Radiat Oncol Biol Phys. 2008 May 1;71(1):226-36. doi: 10.1016/j.ijrobp.2007.10.068. PMID: 18406887.

118. Summerer I, Hess J, Pitea A, Unger K, Hieber L, Selmansberger M, Lauber K, Zitzelsberger H. Integrative analysis of the microRNA-mRNA response to radiochemotherapy in primary head and neck squamous cell carcinoma cells. BMC Genomics. 2015 Sep 2;16(1):654. doi: 10.1186/s12864-015-1865-x. PMID: 26328888; PMCID: PMC4557600.

119. Eschrich S, Zhang H, Zhao H, Boulware D, Lee JH, Bloom G, Torres-Roca JF. Systems biology modeling of the radiation sensitivity network: a biomarker discovery platform. Int J Radiat Oncol Biol Phys. 2009 Oct 1;75(2):497-505. doi: 10.1016/j.ijrobp.2009.05.056. PMID: 19735874; PMCID: PMC2762403.

120. Huang MY, Wang JY, Chang HJ, Kuo CW, Tok TS, Lin SR. CDC25A, VAV1, TP73, BRCA1 and ZAP70 gene overexpression correlates with radiation response in colorectal cancer. Oncol Rep. 2011 May;25(5):1297-306. doi: 10.3892/or.2011.1193. Epub 2011 Feb 22. PMID: 21344162.

121. Kobunai T, Watanabe T, Fukusato T. REG4, NEIL2, and BIRC5 gene expression correlates with gamma-radiation sensitivity in patients with rectal cancer receiving radiotherapy. Anticancer Res. 2011 Dec;31(12):4147-53. PMID: 22199273.

122. Lee YS, Oh JH, Yoon S, Kwon MS, Song CW, Kim KH, Cho MJ, Mollah ML, Je YJ, Kim YD, Kim CD, Lee JH. Differential gene expression profiles of radioresistant non-small-cell lung cancer cell lines established by fractionated irradiation: tumor protein p53-inducible protein 3 confers sensitivity to ionizing radiation. Int J Radiat Oncol Biol Phys. 2010 Jul 1;77(3):858-66. doi: 10.1016/j.ijrobp.2009.12.076. PMID: 20510196.

123. Guo WF, Lin RX, Huang J, Zhou Z, Yang J, Guo GZ, Wang SQ. Identification of differentially expressed genes contributing to radioresistance in lung cancer cells using microarray analysis. Radiat Res. 2005 Jul;164(1):27-35. doi: 10.1667/rr3401. PMID: 15966762.

124. Achary MP, Jaggernauth W, Gross E, Alfieri A, Klinger HP, Vikram B. Cell lines from the same cervical carcinoma but with different radiosensitivities exhibit different cDNA microarray patterns of gene expression. Cytogenet Cell Genet. 2000;91(1-4):39-43. doi: 10.1159/000056815. PMID: 11173827.

125. Ojima E, Inoue Y, Miki C, Mori M, Kusunoki M. Effectiveness of gene expression profiling for response prediction of rectal cancer to preoperative radiotherapy. J Gastroenterol. 2007 Sep;42(9):730-6. doi: 10.1007/s00535-007-2089-x. Epub 2007 Sep 25. PMID: 17876542.

126. Tewari D, Monk BJ, Al-Ghazi MS, Parker R, Heck JD, Burger RA, Fruehauf JP. Gene expression profiling of in vitro radiation resistance in cervical carcinoma: a feasibility study. Gynecol Oncol. 2005 Oct;99(1):84-91. doi: 10.1016/j.ygyno.2005.05.043. PMID: 16109440.

127. Ahn SJ, Choi C, Choi YD, Kim YC, Kim KS, Oh IJ, Ban HJ, Yoon MS, Nam TK, Jeong JU, Song JY, Chung WK. Microarray analysis of gene expression in lung cancer cell lines treated by fractionated irradiation. Anticancer Res. 2014 Sep;34(9):4939-48. PMID: 25202076.

128. Chang JT, Chan SH, Lin CY, Lin TY, Wang HM, Liao CT, Wang TH, Lee LY, Cheng AJ. Differentially expressed genes in radioresistant nasopharyngeal cancer cells: gp96 and GDF15. Mol Cancer Ther. 2007 Aug;6(8):2271-9. doi: 10.1158/1535-7163.MCT-06-0801. Epub 2007 Aug 1. PMID: 17671084.

129. Fukuda K, Sakakura C, Miyagawa K, Kuriu Y, Kin S, Nakase Y, Hagiwara A, Mitsufuji S, Okazaki Y, Hayashizaki Y, Yamagishi H. Differential gene expression profiles of radioresistant oesophageal cancer cell lines established by continuous

fractionated irradiation. Br J Cancer. 2004 Oct 18;91(8):1543-50. doi: 10.1038/sj.bjc.6602187. PMID: 15365572; PMCID: PMC2409931.

130. Khodarev NN, Beckett M, Labay E, Darga T, Roizman B, Weichselbaum RR. STAT1 is overexpressed in tumors selected for radioresistance and confers protection from radiation in transduced sensitive cells. Proc Natl Acad Sci U S A. 2004 Feb 10;101(6):1714-9. doi: 10.1073/pnas.0308102100. Epub 2004 Jan 30. PMID: 14755057; PMCID: PMC341831.

131. Li XH, Qu JQ, Yi H, Zhang PF, Yi HM, Wan XX, He QY, Ye X, Yuan L, Zhu JF, Li JY, Xiao ZQ. Integrated analysis of differential miRNA and mRNA expression profiles in human radioresistant and radiosensitive nasopharyngeal carcinoma cells. PLoS One. 2014 Jan 31;9(1):e87767. doi: 10.1371/journal.pone.0087767. PMID: 24498188; PMCID: PMC3909230.

132. Michna A, Schötz U, Selmansberger M, Zitzelsberger H, Lauber K, Unger K, Hess J. Transcriptomic analyses of the radiation response in head and neck squamous cell carcinoma subclones with different radiation sensitivity: time-course gene expression profiles and gene association networks. Radiat Oncol. 2016 Jul 26;11:94. doi: 10.1186/s13014-016-0672-0. PMID: 27455841; PMCID: PMC4960706.

133. Ogawa R, Ishiguro H, Kuwabara Y, Kimura M, Mitsui A, Mori Y, Mori R, Tomoda K, Katada T, Harada K, Fujii Y. Identification of candidate genes involved in the radiosensitivity of esophageal cancer cells by microarray analysis. Dis Esophagus. 2008;21(4):288-97. doi: 10.1111/j.1442-2050.2007.00759.x. PMID: 18477249.

134. Zhang H, Gao XS, Zhao J, Xiong W, Zhang M, Li HZ, Zhou DM, Jin X, Zhang DS. Differential gene expression profiles of DNA repair genes in esophageal cancer cells after X-ray irradiation. Chin J Cancer. 2010 Oct;29(10):865-72. doi: 10.5732/cjc.010.10149. PMID: 20868556.

135. Zhou FX, Xiong J, Luo ZG, Dai J, Yu HJ, Liao ZK, Lei H, Xie CH, Zhou YF. cDNA expression analysis of a human radiosensitive-radioresistant cell line model

identifies telomere function as a hallmark of radioresistance. Radiat Res. 2010 Nov;174(5):550-7. doi: 10.1667/RR1657.1. Epub 2010 Jul 14. PMID: 20726715.

136. Seifert M, Peitzsch C, Gorodetska I, Börner C, Klink B, Dubrovska A. Network-based analysis of prostate cancer cell lines reveals novel marker gene candidates associated with radioresistance and patient relapse. PLoS Comput Biol. 2019 Nov 4;15(11):e1007460. doi: 10.1371/journal.pcbi.1007460. PMID: 31682594; PMCID: PMC6855562.

137. Todorovic V, Prevc A, Zakelj MN, Savarin M, Brozic A, Groselj B, Strojan P, Cemazar M, Sersa G. Mechanisms of different response to ionizing irradiation in isogenic head and neck cancer cell lines. Radiat Oncol. 2019 Nov 27;14(1):214. doi: 10.1186/s13014-019-1418-6. PMID: 31775835; PMCID: PMC6882348.

138. You GR, Cheng AJ, Lee LY, Huang YC, Liu H, Chen YJ, Chang JT. Prognostic signature associated with radioresistance in head and neck cancer via transcriptomic and bioinformatic analyses. BMC Cancer. 2019 Jan 14;19(1):64. doi: 10.1186/s12885-018-5243-3. PMID: 30642292; PMCID: PMC6332600.

139. Alfonso JCL, Berk L. Modeling the effect of intratumoral heterogeneity of radiosensitivity on tumor response over the course of fractionated radiation therapy. Radiat Oncol. 2019 May 30;14(1):88. doi: 10.1186/s13014-019-1288-y. PMID: 31146751; PMCID: PMC6543639.

140. Corchete LA, Rojas EA, Alonso-López D, De Las Rivas J, Gutiérrez NC, Burguillo FJ. Systematic comparison and assessment of RNA-seq procedures for gene expression quantitative analysis. Sci Rep. 2020 Nov 12;10(1):19737. doi: 10.1038/s41598-020-76881-x. PMID: 33184454; PMCID: PMC7665074.

141. Schweighofer B, Testori J, Sturtzel C, Sattler S, Mayer H, Wagner O, Bilban M, Hofer E. The VEGF-induced transcriptional response comprises gene clusters at the crossroad of angiogenesis and inflammation. Thromb Haemost. 2009

Sep;102(3):544-54. doi: 10.1160/TH08-12-0830. PMID: 19718476; PMCID: PMC2886966.

142. Gu H, Wu XY, Fan RT, Wang X, Guo YZ, Wang R. Side population cells from long-term passage non-small cell lung cancer cells display loss of cancer stem cell-like properties and chemoradioresistance. Oncol Lett. 2016 Oct;12(4):2886-2893. doi: 10.3892/ol.2016.4934. Epub 2016 Aug 2. PMID: 27698875; PMCID: PMC5038156.

143. Lonati L, Barbieri S, Guardamagna I, Ottolenghi A, Baiocco G. Radiation-induced cell cycle perturbations: a computational tool validated with flow-cytometry data. Sci Rep. 2021 Jan 13;11(1):925. doi: 10.1038/s41598-020-79934-3. PMID: 33441727; PMCID: PMC7806866.

144. Aruga T, Ando K, Iizuka M, Koike S, Fukutsu K, Itsukaichi H, Arimizu N. Radiosensitivity and cell cycle redistribution of cultured human tumour cells during fractionated daily 2-Gy irradiations. Int J Radiat Biol. 1995 Jan;67(1):65-70. doi: 10.1080/09553009514550081. PMID: 7852818

145. Peng Y, Fu S, Hu W, Qiu Y, Zhang L, Tan R, Sun LQ. Glutamine synthetase facilitates cancer cells to recover from irradiation-induced G2/M arrest. Cancer Biol Ther. 2020;21(1):43-51. doi: 10.1080/15384047.2019.1665394. Epub 2019 Sep 17. PMID: 31526079; PMCID: PMC7012188.

146. A, Swanton C; TRACERx Consortium. Tracking the Evolution of Non-Small-Cell Lung Cancer. N Engl J Med. 2017 Jun 1;376(22):2109-2121. doi: 10.1056/NEJMoa1616288. Epub 2017 Apr 26. PMID: 28445112.

147. Paczkowski M, Kretzschmar WW, Markelc B, Liu SK, Kunz-Schughart LA, Harris AL, Partridge M, Byrne HM, Kannan P. Reciprocal interactions between tumour cell populations enhance growth and reduce radiation sensitivity in prostate cancer. Commun Biol. 2021 Jan 4;4(1):6. doi: 10.1038/s42003-020-01529-5. PMID: 33398023; PMCID: PMC7782740.

148. Fukui R, Saga R, Matsuya Y, Tomita K, Kuwahara Y, Ohuchi K, Sato T, Okumura K, Date H, Fukumoto M, Hosokawa Y. Tumor radioresistance caused by radiation-induced changes of stem-like cell content and sub-lethal damage repair capability. Sci Rep. 2022 Jan 20;12(1):1056. doi: 10.1038/s41598-022-05172-4. PMID: 35058559; PMCID: PMC8776741.

149. Kim JH, Brown SL, Gordon MN. Radiation-induced senescence: therapeutic opportunities. Radiat Oncol. 2023 Jan 13;18(1):10. doi: 10.1186/s13014-022-02184-2. PMID: 36639774; PMCID: PMC9837958.

150. Tabasso AFS, Jones DJL, Jones GDD, Macip S. Radiotherapy-Induced Senescence and its Effects on Responses to Treatment. Clin Oncol (R Coll Radiol). 2019 May;31(5):283-289. doi: 10.1016/j.clon.2019.02.003. Epub 2019 Feb 27. PMID: 30826201.

151. Benadjaoud MA, Soysouvanh F, Tarlet G, Paget V, Buard V, Santos de Andrade H, Morilla I, Dos Santos M, Bertho A, l'Homme B, Gruel G, François A, Mondini M, Deutsch E, Guipaud O, Milliat F. Deciphering the Dynamic Molecular Program of Radiation-Induced Endothelial Senescence. Int J Radiat Oncol Biol Phys. 2022 Mar 15;112(4):975-985. doi: 10.1016/j.ijrobp.2021.11.019. Epub 2021 Nov 20. PMID: 34808254.

152. Patel NH, Sohal SS, Manjili MH, Harrell JC, Gewirtz DA. The Roles of Autophagy and Senescence in the Tumor Cell Response to Radiation. Radiat Res. 2020 Aug 1;194(2):103-115. doi: 10.1667/RADE-20-00009. PMID: 32845995; PMCID: PMC7482104.

153. Lee S, Schmitt CA. The dynamic nature of senescence in cancer. Nat Cell Biol. 2019 Jan;21(1):94-101. doi: 10.1038/s41556-018-0249-2. Epub 2019 Jan 2. PMID: 30602768.

154. Schmitt CA, Wang B, Demaria M. Senescence and cancer - role and therapeutic opportunities. Nat Rev Clin Oncol. 2022 Oct;19(10):619-636. doi:

10.1038/s41571-022-00668-4. Epub 2022 Aug 31. PMID: 36045302; PMCID: PMC9428886.

155. Milanovic M, Fan DNY, Belenki D, Däbritz JHM, Zhao Z, Yu Y, Dörr JR, Dimitrova L, Lenze D, Monteiro Barbosa IA, Mendoza-Parra MA, Kanashova T, Metzner M, Pardon K, Reimann M, Trumpp A, Dörken B, Zuber J, Gronemeyer H, Hummel M, Dittmar G, Lee S, Schmitt CA. Senescence-associated reprogramming promotes cancer stemness. Nature. 2018 Jan 4;553(7686):96-100. doi: 10.1038/nature25167. Epub 2017 Dec 20. PMID: 29258294.

156. von Joest M, Chen C, Douché T, Chantrel J, Chiche A, Gianetto QG, Matondo M, Li H. Amphiregulin mediates non-cell-autonomous effect of senescence on reprogramming. Cell Rep. 2022 Jul 12;40(2):111074. doi: 10.1016/j.celrep.2022.111074. PMID: 35830812.

157. Koch J, Mönch D, Maaß A, Gromoll C, Hehr T, Leibold T, Schlitt HJ, Dahlke MH, Renner P. Three dimensional cultivation increases chemo- and radioresistance of colorectal cancer cell lines. PLoS One. 2021 Jan 4;16(1):e0244513. doi: 10.1371/journal.pone.0244513. PMID: 33395433; PMCID: PMC7781370.

158. Shi D, Shi G, Huang G, Zhang J, Lartigau E. Chemosensitivity of radioresistant cells in the multicellular spheroids of A549 lung adenocarcinoma. J Exp Clin Cancer Res. 2009 Jun 2;28(1):72. doi: 10.1186/1756-9966-28-72. PMID: 19490637; PMCID: PMC2695815.

159. Roy A, Bera S, Saso L, Dwarakanath BS. Role of autophagy in tumor response to radiation: Implications for improving radiotherapy. Front Oncol. 2022 Sep 12;12:957373. doi: 10.3389/fonc.2022.957373. PMID: 36172166; PMCID: PMC9510974.

160. Chen H, Ma Z, Vanderwaal RP, Feng Z, Gonzalez-Suarez I, Wang S, Zhang J, Roti Roti JL, Gonzalo S, Zhang J. The mTOR inhibitor rapamycin suppresses DNA double-strand break repair. Radiat Res. 2011 Feb;175(2):214-24. doi: 10.1667/rr2323.1. Epub 2010 Oct 26. PMID: 21268715; PMCID: PMC4412148.

161. Casella G, Munk R, Kim KM, Piao Y, De S, Abdelmohsen K, Gorospe M. Transcriptome signature of cellular senescence. Nucleic Acids Res. 2019 Dec 2;47(21):11476. doi: 10.1093/nar/gkz879. Erratum for: Nucleic Acids Res. 2019 Aug 22;47(14):7294-7305. PMID: 31612919; PMCID: PMC6868356.

162. Pommer M, Kuphal S, Bosserhoff AK. Amphiregulin Regulates Melanocytic Senescence. Cells. 2021 Feb 5;10(2):326. doi: 10.3390/cells10020326. PMID: 33562468; PMCID: PMC7914549.

163. Yu Y, Schleich K, Yue B, Ji S, Lohneis P, Kemper K, Silvis MR, Qutob N, van Rooijen E, Werner-Klein M, Li L, Dhawan D, Meierjohann S, Reimann M, Elkahloun A, Treitschke S, Dörken B, Speck C, Mallette FA, Zon LI, Holmen SL, Peeper DS, Samuels Y, Schmitt CA, Lee S. Targeting the Senescence-Overriding Cooperative Activity of Structurally Unrelated H3K9 Demethylases in Melanoma. Cancer Cell. 2018 Apr 9;33(4):785. doi: 10.1016/j.ccell.2018.03.009. Erratum for: Cancer Cell. 2018 Feb 12;33(2):322-336.e8. doi: 10.1016/j.ccell.2018.01.002. PMID: 29634951.

164. Zheng YW, Li ZH, Lei L, Liu CC, Wang Z, Fei LR, Yang MQ, Huang WJ, Xu HT. FAM83A Promotes Lung Cancer Progression by Regulating the Wnt and Hippo Signaling Pathways and Indicates Poor Prognosis. Front Oncol. 2020 Mar 5;10:180. doi: 10.3389/fonc.2020.00180. PMID: 32195172; PMCID: PMC7066079.

165. Zhou C, Zhu X, Liu N, Dong X, Zhang X, Huang H, Tang Y, Liu S, Hu M, Wang M, Deng X, Li S, Zhang R, Huang Y, Lyu H, Xiao S, Luo S, Ali DW, Michalak M, Chen XZ, Wang Z, Tang J. B-lymphoid tyrosine kinase-mediated FAM83A phosphorylation elevates pancreatic tumorigenesis through interacting with β-catenin. Signal Transduct Target Ther. 2023 Feb 17;8(1):66. doi: 10.1038/s41392-022-01268-5. PMID: 36797256; PMCID: PMC9935901.

166. Rath BH, Waung I, Camphausen K, Tofilon PJ. Inhibition of the Histone H3K27 Demethylase UTX Enhances Tumor Cell Radiosensitivity. Mol Cancer Ther. 2018 May;17(5):1070-1078. doi: 10.1158/1535-7163.MCT-17-1053. Epub 2018 Feb 26.

PMID: 29483212; PMCID: PMC5932086.

167. Fridman AL, Tainsky MA. Critical pathways in cellular senescence and immortalization revealed by gene expression profiling. Oncogene. 2008 Oct 9;27(46):5975-87. doi: 10.1038/onc.2008.213. Epub 2008 Aug 18. PMID: 18711403; PMCID: PMC3843241.

168. Chen H, Ma Z, Vanderwaal RP, Feng Z, Gonzalez-Suarez I, Wang S, Zhang J, Roti Roti JL, Gonzalo S, Zhang J. The mTOR inhibitor rapamycin suppresses DNA double-strand break repair. Radiat Res. 2011 Feb;175(2):214-24. doi: 10.1667/rr2323.1. Epub 2010 Oct 26. PMID: 21268715; PMCID: PMC4412148.

169. Kordaß T, Osen W, Eichmüller SB. Controlling the Immune Suppressor: Transcription Factors and MicroRNAs Regulating CD73/NT5E. Front Immunol. 2018 Apr 18;9:813. doi: 10.3389/fimmu.2018.00813. PMID: 29720980; PMCID: PMC5915482.

170. Hirano T. IL-6 in inflammation, autoimmunity and cancer. Int Immunol. 2021 Mar 1;33(3):127-148. doi: 10.1093/intimm/dxaa078. PMID: 33337480; PMCID: PMC7799025.

171. Ferrara N, Hillan KJ, Novotny W. Bevacizumab (Avastin), a humanized anti-VEGF monoclonal antibody for cancer therapy. Biochem Biophys Res Commun. 2005 Jul 29;333(2):328-35. doi: 10.1016/j.bbrc.2005.05.132. PMID: 15961063.

172. Rosell R, Karachaliou N. Gene Expression Signatures Predicting Survival and Chemotherapy Benefit in Patients with Resected Non-small-Cell Lung Cancer. EBioMedicine. 2018 Jul;33:16-17. doi: 10.1016/j.ebiom.2018.06.007. PMID: 29929887; PMCID: PMC6085495.

173. Dehing-Oberije C, Aerts H, Yu S, De Ruysscher D, Menheere P, Hilvo M, van der Weide H, Rao B, Lambin P. Development and validation of a prognostic model using blood biomarker information for prediction of survival of non-small-cell lung cancer patients treated with combined chemotherapy and radiation or radiotherapy alone (NCT00181519, NCT00573040, and NCT00572325). Int J Radiat Oncol Biol

Phys. 2011 Oct 1;81(2):360-8. doi: 10.1016/j.ijrobp.2010.06.011. Epub 2010 Oct 1. PMID: 20888135.

174. Peinado-Serrano J, Quintanal-Villalonga Á, Muñoz-Galvan S, Verdugo-Sivianes EM, Mateos JC, Ortiz-Gordillo MJ, Carnero A. A Six-Gene Prognostic and Predictive Radiotherapy-Based Signature for Early and Locally Advanced Stages in Non-Small-Cell Lung Cancer. Cancers (Basel). 2022 Apr 19;14(9):2054. doi: 10.3390/cancers14092054. PMID: 35565183; PMCID: PMC9099638.

175. Hamdy FC, Donovan JL, Lane JA, Mason M, Metcalfe C, Holding P, Davis M, Peters TJ, Turner EL, Martin RM, Oxley J, Robinson M, Staffurth J, Walsh E, Bollina P, Catto J, Doble A, Doherty A, Gillatt D, Kockelbergh R, Kynaston H, Paul A, Powell P, Prescott S, Rosario DJ, Rowe E, Neal DE; ProtecT Study Group. 10-Year Outcomes after Monitoring, Surgery, or Radiotherapy for Localized Prostate Cancer. N Engl J Med. 2016 Oct 13;375(15):1415-1424. doi: 10.1056/NEJMoa1606220. Epub 2016 Sep 14. PMID: 27626136.

176. Donovan JL, Hamdy FC, Lane JA, Mason M, Metcalfe C, Walsh E, Blazeby JM, Peters TJ, Holding P, Bonnington S, Lennon T, Bradshaw L, Cooper D, Herbert P, Howson J, Jones A, Lyons N, Salter E, Thompson P, Tidball S, Blaikie J, Gray C, Bollina P, Catto J, Doble A, Doherty A, Gillatt D, Kockelbergh R, Kynaston H, Paul A, Powell P, Prescott S, Rosario DJ, Rowe E, Davis M, Turner EL, Martin RM, Neal DE; ProtecT Study Group*. Patient-Reported Outcomes after Monitoring, Surgery, or Radiotherapy for Prostate Cancer. N Engl J Med. 2016 Oct 13;375(15):1425-1437. doi: 10.1056/NEJMoa1606221. Epub 2016 Sep 14. PMID: 27626365; PMCID: PMC5134995.

177. D'Amico AV, Whittington R, Malkowicz SB, Schultz D, Blank K, Broderick GA, Tomaszewski JE, Renshaw AA, Kaplan I, Beard CJ, Wein A. Biochemical outcome after radical prostatectomy, external beam radiation therapy, or interstitial radiation therapy for clinically localized prostate cancer. JAMA. 1998 Sep 16;280(11):969-74. doi: 10.1001/jama.280.11.969. PMID: 9749478.

178. Cooperberg MR, Pasta DJ, Elkin EP, Litwin MS, Latini DM, Du Chane J, Carroll PR. The University of California, San Francisco Cancer of the Prostate Risk Assessment score: a straightforward and reliable preoperative predictor of disease recurrence after radical prostatectomy. J Urol. 2005 Jun;173(6):1938-42. doi: 10.1097/01.ju.0000158155.33890.e7. Erratum in: J Urol. 2006 Jun;175(6):2369. PMID: 15879786; PMCID: PMC2948569.

179. Klein EA, Cooperberg MR, Magi-Galluzzi C, Simko JP, Falzarano SM, Maddala T, Chan JM, Li J, Cowan JE, Tsiatis AC, Cherbavaz DB, Pelham RJ, Tenggara-Hunter I, Baehner FL, Knezevic D, Febbo PG, Shak S, Kattan MW, Lee M, Carroll PR. A 17-gene assay to predict prostate cancer aggressiveness in the context of Gleason grade heterogeneity, tumor multifocality, and biopsy undersampling. Eur Urol. 2014 Sep;66(3):550-60. doi: 10.1016/j.eururo.2014.05.004. Epub 2014 May 16. PMID: 24836057.

180. Nakagawa T, Kollmeyer TM, Morlan BW, Anderson SK, Bergstralh EJ, Davis BJ, Asmann YW, Klee GG, Ballman KV, Jenkins RB. A tissue biomarker panel predicting systemic progression after PSA recurrence post-definitive prostate cancer therapy. PLoS One. 2008;3(5):e2318. doi: 10.1371/journal.pone.0002318. Epub 2008 May 28. PMID: 18846227; PMCID: PMC2565588.

181. Cuzick J, Swanson GP, Fisher G, Brothman AR, Berney DM, Reid JE, Mesher D, Speights VO, Stankiewicz E, Foster CS, Møller H, Scardino P, Warren JD, Park J, Younus A, Flake DD 2nd, Wagner S, Gutin A, Lanchbury JS, Stone S; Transatlantic Prostate Group. Prognostic value of an RNA expression signature derived from cell cycle proliferation genes in patients with prostate cancer: a retrospective study. Lancet Oncol. 2011 Mar;12(3):245-55. doi: 10.1016/S1470-2045(10)70295-3. PMID: 21310658; PMCID: PMC3091030.

182. Jairath NK, Dal Pra A, Vince R Jr, Dess RT, Jackson WC, Tosoian JJ, McBride SM, Zhao SG, Berlin A, Mahal BA, Kishan AU, Den RB, Freedland SJ, Salami SS, Kaffenberger SD, Pollack A, Tran P, Mehra R, Morgan TM, Weiner AB,

Mohamad O, Carroll PR, Cooperberg MR, Karnes RJ, Nguyen PL, Michalski JM, Tward JD, Feng FY, Schaeffer EM, Spratt DE. A Systematic Review of the Evidence for the Decipher Genomic Classifier in Prostate Cancer. Eur Urol. 2021 Mar;79(3):374-383. doi: 10.1016/j.eururo.2020.11.021. Epub 2020 Dec 5. PMID: 33293078.

183. Walker SM, Knight LA, McCavigan AM, Logan GE, Berge V, Sherif A, Pandha H, Warren AY, Davidson C, Uprichard A, Blayney JK, Price B, Jellema GL, Steele CJ, Svindland A, McDade SS, Eden CG, Foster C, Mills IG, Neal DE, Mason MD, Kay EW, Waugh DJ, Harkin DP, Watson RW, Clarke NW, Kennedy RD. Molecular Subgroup of Primary Prostate Cancer Presenting with Metastatic Biology. Eur Urol. 2017 Oct;72(4):509-518. doi: 10.1016/j.eururo.2017.03.027. Epub 2017 Apr 10. PMID: 28408174.

184. Yang L, Roberts D, Takhar M, Erho N, Bibby BAS, Thiruthaneeswaran N, Bhandari V, Cheng WC, Haider S, McCorry AMB, McArt D, Jain S, Alshalalfa M, Ross A, Schaffer E, Den RB, Jeffrey Karnes R, Klein E, Hoskin PJ, Freedland SJ, Lamb AD, Neal DE, Buffa FM, Bristow RG, Boutros PC, Davicioni E, Choudhury A, West CML. Development and Validation of a 28-gene Hypoxia-related Prognostic Signature for Localized Prostate Cancer. EBioMedicine. 2018 May;31:182-189. doi: 10.1016/j.ebiom.2018.04.019. Epub 2018 Apr 23. PMID: 29729848; PMCID: PMC6014579.

185. Jairath NK, Dal Pra A, Vince R Jr, Dess RT, Jackson WC, Tosoian JJ, McBride SM, Zhao SG, Berlin A, Mahal BA, Kishan AU, Den RB, Freedland SJ, Salami SS, Kaffenberger SD, Pollack A, Tran P, Mehra R, Morgan TM, Weiner AB, Mohamad O, Carroll PR, Cooperberg MR, Karnes RJ, Nguyen PL, Michalski JM, Tward JD, Feng FY, Schaeffer EM, Spratt DE. A Systematic Review of the Evidence for the Decipher Genomic Classifier in Prostate Cancer. Eur Urol. 2021 Mar;79(3):374-383. doi: 10.1016/j.eururo.2020.11.021. Epub 2020 Dec 5. PMID: 33293078.

186. Wang M, Nagle RB, Knudsen BS, Cress AE, Rogers GC. Centrosome loss results in an unstable genome and malignant prostate tumors. Oncogene. 2020 Jan;39(2):399-413. doi: 10.1038/s41388-019-0995-z. Epub 2019 Sep 2. PMID: 31477840.

187. Janes JL, Boyer MJ, Bennett JP, Thomas VM, De Hoedt AM, Edwards V DK, Singla PK, Abran JM, Aboushwareb T, Salama JK, Freedland SJ. The 17-Gene Genomic Prostate Score Test Is Prognostic for Outcomes After Primary External Beam Radiation Therapy in Men With Clinically Localized Prostate Cancer. Int J Radiat Oncol Biol Phys. 2023 Jan 1;115(1):120-131. doi: 10.1016/j.ijrobp.2022.06.101. Epub 2022 Oct 25. PMID: 36306979.

188. Tward J, Lenz L, Flake DD II,, Rajamani S, Yonover P, Olsson C, Kapoor DA, Mantz C, Liauw SL, Antic T, Fabrizio M, Salzstein D, Shore N, Albertson D, Henderson J, Lee SP, Gay HA, Michalski J, Hung A, Raben D, Garraway I, Lewis MS, Nguyen PL, Marshall DT, Brawer MK, Stone S, Cohen T. The Clinical Cell-Cycle Risk (CCR) Score Is Associated With Metastasis After Radiation Therapy and Provides Guidance on When to Forgo Combined Androgen Deprivation Therapy With Dose-Escalated Radiation. Int J Radiat Oncol Biol Phys. 2022 May 1;113(1):66-76. doi: 10.1016/j.ijrobp.2021.09.034. Epub 2021 PMID: 34610388.

189. Freedland SJ, Gerber L, Reid J, Welbourn W, Tikishvili E, Park J, Younus A, Gutin A, Sangale Z, Lanchbury JS, Salama JK, Stone S. Prognostic utility of cell cycle progression score in men with prostate cancer after primary external beam radiation therapy. Int J Radiat Oncol Biol Phys. 2013 Aug 1;86(5):848-53. doi: 10.1016/j.ijrobp.2013.04.043 PMID: 23755923; PMCID: PMC3710548.

190. Nguyen PL, Martin NE, Choeurng V, Palmer-Aronsten B, Kolisnik T, Beard CJ, Orio PF, Nezolosky MD, Chen YW, Shin H, Davicioni E, Feng FY. Utilization of biopsy-based genomic classifier to predict distant metastasis after definitive radiation and short-course ADT for intermediate and high-risk prostate cancer. Prostate

Cancer Prostatic Dis. 2017 Jun;20(2):186-192. doi: 10.1038/pcan.2016.58 PMID: 28117383; PMCID: PMC5435968.

191. Walker SM, Knight LA, McCavigan AM, Logan GE, Berge V, Sherif A, Pandha H, Warren AY, Davidson C, Uprichard A, Blayney JK, Price B, Jellema GL, Steele CJ, Svindland A, McDade SS, Eden CG, Foster C, Mills IG, Neal DE, Mason MD, Kay EW, Waugh DJ, Harkin DP, Watson RW, Clarke NW, Kennedy RD. Molecular Subgroup of Primary Prostate Cancer Presenting with Metastatic Biology. Eur Urol. 2017 Oct;72(4):509-518. doi: 10.1016/j.eururo.2017.03.027. Epub 2017 Apr 10. PMID: 28408174.

192. Lam D, Clark S, Stirzaker C, Pidsley R. Advances in Prognostic Methylation Biomarkers for Prostate Cancer. Cancers (Basel). 2020 Oct 15;12(10):2993. doi: 10.3390/cancers12102993. PMID: 33076494; PMCID: PMC7602626.

193. Jeyapala R, Kamdar S, Olkhov-Mitsel E, Savio AJ, Zhao F, Cuizon C, Liu RS, Zlotta A, Fleshner N, van der Kwast T, Bapat B. An integrative DNA methylation model for improved prognostication of postsurgery recurrence and therapy in prostate cancer patients. Urol Oncol. 2020 Feb;38(2):39.e1-39.e9. doi: 10.1016/j.urolonc.2019.08.017. Epub 2019 Sep 24. PMID: 31558364.

194. Savio AJ, Kamdar S, Jeyapala R, Olkhov-Mitsel E, Cuizon C, Finelli A, Zlotta AR, Toi A, Fleshner NE, van der Kwast T, Bapat B. Methylation Markers in Prostate Biopsies Are Prognosticators for Late Biochemical Recurrence and Therapy after Surgery in Prostate Cancer Patients. J Mol Diagn. 2020 Jan;22(1):30-39. doi: 10.1016/j.jmoldx.2019.08.003. Epub 2019 Oct 9. PMID: 31605802.

195. Massie CE, Mills IG, Lynch AG. The importance of DNA methylation in prostate cancer development. J Steroid Biochem Mol Biol. 2017 Feb;166:1-15. doi: 10.1016/j.jsbmb.2016.04.009. Epub 2016 Apr 24. PMID: 27117390.

196. Jain S, Lyons CA, Walker SM, McQuaid S, Hynes SO, Mitchell DM, Pang B, Logan GE, McCavigan AM, O'Rourke D, McArt DG, McDade SS, Mills IG, Prise KM, Knight LA, Steele CJ, Medlow PW, Berge V, Katz B, Loblaw DA, Harkin DP, James

JA, O'Sullivan JM, Kennedy RD, Waugh DJ. Validation of a Metastatic Assay using biopsies to improve risk stratification in patients with prostate cancer treated with radical radiation therapy. Ann Oncol. 2018 Jan 1;29(1):215-222. doi: 10.1093/annonc/mdx637. PMID: 29045551; PMCID: PMC5834121.

197. Nakagawa T, Kollmeyer TM, Morlan BW, Anderson SK, Bergstralh EJ, Davis BJ, Asmann YW, Klee GG, Ballman KV, Jenkins RB. A tissue biomarker panel predicting systemic progression after PSA recurrence post-definitive prostate cancer therapy. PLoS One. 2008;3(5):e2318. doi: 10.1371/journal.pone.0002318. Epub 2008 May 28. PMID: 18846227; PMCID: PMC2565588.

198. Cuzick J, Swanson GP, Fisher G, Brothman AR, Berney DM, Reid JE, Mesher D, Speights VO, Stankiewicz E, Foster CS, Møller H, Scardino P, Warren JD, Park J, Younus A, Flake DD 2nd, Wagner S, Gutin A, Lanchbury JS, Stone S; Transatlantic Prostate Group. Prognostic value of an RNA expression signature derived from cell cycle proliferation genes in patients with prostate cancer: a retrospective study. Lancet Oncol. 2011 Mar;12(3):245-55. doi: 10.1016/S1470-2045(10)70295-3. PMID: 21310658; PMCID: PMC3091030.

199. Jairath NK, Dal Pra A, Vince R Jr, Dess RT, Jackson WC, Tosoian JJ, McBride SM, Zhao SG, Berlin A, Mahal BA, Kishan AU, Den RB, Freedland SJ, Salami SS, Kaffenberger SD, Pollack A, Tran P, Mehra R, Morgan TM, Weiner AB, Mohamad O, Carroll PR, Cooperberg MR, Karnes RJ, Nguyen PL, Michalski JM, Tward JD, Feng FY, Schaeffer EM, Spratt DE. A Systematic Review of the Evidence for the Decipher Genomic Classifier in Prostate Cancer. Eur Urol. 2021 Mar;79(3):374-383. doi: 10.1016/j.eururo.2020.11.021. Epub 2020 Dec 5. PMID: 33293078.

200. Walker SM, Knight LA, McCavigan AM, Logan GE, Berge V, Sherif A, Pandha H, Warren AY, Davidson C, Uprichard A, Blayney JK, Price B, Jellema GL, Steele CJ, Svindland A, McDade SS, Eden CG, Foster C, Mills IG, Neal DE, Mason MD, Kay EW, Waugh DJ, Harkin DP, Watson RW, Clarke NW, Kennedy RD. Molecular

Subgroup of Primary Prostate Cancer Presenting with Metastatic Biology. Eur Urol. 2017 Oct;72(4):509-518. doi: 10.1016/j.eururo.2017.03.027. Epub 2017 Apr 10. PMID: 28408174.

201. Yang L, Roberts D, Takhar M, Erho N, Bibby BAS, Thiruthaneeswaran N, Bhandari V, Cheng WC, Haider S, McCorry AMB, McArt D, Jain S, Alshalalfa M, Ross A, Schaffer E, Den RB, Jeffrey Karnes R, Klein E, Hoskin PJ, Freedland SJ, Lamb AD, Neal DE, Buffa FM, Bristow RG, Boutros PC, Davicioni E, Choudhury A, West CML. Development and Validation of a 28-gene Hypoxia-related Prognostic Signature for Localized Prostate Cancer. EBioMedicine. 2018 May;31:182-189. doi: 10.1016/j.ebiom.2018.04.019. Epub 2018 Apr 23. PMID: 29729848; PMCID: PMC6014579.

202. Varambally S, Dhanasekaran SM, Zhou M, Barrette TR, Kumar-Sinha C, Sanda MG, Ghosh D, Pienta KJ, Sewalt RG, Otte AP, Rubin MA, Chinnaiyan AM. The polycomb group protein EZH2 is involved in progression of prostate cancer. Nature. 2002 Oct 10;419(6907):624-9. doi: 10.1038/nature01075. PMID: 12374981.

203. Robinson D, Van Allen EM, Wu YM, Schultz N, Lonigro RJ, Mosquera JM, Montgomery B, Taplin ME, Pritchard CC, Attard G, Beltran H, Abida W, Bradley RK, Vinson J, Cao X, Vats P, Kunju LP, Hussain M, Feng FY, Tomlins SA, Cooney KA, Smith DC, Brennan C, Siddiqui J, Mehra R, Chen Y, Rathkopf DE, Morris MJ, Solomon SB, Durack JC, Reuter VE, Gopalan A, Gao J, Loda M, Lis RT, Bowden M, Balk SP, Gaviola G, Sougnez C, Gupta M, Yu EY, Mostaghel EA, Cheng HH, Mulcahy H, True LD, Plymate SR, Dvinge H, Ferraldeschi R, Flohr P, Miranda S, Zafeiriou Z, Tunariu N, Mateo J, Perez-Lopez R, Demichelis F, Robinson BD, Schiffman M, Nanus DM, Tagawa ST, Sigaras A, Eng KW, Elemento O, Sboner A, Heath EI, Scher HI, Pienta KJ, Kantoff P, de Bono JS, Rubin MA, Nelson PS, Garraway LA, Sawyers CL, Chinnaiyan AM. Integrative clinical genomics of advanced prostate cancer. Cell. 2015 May 21;161(5):1215-1228. doi:

10.1016/j.cell.2015.05.001. Erratum in: Cell. 2015 Jul 16;162(2):454. PMID: 26000489; PMCID: PMC4484602.

204. Oh S, Shin S, Song H, Grande JP, Janknecht R. Relationship between ETS Transcription Factor ETV1 and TGF-β-regulated SMAD Proteins in Prostate Cancer. Sci Rep. 2019 Jun 3;9(1):8186. doi: 10.1038/s41598-019-44685-3. PMID: 31160676; PMCID: PMC6546734.

205. Kim TD, Jin F, Shin S, Oh S, Lightfoot SA, Grande JP, Johnson AJ, van Deursen JM, Wren JD, Janknecht R. Histone demethylase JMJD2A drives prostate tumorigenesis through transcription factor ETV1. J Clin Invest. 2016 Feb;126(2):706-20. doi: 10.1172/JCI78132. Epub 2016 Jan 5. PMID: 26731476; PMCID: PMC4731184.

206. Thibodeau SN, French AJ, McDonnell SK, Cheville J, Middha S, Tillmans L, Riska S, Baheti S, Larson MC, Fogarty Z, Zhang Y, Larson N, Nair A, O'Brien D, Wang L, Schaid DJ. Identification of candidate genes for prostate cancer-risk SNPs utilizing a normal prostate tissue eQTL data set. Nat Commun. 2015 Nov 27;6:8653. doi: 10.1038/ncomms9653. PMID: 26611117; PMCID: PMC4663677.

207. Murdocca M, De Masi C, Pucci S, Mango R, Novelli G, Di Natale C, Sangiuolo F. LOX-1 and cancer: an indissoluble liaison. Cancer Gene Ther. 2021 Nov;28(10-11):1088-1098. doi: 10.1038/s41417-020-00279-0. Epub 2021 Jan 5. PMID: 33402733; PMCID: PMC8571092.

208. González-Chavarría I, Cerro RP, Parra NP, Sandoval FA, Zuñiga FA, Omazábal VA, Lamperti LI, Jiménez SP, Fernandez EA, Gutiérrez NA, Rodriguez FS, Onate SA, Sánchez O, Vera JC, Toledo JR. Lectin-like oxidized LDL receptor-1 is an enhancer of tumor angiogenesis in human prostate cancer cells. PLoS One. 2014 Aug 29;9(8):e106219. doi: 10.1371/journal.pone.0106219. PMID: 25170920; PMCID: PMC4149537.

209. Hayashi T, Oue N, Sakamoto N, Anami K, Oo HZ, Sentani K, Ohara S, Teishima J, Matsubara A, Yasui W. Identification of transmembrane protein in

prostate cancer by the Escherichia coli ampicillin secretion trap: expression of CDON is involved in tumor cell growth and invasion. Pathobiology. 2011;78(5):277-84. doi: 10.1159/000329588. Epub 2011 Aug 17. PMID: 21849809.

210. Nakata D, Nakao S, Nakayama K, Araki S, Nakayama Y, Aparicio S, Hara T, Nakanishi A. The RNA helicase DDX39B and its paralog DDX39A regulate androgen receptor splice variant AR-V7 generation. Biochem Biophys Res Commun. 2017 Jan 29;483(1):271-276. doi: 10.1016/j.bbrc.2016.12.153. Epub 2016 Dec 23. PMID: 28025139.

211. Zhao SG, Chen WS, Li H, Foye A, Zhang M, Sjöström M, Aggarwal R, Playdle D, Liao A, Alumkal JJ, Das R, Chou J, Hua JT, Barnard TJ, Bailey AM, Chow ED, Perry MD, Dang HX, Yang R, Moussavi-Baygi R, Zhang L, Alshalalfa M, Laura Chang S, Houlahan KE, Shiah YJ, Beer TM, Thomas G, Chi KN, Gleave M, Zoubeidi A, Reiter RE, Rettig MB, Witte O, Yvonne Kim M, Fong L, Spratt DE, Morgan TM, Bose R, Huang FW, Li H, Chesner L, Shenoy T, Goodarzi H, Asangani IA, Sandhu S, Lang JM, Mahajan NP, Lara PN, Evans CP, Febbo P, Batzoglou S, Knudsen KE, He HH, Huang J, Zwart W, Costello JF, Luo J, Tomlins SA, Wyatt AW, Dehm SM, Ashworth A, Gilbert LA, Boutros PC, Farh K, Chinnaiyan AM, Maher CA, Small EJ, Quigley DA, Feng FY. The DNA methylation landscape of advanced prostate cancer. Nat Genet. 2020 Aug;52(8):778-789. doi: 10.1038/s41588-020-0648-8. Epub 2020 Jul 13. PMID: 32661416; PMCID: PMC7454228.

212. Mundbjerg K, Chopra S, Alemozaffar M, Duymich C, Lakshminarasimhan R, Nichols PW, Aron M, Siegmund KD, Ukimura O, Aron M, Stern M, Gill P, Carpten JD, Ørntoft TF, Sørensen KD, Weisenberger DJ, Jones PA, Duddalwar V, Gill I, Liang G. Identifying aggressive prostate cancer foci using a DNA methylation classifier. Genome Biol. 2017 Jan 12;18(1):3. doi: 10.1186/s13059-016-1129-3. PMID: 28081708; PMCID: PMC5234101.

213. Zhao SG, Chen WS, Li H, Foye A, Zhang M, Sjöström M, Aggarwal R, Playdle D, Liao A, Alumkal JJ, Das R, Chou J, Hua JT, Barnard TJ, Bailey AM, Chow

ED, Perry MD, Dang HX, Yang R, Moussavi-Baygi R, Zhang L, Alshalalfa M, Laura Chang S, Houlahan KE, Shiah YJ, Beer TM, Thomas G, Chi KN, Gleave M, Zoubeidi A, Reiter RE, Rettig MB, Witte O, Yvonne Kim M, Fong L, Spratt DE, Morgan TM, Bose R, Huang FW, Li H, Chesner L, Shenoy T, Goodarzi H, Asangani IA, Sandhu S, Lang JM, Mahajan NP, Lara PN, Evans CP, Febbo P, Batzoglou S, Knudsen KE, He HH, Huang J, Zwart W, Costello JF, Luo J, Tomlins SA, Wyatt AW, Dehm SM, Ashworth A, Gilbert LA, Boutros PC, Farh K, Chinnaiyan AM, Maher CA, Small EJ, Quigley DA, Feng FY. The DNA methylation landscape of advanced prostate cancer. Nat Genet. 2020 Aug;52(8):778-789. doi: 10.1038/s41588-020-0648-8. Epub 2020 Jul 13. PMID: 32661416; PMCID: PMC7454228.

214. Ylitalo EB, Thysell E, Landfors M, Brattsand M, Jernberg E, Crnalic S, Widmark A, Hultdin M, Bergh A, Degerman S, Wikström P. A novel DNA methylation signature is associated with androgen receptor activity and patient prognosis in bone metastatic prostate cancer. Clin Epigenetics. 2021 Jun 30;13(1):133. doi: 10.1186/s13148-021-01119-0. PMID: 34193246; PMCID: PMC8244194.

215. Rauluseviciute I, Drabløs F, Rye MB. DNA hypermethylation associated with upregulated gene expression in prostate cancer demonstrates the diversity of epigenetic regulation. BMC Med Genomics. 2020 Jan 8;13(1):6. doi: 10.1186/s12920-020-0657-6. PMID: 31914996; PMCID: PMC6950795.

216. Ramanand SG, Chen Y, Yuan J, Daescu K, Lambros MB, Houlahan KE, Carreira S, Yuan W, Baek G, Sharp A, Paschalis A, Kanchwala M, Gao Y, Aslam A, Safdar N, Zhan X, Raj GV, Xing C, Boutros PC, de Bono J, Zhang MQ, Mani RS. The landscape of RNA polymerase II-associated chromatin interactions in prostate cancer. J Clin Invest. 2020 Aug 3;130(8):3987-4005. doi: 10.1172/JCI134260. PMID: 32343676; PMCID: PMC7410051.

217. Park SH, Fong KW, Mong E, Martin MC, Schiltz GE, Yu J. Going beyond Polycomb: EZH2 functions in prostate cancer. Oncogene. 2021 Sep;40(39):5788-

5798. doi: 10.1038/s41388-021-01982-4. Epub 2021 Aug 4. PMID: 34349243; PMCID: PMC8487936.

218. Xu K, Wu ZJ, Groner AC, He HH, Cai C, Lis RT, Wu X, Stack EC, Loda M, Liu T, Xu H, Cato L, Thornton JE, Gregory RI, Morrissey C, Vessella RL, Montironi R, Magi-Galluzzi C, Kantoff PW, Balk SP, Liu XS, Brown M. EZH2 oncogenic activity in castration-resistant prostate cancer cells is Polycomb-independent. Science. 2012 Dec 14;338(6113):1465-9. doi: 10.1126/science.1227604. PMID: 23239736; PMCID: PMC3625962.

219. Wu X, Scott H, Carlsson SV, Sjoberg DD, Cerundolo L, Lilja H, Prevo R, Rieunier G, Macaulay V, Higgins GS, Verrill CL, Lamb AD, Cunliffe VT, Bountra C, Hamdy FC, Bryant RJ. Increased EZH2 expression in prostate cancer is associated with metastatic recurrence following external beam radiotherapy. Prostate. 2019 Jul;79(10):1079-1089. doi: 10.1002/pros.23817. Epub 2019 May 18. PMID: 31104332; PMCID: PMC6563086.

220. Gabriel N, Balaji K, Jayachandran K, Inkman M, Zhang J, Dahiya S, Goldstein M. Loss of H3K27 Trimethylation Promotes Radiotherapy Resistance in Medulloblastoma and Induces an Actionable Vulnerability to BET Inhibition. Cancer Res. 2022 May 16;82(10):2019-2030. doi: 10.1158/0008-5472.CAN-21-0871. PMID: 35315927; PMCID: PMC9117495.

221. Spratt DE, Tang S, Sun Y, Huang HC, Chen E, Mohamad O, Armstrong AJ, Tward JD, Nguyen PL, Lang JM, Zhang J, Mitani A, Simko JP, DeVries S, van der Wal D, Pinckaers H, Monson JM, Campbell HA, Wallace J, Ferguson MJ, Bahary JP, Schaeffer EM; NRG Prostate Cancer AI Consortium; Sandler HM, Tran PT, Rodgers JP, Esteva A, Yamashita R, Feng FY. Artificial Intelligence Predictive Model for Hormone Therapy Use in Prostate Cancer. Res Sq [Preprint]. 2023 Apr 21:rs.3.rs-2790858. doi: 10.21203/rs.3.rs-2790858/v1. Update in: NEJM Evid. 2023 Aug;2(8):EVIDoa2300023. doi: 10.1056/EVIDoa2300023. PMID: 37131691; PMCID: PMC10153374.

222. Pizurica M, Zheng Y, Carrillo-Perez F, Noor H, Yao W, Wohlfart C, Vladimirova A, Marchal K, Gevaert O. Digital profiling of gene expression from histology images with linearized attention. Nat Commun. 2024 Nov 14;15(1):9886. doi: 10.1038/s41467-024-54182-5. PMID: 39543087; PMCID: PMC11564640.

APPENDIX:

Table 1 Literature Review Search Terms

Search	TERM
Pubmed	"(((((gene expression[MeSH Terms]) OR microrna [MeSH Terms]) OR micrornas [MeSH Terms]) OR biomarkers[MeSH Terms]) OR expression signature, gene[MeSH Terms]) OR gene expression signatures[MeSH Terms])) AND ((radiotherapy[Title] OR radiosensitivity[Title] OR radiation therapy[Title]))) NOT (((Review [Publication Type] OR case report[Publication Type]))))) NOT ((hypoxia OR imaging))". "MeSH" Terms represent medical subject headings in the National Library of Medicine. These identify descriptors in a hierarchical structure which enable various levels of searching specificity. An Additional advance search te rm was performed, it included, "((gene expression [MeSH Terms]) OR microrna [MeSH Terms]) OR micrornas [MeSH Terms]) OR biomarkers[MeSH Terms]) OR expression signature, gene[MeSH Terms]) OR gene expression signatures[MeSH Terms])) AND ((radiochemotherapy [Title]))"
Web of Science	"TITLE (radiotherapy OR radiosensitivity OR (radiation AND therapy)) AND (TITLE -ABS (signature OR (gene AND expression) OR biomarker OR microRNA OR microRNAs)) AND (LIMIT -TO (DOCTYPE, "ar")) AND NOT TITLE (hypoxia OR imaging)". "TITLE" represents the phrase present in the title, whereas "TITLE -ABS" represents the phrase present in the title or abstract. An additional advan ce search was included. It contained "TITLE (radiochemotherapy) AND (TITLE-ABS (signature OR (gene AND expression) OR biomarker OR microRNA OR microRNAs))"
Scopus	"(TI= (radiotherapy OR radiosensitivity OR radiation therapy) AND TS = (microRNA or microRNAs OR signature OR gene expression OR biomarker)) Reduce to Articles Only NOT TS= (hypoxia OR imaging)"

The above search terms were used for the literature review in; A) Pubmed, B) Web of Science and C) Scopus.

Appendix Figure 1. Framework for comparing radiosensitivity gene expression studies

Appendix Figure 2. Genes Increased in 3 Studies of Generated Radiation Resistant Cell Lines

Ensembl	name	description
ENSG00000136068	FLNB	filamin B [Source:HGNC Symbol;Acc:HGNC:3755]
ENSG00000136244	IL6	interleukin 6 [Source:HGNC Symbol;Acc:HGNC:6018]
ENSG00000112715	VEGFA	vascular endothelial growth factor A [Source:HGNC Symbol;Acc:HGNC:12680]
ENSG00000107796	ACTA2	actin alpha 2, smooth muscle [Source:HGNC Symbol;Acc:HGNC:130]
ENSG00000099139	PCSK5	proprotein convertase subtilisin/kexin type 5 [Source:HGNC Symbol;Acc:HGNC:8747]
ENSG00000135318	NT5E	5'-nucleotidase ecto [Source:HGNC Symbol;Acc:HGNC:8021]
ENSG00000118515	SGK1	serum/glucocorticoid regulated kinase 1 [Source:HGNC Symbol;Acc:HGNC:10810]
ENSG00000116717	GADD45A	growth arrest and DNA damage inducible alpha [Source:HGNC Symbol;Acc:HGNC:4095]

Appendix Figure 3. Genes Increased in 2 Studies of Generated Radiation Resistant Cell Lines

GDF15	ENSG0000	GDF15	growth differentiation factor 15 [Source:HGNC Symbol;Acc:HGNC:30142]
HERPUD1	ENSG0000	HERPUD1	homocysteine inducible ER protein with ubiquitin like domain 1 [Source:HGNC Symbol;Acc:HGN
IL1B	ENSG0000	IL1B	interleukin 1 beta [Source:HGNC Symbol;Acc:HGNC:5992]
GTPBP2	ENSG0000	GTPBP2	GTP binding protein 2 [Source:HGNC Symbol;Acc:HGNC:4670]
PLAU	ENSG0000	PLAU	plasminogen activator, urokinase [Source:HGNC Symbol;Acc:HGNC:9052]
CYP1A1	ENSG0000	CYP1A1	cytochrome P450 family 1 subfamily A member 1 [Source:HGNC Symbol;Acc:HGNC:2595]
EGR1	ENSG0000	EGR1	early growth response 1 [Source:HGNC Symbol;Acc:HGNC:3238]
CEP72	ENSG0000	CEP72	centrosomal protein 72 [Source:HGNC Symbol;Acc:HGNC:25547]
GPRC5A	ENSG0000	GPRC5A	G protein-coupled receptor class C group 5 member A [Source:HGNC Symbol;Acc:HGNC:9836]
CCND2	ENSG0000	CCND2	cyclin D2 [Source:HGNC Symbol;Acc:HGNC:1583]
TMPRSS3	ENSG0000	TMPRSS3	transmembrane serine protease 3 [Source:HGNC Symbol;Acc:HGNC:11877]
FOS	ENSG0000	FOS	Fos proto-oncogene, AP-1 transcription factor subunit [Source:HGNC Symbol;Acc:HGNC:3796]
TNFAIP3	ENSG0000	TNFAIP3	TNF alpha induced protein 3 [Source:HGNC Symbol;Acc:HGNC:11896]
RPS6KA2	ENSG0000	RPS6KA2	ribosomal protein S6 kinase A2 [Source:HGNC Symbol;Acc:HGNC:10431]
ALDH6A1	ENSG0000	ALDH6A1	aldehyde dehydrogenase 6 family member A1 [Source:HGNC Symbol;Acc:HGNC:7179]
RSAD2	ENSG0000	RSAD2	radical S-adenosyl methionine domain containing 2 [Source:HGNC Symbol;Acc:HGNC:30908]
PHLDB2	ENSG0000	PHLDB2	pleckstrin homology like domain family B member 2 [Source:HGNC Symbol;Acc:HGNC:29573]
IRAK2	ENSG0000	IRAK2	interleukin 1 receptor associated kinase 2 [Source:HGNC Symbol;Acc:HGNC:6113]
ATF3	ENSG0000	ATF3	activating transcription factor 3 [Source:HGNC Symbol;Acc:HGNC:785]
SDC4	ENSG0000	SDC4	syndecan 4 [Source:HGNC Symbol;Acc:HGNC:10661]
HBEGF	ENSG0000	HBEGF	heparin binding EGF like growth factor [Source:HGNC Symbol;Acc:HGNC:3059]
CXCL2	ENSG0000	CXCL2	C-X-C motif chemokine ligand 2 [Source:HGNC Symbol;Acc:HGNC:4603]
ERRFI1	ENSG0000	ERRFI1	ERBB receptor feedback inhibitor 1 [Source:HGNC Symbol;Acc:HGNC:18185]
MAFF	ENSG0000	MAFF	MAF bZIP transcription factor F [Source:HGNC Symbol;Acc:HGNC:6780]
IRF9	ENSG0000	IRF9	interferon regulatory factor 9 [Source:HGNC Symbol;Acc:HGNC:6131]
MUC1	ENSG0000	MUC1	mucin 1, cell surface associated [Source:HGNC Symbol;Acc:HGNC:7508]
PLAC8	ENSG0000	PLAC8	placenta associated 8 [Source:HGNC Symbol;Acc:HGNC:19254]
TRIB3	ENSG0000	TRIB3	tribbles pseudokinase 3 [Source:HGNC Symbol;Acc:HGNC:16228]
ANG	ENSG0000	ANG	angiogenin [Source:HGNC Symbol;Acc:HGNC:483]
COL5A1	ENSG0000	COL5A1	collagen type V alpha 1 chain [Source:HGNC Symbol;Acc:HGNC:2209]
TFPI	ENSG0000	TFPI	tissue factor pathway inhibitor [Source:HGNC Symbol;Acc:HGNC:11760]
CA9	ENSG0000	CA9	carbonic anhydrase 9 [Source:HGNC Symbol;Acc:HGNC:1383]
ACSL5	ENSG0000	ACSL5	acyl-CoA synthetase long chain family member 5 [Source:HGNC Symbol;Acc:HGNC:16526]
BIRC3	ENSG0000	BIRC3	baculoviral IAP repeat containing 3 [Source:HGNC Symbol;Acc:HGNC:591]
C11ORF86	ENSG0000	C11orf86	chromosome 11 open reading frame 86 [Source:HGNC Symbol;Acc:HGNC:34442]
TRIM31	ENSG0000	TRIM31	tripartite motif containing 31 [Source:HGNC Symbol;Acc:HGNC:16289]
TRIM2	ENSG0000	TRIM2	tripartite motif containing 2 [Source:HGNC Symbol;Acc:HGNC:15974]
UNC5B	ENSG0000	UNC5B	unc-5 netrin receptor B [Source:HGNC Symbol;Acc:HGNC:12568]
MB	ENSG0000	MB	myoglobin [Source:HGNC Symbol;Acc:HGNC:6915]
PHGDH	ENSG0000	PHGDH	phosphoglycerate dehydrogenase [Source:HGNC Symbol;Acc:HGNC:8923]
TMEM45A	ENSG0000	TMEM45A	transmembrane protein 45A [Source:HGNC Symbol;Acc:HGNC:25480]
MAML2	ENSG0000	MAML2	mastermind like transcriptional coactivator 2 [Source:HGNC Symbol;Acc:HGNC:16259]
CFI	ENSG0000	CFI	complement factor I [Source:HGNC Symbol;Acc:HGNC:5394]
SH3BP5	ENSG0000	SH3BP5	SH3 domain binding protein 5 [Source:HGNC Symbol;Acc:HGNC:10827]
NUPR1	ENSG0000	NUPR1	nuclear protein 1, transcriptional regulator [Source:HGNC Symbol;Acc:HGNC:29990]
EMP1	ENSG0000	EMP1	epithelial membrane protein 1 [Source:HGNC Symbol;Acc:HGNC:3333]
CPS1	ENSG0000	CPS1	carbamoyl-phosphate synthase 1 [Source:HGNC Symbol;Acc:HGNC:2323]
SLC1A4	ENSG0000	SLC1A4	solute carrier family 1 member 4 [Source:HGNC Symbol;Acc:HGNC:10942]
C11ORF96	ENSG0000	C11orf96	chromosome 11 open reading frame 96 [Source:HGNC Symbol;Acc:HGNC:38675]
SNAI2	ENSG0000	SNAI2	snail family transcriptional repressor 2 [Source:HGNC Symbol;Acc:HGNC:11094]
SAMD11	ENSG0000	SAMD11	sterile alpha motif domain containing 11 [Source:HGNC Symbol;Acc:HGNC:28706]
ASS1	ENSG0000	ASS1	argininosuccinate synthase 1 [Source:HGNC Symbol;Acc:HGNC:758]
PLEKHA2	ENSG0000	PLEKHA2	pleckstrin homology domain containing A2 [Source:HGNC Symbol;Acc:HGNC:14336]
METTL7A	ENSG0000	METTL7A	methyltransferase like 7A [Source:HGNC Symbol;Acc:HGNC:24550]
HEY1	ENSG0000	HEY1	hes related family bHLH transcription factor with YRPW motif 1 [Source:HGNC Symbol;Acc:HGNC
SCNN1A	ENSG0000	SCNN1A	sodium channel epithelial 1 subunit alpha [Source:HGNC Symbol;Acc:HGNC:10599]
AREG	ENSG0000	AREG	amphiregulin [Source:HGNC Symbol;Acc:HGNC:651]
RUFY2	ENSG0000	RUFY2	RUN and FYVE domain containing 2 [Source:HGNC Symbol;Acc:HGNC:19761]
TRIB2	ENSG0000	TRIB2	tribbles pseudokinase 2 [Source:HGNC Symbol;Acc:HGNC:30809]

SPOCK1	ENSG0000	SPOCK1	SPARC (osteonectin), cwcv and kazal like domains proteoglycan 1 [Source:HGNC Symbol;Acc:HGI
INHBA	ENSG0000	INHBA	inhibin subunit beta A [Source:HGNC Symbol;Acc:HGNC:6066]
IL7R	ENSG0000	IL7R	interleukin 7 receptor [Source:HGNC Symbol;Acc:HGNC:6024]
ABTB1	ENSG0000	ABTB1	ankyrin repeat and BTB domain containing 1 [Source:HGNC Symbol;Acc:HGNC:18275]
CCL2	ENSG0000	CCL2	C-C motif chemokine ligand 2 [Source:HGNC Symbol;Acc:HGNC:10618]
IFITM1	ENSG0000	IFITM1	interferon induced transmembrane protein 1 [Source:HGNC Symbol;Acc:HGNC:5412]
GNG11	ENSG0000	GNG11	G protein subunit gamma 11 [Source:HGNC Symbol;Acc:HGNC:4403]
IFITM2	ENSG0000	IFITM2	interferon induced transmembrane protein 2 [Source:HGNC Symbol;Acc:HGNC:5413]
RARRES3	None	None	None
MFNG	ENSG0000	MFNG	MFNG O-fucosylpeptide 3-beta-N-acetylglucosaminyltransferase [Source:HGNC Symbol;Acc:HGI
HIC1	ENSG0000	HIC1	HIC ZBTB transcriptional repressor 1 [Source:HGNC Symbol;Acc:HGNC:4909]
SLC1A3	ENSG0000	SLC1A3	solute carrier family 1 member 3 [Source:HGNC Symbol;Acc:HGNC:10941]
BST2	ENSG0000	BST2	bone marrow stromal cell antigen 2 [Source:HGNC Symbol;Acc:HGNC:1119]
TIMP3	ENSG0000	TIMP3	TIMP metallopeptidase inhibitor 3 [Source:HGNC Symbol;Acc:HGNC:11822]
SLC8A1	ENSG0000	SLC8A1	solute carrier family 8 member A1 [Source:HGNC Symbol;Acc:HGNC:11068]
CTIF	ENSG0000	CTIF	cap binding complex dependent translation initiation factor [Source:HGNC Symbol;Acc:HGNC:23
LYRM9	ENSG0000	LYRM9	LYR motif containing 9 [Source:HGNC Symbol;Acc:HGNC:27314]
SEMA7A	ENSG0000	SEMA7A	semaphorin 7A (John Milton Hagen blood group) [Source:HGNC Symbol;Acc:HGNC:10741]
SYNGR3	ENSG0000	SYNGR3	synaptogyrin 3 [Source:HGNC Symbol;Acc:HGNC:11501]
APOL6	ENSG0000	APOL6	apolipoprotein L6 [Source:HGNC Symbol;Acc:HGNC:14870]
IFNL1	ENSG0000	IFNL1	interferon lambda 1 [Source:HGNC Symbol;Acc:HGNC:18363]
PINK1	ENSG0000	PINK1	PTEN induced kinase 1 [Source:HGNC Symbol;Acc:HGNC:14581]
CPA4	ENSG0000	CPA4	carboxypeptidase A4 [Source:HGNC Symbol;Acc:HGNC:15740]
COL4A2	ENSG0000	COL4A2	collagen type IV alpha 2 chain [Source:HGNC Symbol;Acc:HGNC:2203]
GGTLC2	ENSG0000	GGTLC2	gamma-glutamyltransferase light chain 2 [Source:HGNC Symbol;Acc:HGNC:18596]
FSTL1	ENSG0000	FSTL1	follistatin like 1 [Source:HGNC Symbol;Acc:HGNC:3972]
GGTLC1	ENSG0000	GGTLC1	gamma-glutamyltransferase light chain 1 [Source:HGNC Symbol;Acc:HGNC:16437]
GGT1	ENSG0000	GGT1	gamma-glutamyltransferase 1 [Source:HGNC Symbol;Acc:HGNC:4250]
TTLL1	ENSG0000	TTLL1	tubulin tyrosine ligase like 1 [Source:HGNC Symbol;Acc:HGNC:1312]
DLX3	ENSG0000	DLX3	distal-less homeobox 3 [Source:HGNC Symbol;Acc:HGNC:2916]
TGFB1I1	ENSG0000	TGFB1I1	transforming growth factor beta 1 induced transcript 1 [Source:HGNC Symbol;Acc:HGNC:11767]
PODXL	ENSG0000	PODXL	podocalyxin like [Source:HGNC Symbol;Acc:HGNC:9171]
CD55	ENSG0000	CD55	CD55 molecule (Cromer blood group) [Source:HGNC Symbol;Acc:HGNC:2665]
LAMA3	ENSG0000	LAMA3	laminin subunit alpha 3 [Source:HGNC Symbol;Acc:HGNC:6483]
NR0B1	ENSG0000	NR0B1	nuclear receptor subfamily 0 group B member 1 [Source:HGNC Symbol;Acc:HGNC:7960]
KCTD4	ENSG0000	KCTD4	potassium channel tetramerization domain containing 4 [Source:HGNC Symbol;Acc:HGNC:23227
POT1	ENSG0000	POT1	protection of telomeres 1 [Source:HGNC Symbol;Acc:HGNC:17284]
GJA1	ENSG0000	GJA1	gap junction protein alpha 1 [Source:HGNC Symbol;Acc:HGNC:4274]
IGF2BP3	ENSG0000	IGF2BP3	insulin like growth factor 2 mRNA binding protein 3 [Source:HGNC Symbol;Acc:HGNC:28868]
ARNTL	ENSG0000	ARNTL	aryl hydrocarbon receptor nuclear translocator like [Source:HGNC Symbol;Acc:HGNC:701]
RRAD	ENSG0000	RRAD	RRAD, Ras related glycolysis inhibitor and calcium channel regulator [Source:HGNC Symbol;Acc:H
CLU	ENSG0000	CLU	clusterin [Source:HGNC Symbol;Acc:HGNC:2095]
DDIT4	ENSG0000	DDIT4	DNA damage inducible transcript 4 [Source:HGNC Symbol;Acc:HGNC:24944]
CITED2	ENSG0000	CITED2	Cbp/p300 interacting transactivator with Glu/Asp rich carboxy-terminal domain 2 [Source:HGNC
LOX	ENSG0000	LOX	lysyl oxidase [Source:HGNC Symbol;Acc:HGNC:6664]
BNIP3L	ENSG0000	BNIP3L	BCL2 interacting protein 3 like [Source:HGNC Symbol;Acc:HGNC:1085]
SLC2A1	ENSG0000	SLC2A1	solute carrier family 2 member 1 [Source:HGNC Symbol;Acc:HGNC:11005]
BHLHE40	ENSG0000	BHLHE40	basic helix-loop-helix family member e40 [Source:HGNC Symbol;Acc:HGNC:1046]
PLOD2	ENSG0000	PLOD2	procollagen-lysine,2-oxoglutarate 5-dioxygenase 2 [Source:HGNC Symbol;Acc:HGNC:9082]
NFIL3	ENSG0000	NFIL3	nuclear factor, interleukin 3 regulated [Source:HGNC Symbol;Acc:HGNC:7787]
FOSL2	ENSG0000	FOSL2	FOS like 2, AP-1 transcription factor subunit [Source:HGNC Symbol;Acc:HGNC:3798]
FZD10	ENSG0000	FZD10	frizzled class receptor 10 [Source:HGNC Symbol;Acc:HGNC:4039]
STC2	ENSG0000	STC2	stanniocalcin 2 [Source:HGNC Symbol;Acc:HGNC:11374]
IRS2	ENSG0000	IRS2	insulin receptor substrate 2 [Source:HGNC Symbol;Acc:HGNC:6126]
P4HA1	ENSG0000	P4HA1	prolyl 4-hydroxylase subunit alpha 1 [Source:HGNC Symbol;Acc:HGNC:8546]
SLC2A3	ENSG0000	SLC2A3	solute carrier family 2 member 3 [Source:HGNC Symbol;Acc:HGNC:11007]
TRIB1	ENSG0000	TRIB1	tribbles pseudokinase 1 [Source:HGNC Symbol;Acc:HGNC:16891]
ENO2	ENSG0000	ENO2	enolase 2 [Source:HGNC Symbol;Acc:HGNC:3353]

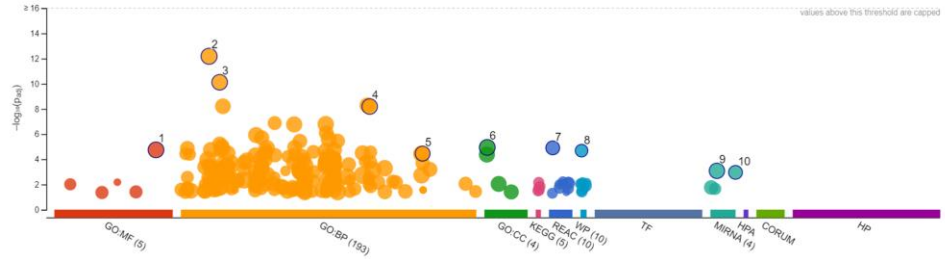
Appendix Figure 4 Genes Decreased in 3 Studies of Generated Radiation Resistant Cell Lines

initial_ali	converted	name	description
KRT14	ENSG0000	KRT14	keratin 14 [Source:HGNC Symbol;Acc:HGNC:6416]

Appendix Figure 5 Genes Decreased in 2 Studies of Generated Radiation Resistant Cell Lines

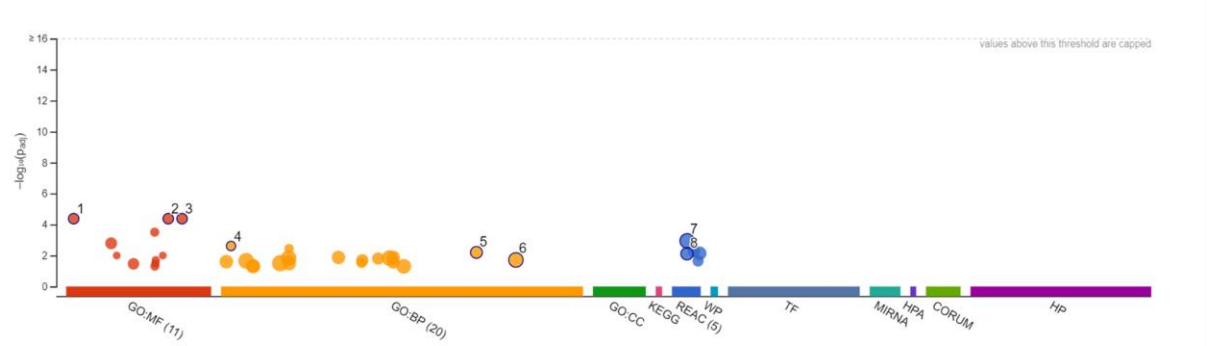
initial_ali	converted	name	description
LCN2	ENSG0000	LCN2	lipocalin 2 [Source:HGNC Symbol;Acc:HGNC:6526]
PLAU	ENSG0000	PLAU	plasminogen activator, urokinase [Source:HGNC Symbol;Acc:HGNC:9052]
SEPHS2	ENSG0000	SEPHS2	selenophosphate synthetase 2 [Source:HGNC Symbol;Acc:HGNC:19686]
AKR1C1	ENSG0000	AKR1C1	aldo-keto reductase family 1 member C1 [Source:HGNC Symbol;Acc:HGNC:384]
AKR1C3	ENSG0000	AKR1C3	aldo-keto reductase family 1 member C3 [Source:HGNC Symbol;Acc:HGNC:386]
GJA1	ENSG0000	GJA1	gap junction protein alpha 1 [Source:HGNC Symbol;Acc:HGNC:4274]
NFYA	ENSG0000	NFYA	nuclear transcription factor Y subunit alpha [Source:HGNC Symbol;Acc:HGNC:7804]
RMDN2	ENSG0000	RMDN2	regulator of microtubule dynamics 2 [Source:HGNC Symbol;Acc:HGNC:26567]
UGT8	ENSG0000	UGT8	UDP glycosyltransferase 8 [Source:HGNC Symbol;Acc:HGNC:12555]
ANKRD22	ENSG0000	ANKRD22	ankyrin repeat domain 22 [Source:HGNC Symbol;Acc:HGNC:28321]
PLA2G4A	ENSG0000	PLA2G4A	phospholipase A2 group IVA [Source:HGNC Symbol;Acc:HGNC:9035]
ANKRD29	ENSG0000	ANKRD29	ankyrin repeat domain 29 [Source:HGNC Symbol;Acc:HGNC:27110]
KRT6A	ENSG0000	KRT6A	keratin 6A [Source:HGNC Symbol;Acc:HGNC:6443]
POLR3G	ENSG0000	POLR3G	RNA polymerase III subunit G [Source:HGNC Symbol;Acc:HGNC:30075]
SCG5	ENSG0000	SCG5	secretogranin V [Source:HGNC Symbol;Acc:HGNC:10816]
HSPA1A	ENSG0000	HSPA1A	heat shock protein family A (Hsp70) member 1A [Source:HGNC Symbol;Acc:HGNC:5232]
IL1RL1	ENSG0000	IL1RL1	interleukin 1 receptor like 1 [Source:HGNC Symbol;Acc:HGNC:5998]
HS3ST1	ENSG0000	HS3ST1	heparan sulfate-glucosamine 3-sulfotransferase 1 [Source:HGNC Symbol;Acc:HGNC:5194]
PLD6	ENSG0000	PLD6	phospholipase D family member 6 [Source:HGNC Symbol;Acc:HGNC:30447]
UBASH3B	ENSG0000	UBASH3B	ubiquitin associated and SH3 domain containing B [Source:HGNC Symbol;Acc:HGNC:29884]
FAR2	ENSG0000	FAR2	fatty acyl-CoA reductase 2 [Source:HGNC Symbol;Acc:HGNC:25531]
RAB7B	ENSG0000	RAB7B	RAB7B, member RAS oncogene family [Source:HGNC Symbol;Acc:HGNC:30513]
BST1	ENSG0000	BST1	bone marrow stromal cell antigen 1 [Source:HGNC Symbol;Acc:HGNC:1118]
HLX	ENSG0000	HLX	H2.0 like homeobox [Source:HGNC Symbol;Acc:HGNC:4978]
VCX2	ENSG0000	VCX2	variable charge X-linked 2 [Source:HGNC Symbol;Acc:HGNC:18158]
PFKFB3	ENSG0000	PFKFB3	6-phosphofructo-2-kinase/fructose-2,6-biphosphatase 3 [Source:HGNC Symbol;Acc:HGNC:887
SYK	ENSG0000	SYK	spleen associated tyrosine kinase [Source:HGNC Symbol;Acc:HGNC:11491]
JPH1	ENSG0000	JPH1	junctophilin 1 [Source:HGNC Symbol;Acc:HGNC:14201]
NRG4	ENSG0000	NRG4	neuregulin 4 [Source:HGNC Symbol;Acc:HGNC:29862]
NPR3	ENSG0000	NPR3	natriuretic peptide receptor 3 [Source:HGNC Symbol;Acc:HGNC:7945]
HMX1	ENSG0000	HMX1	H6 family homeobox 1 [Source:HGNC Symbol;Acc:HGNC:5017]
EML1	ENSG0000	EML1	EMAP like 1 [Source:HGNC Symbol;Acc:HGNC:3330]
SLC2A9	ENSG0000	SLC2A9	solute carrier family 2 member 9 [Source:HGNC Symbol;Acc:HGNC:13446]
TLR2	ENSG0000	TLR2	toll like receptor 2 [Source:HGNC Symbol;Acc:HGNC:11848]
CDCA4	ENSG0000	CDCA4	cell division cycle associated 4 [Source:HGNC Symbol;Acc:HGNC:14625]
DHRS9	ENSG0000	DHRS9	dehydrogenase/reductase 9 [Source:HGNC Symbol;Acc:HGNC:16888]
RBP1	ENSG0000	RBP1	retinol binding protein 1 [Source:HGNC Symbol;Acc:HGNC:9919]
PLBD1	ENSG0000	PLBD1	phospholipase B domain containing 1 [Source:HGNC Symbol;Acc:HGNC:26215]
EPB41L4A	ENSG0000	EPB41L4A	erythrocyte membrane protein band 4.1 like 4A [Source:HGNC Symbol;Acc:HGNC:13278]
ASB2	ENSG0000	ASB2	ankyrin repeat and SOCS box containing 2 [Source:HGNC Symbol;Acc:HGNC:16012]
NMU	ENSG0000	NMU	neuromedin U [Source:HGNC Symbol;Acc:HGNC:7859]
CD38	ENSG0000	CD38	CD38 molecule [Source:HGNC Symbol;Acc:HGNC:1667]
ST6GALNA	ENSG0000	ST6GALNA	ST6 N-acetylgalactosaminide alpha-2,6-sialyltransferase 1 [Source:HGNC Symbol;Acc:HGNC:23
CA2	ENSG0000	CA2	carbonic anhydrase 2 [Source:HGNC Symbol;Acc:HGNC:1373]
CCDC3	ENSG0000	CCDC3	coiled-coil domain containing 3 [Source:HGNC Symbol;Acc:HGNC:23813]
GALNT14	ENSG0000	GALNT14	polypeptide N-acetylgalactosaminyltransferase 14 [Source:HGNC Symbol;Acc:HGNC:22946]
LAMA1	ENSG0000	LAMA1	laminin subunit alpha 1 [Source:HGNC Symbol;Acc:HGNC:6481]
BMP7	ENSG0000	BMP7	bone morphogenetic protein 7 [Source:HGNC Symbol;Acc:HGNC:1074]
HNRNPD	ENSG0000	HNRNPD	heterogeneous nuclear ribonucleoprotein D [Source:HGNC Symbol;Acc:HGNC:5036]
LRAT	ENSG0000	LRAT	lecithin retinol acyltransferase [Source:HGNC Symbol;Acc:HGNC:6685]
PLP1	ENSG0000	PLP1	proteolipid protein 1 [Source:HGNC Symbol;Acc:HGNC:9086]
GCH1	ENSG0000	GCH1	GTP cyclohydrolase 1 [Source:HGNC Symbol;Acc:HGNC:4193]

Appendix Figure 6 Pathway Analysis of Genes Increased in Radioresistant Studies



ID	Source	Term ID	Term Name	p_{adj} (query_1)
1	GO:MF	GO:0098772	molecular function regulator	1.785×10^{-5}
2	GO:BP	GO:0006950	response to stress	6.524×10^{-13}
3	GO:BP	GO:0009605	response to external stimulus	7.604×10^{-11}
4	GO:BP	GO:0071310	cellular response to organic substance	6.436×10^{-9}
5	GO:BP	GO:1901700	response to oxygen-containing compound	3.510×10^{-5}
6	GO:CC	GO:0005615	extracellular space	1.136×10^{-5}
7	REAC	REAC:R-HSA-12...	Cytokine Signaling in Immune system	1.239×10^{-5}
8	WP	WP:WP2882	Nuclear Receptors Meta-Pathway	2.047×10^{-5}
9	MIRNA	MIRNA:hsa-miR...	hsa-miR-335-5p	8.238×10^{-4}
10	MIRNA	MIRNA:hsa-miR...	hsa-miR-98-5p	1.089×10^{-3}

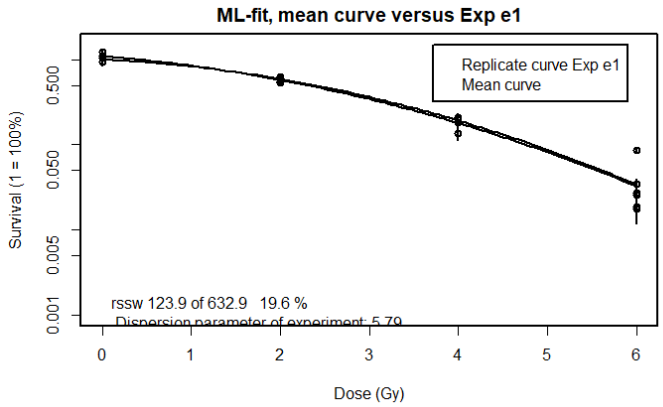
Appendix Figure 7 Pathway Analysis of Genes Decreased in Radioresistant Studies

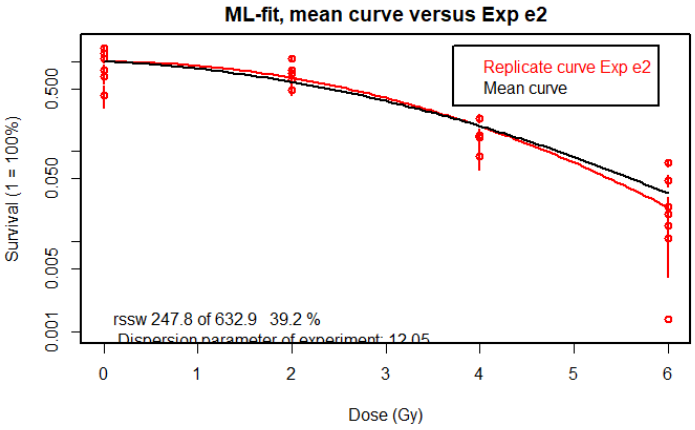


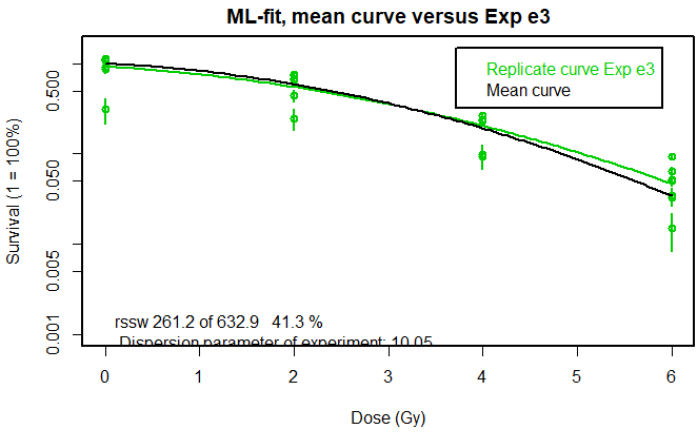
ID	Source	Term ID	Term Name	p_{adj} (query_1)
1	GO:MF	GO:0003953	NAD+ nucleosidase activity	4.190×10^{-5}
2	GO:MF	GO:0050135	NAD(P)+ nucleosidase activity	4.190×10^{-5}
3	GO:MF	GO:0061809	NAD+ nucleotidase, cyclic ADP-ribose generating	4.190×10^{-5}
4	GO:BP	GO:0002138	retinoic acid biosynthetic process	2.424×10^{-3}
5	GO:BP	GO:0090322	regulation of superoxide metabolic process	6.294×10^{-3}
6	GO:BP	GO:1901568	fatty acid derivative metabolic process	1.930×10^{-2}
7	REAC	REAC:R-HSA-19...	Metabolism of vitamins and cofactors	1.142×10^{-3}
8	REAC	REAC:R-HSA-68...	Metabolism of fat-soluble vitamins	7.694×10^{-3}

Appendix Figure1.1 Additional Survival Curves HT29

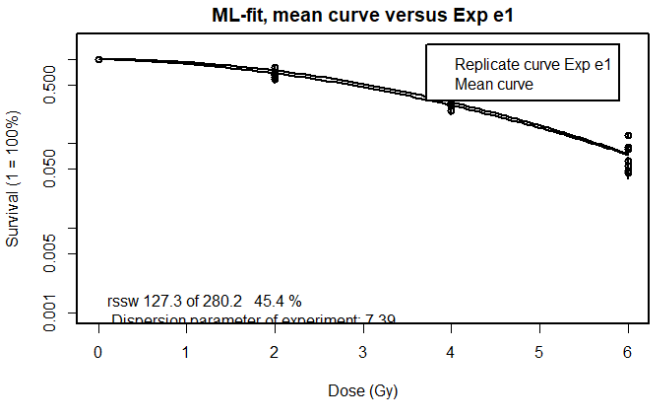
HT29_BASE

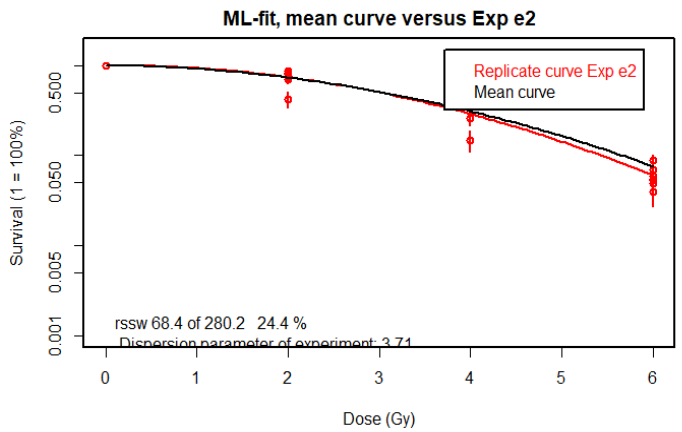


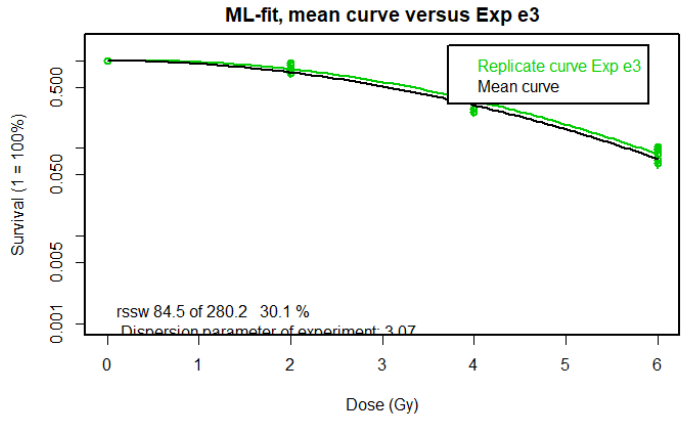




HT29 60A

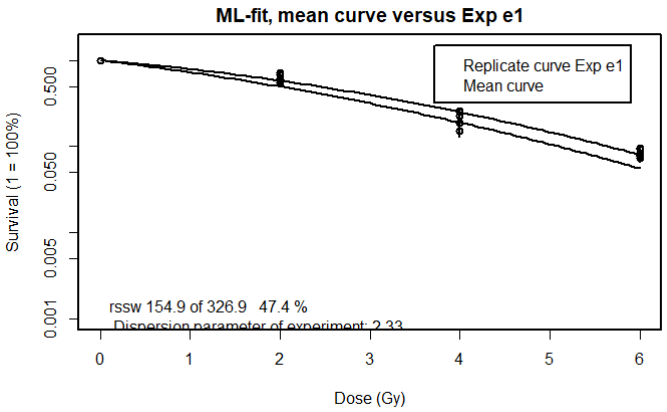


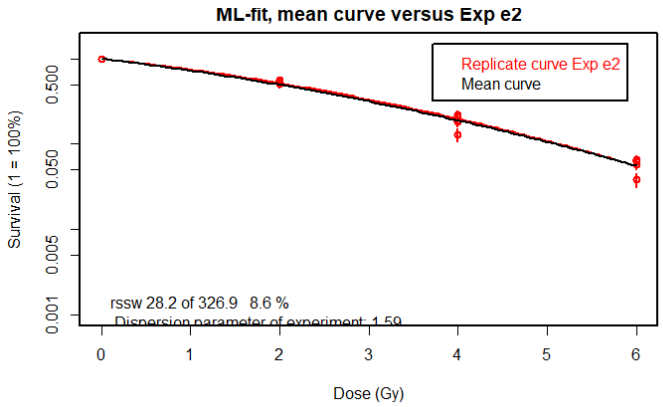


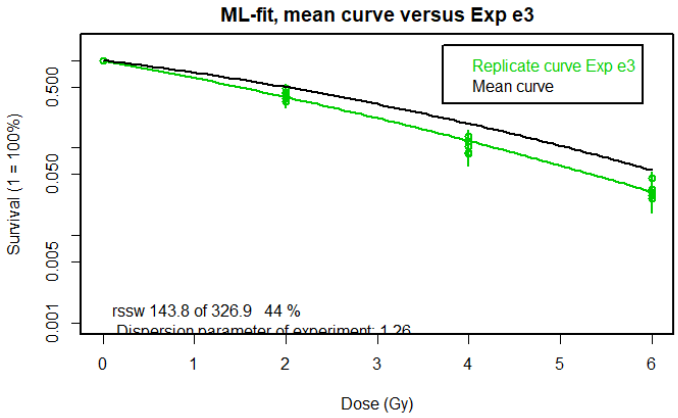


Appendix Figure 1.2 Additional Survival Curves A549

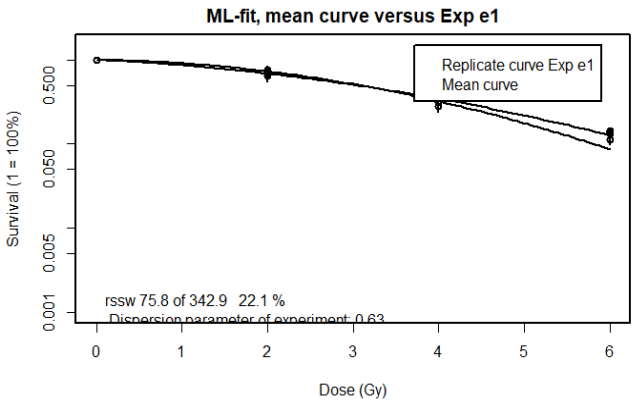
A549 BASE

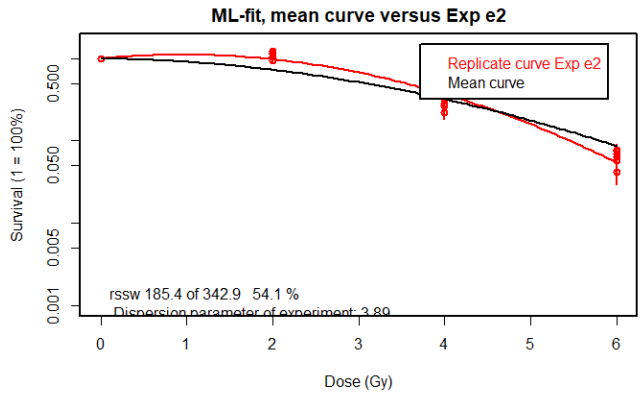


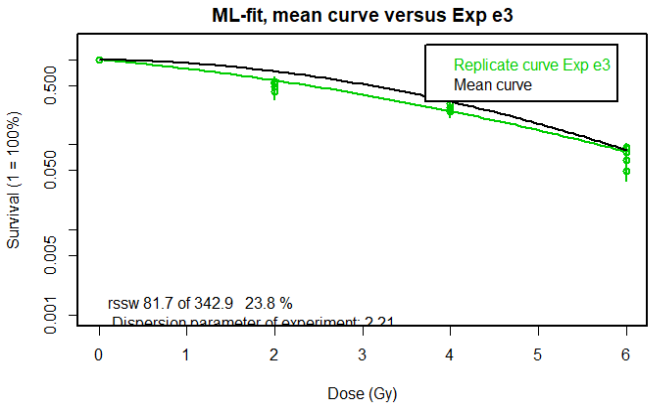




A549-60A

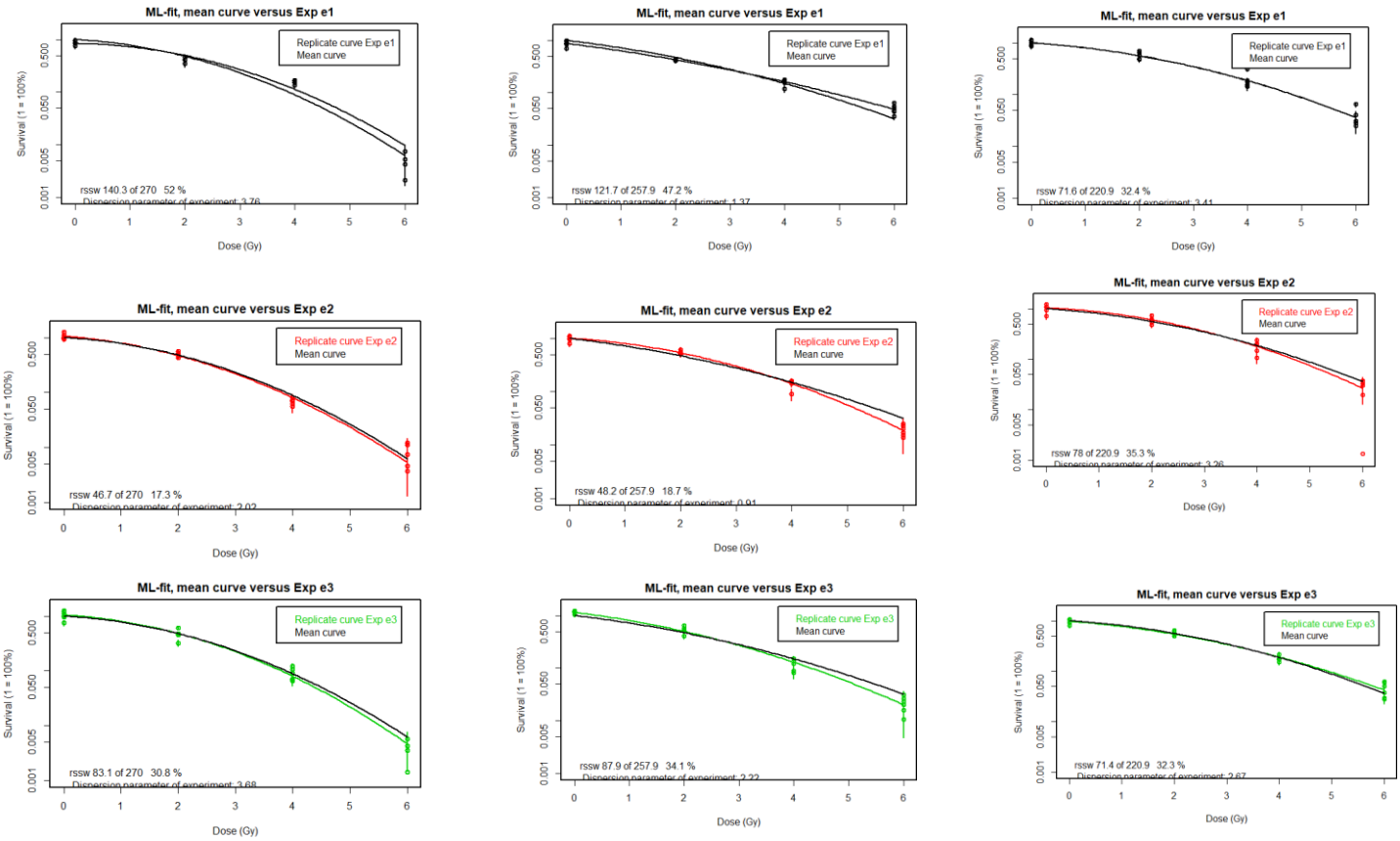






Appendix Figure 1.3 Additional Survival Curves H460

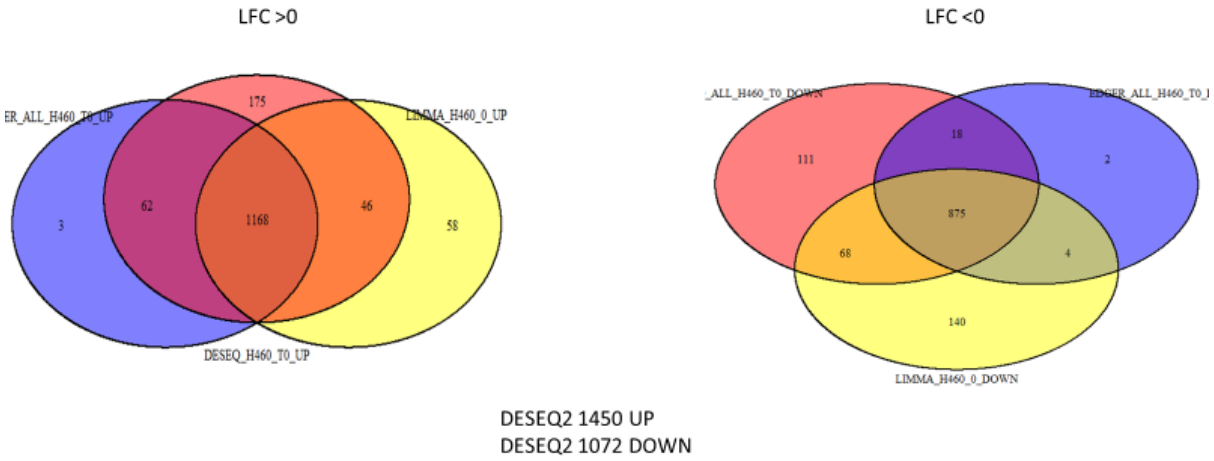
H460 Base **H46060A** **H46050B**



Appendix Table 2.1 Comparison metrics for different mapping pipelines and techniques - Quality Control metrics were obtained for samples using FastQC

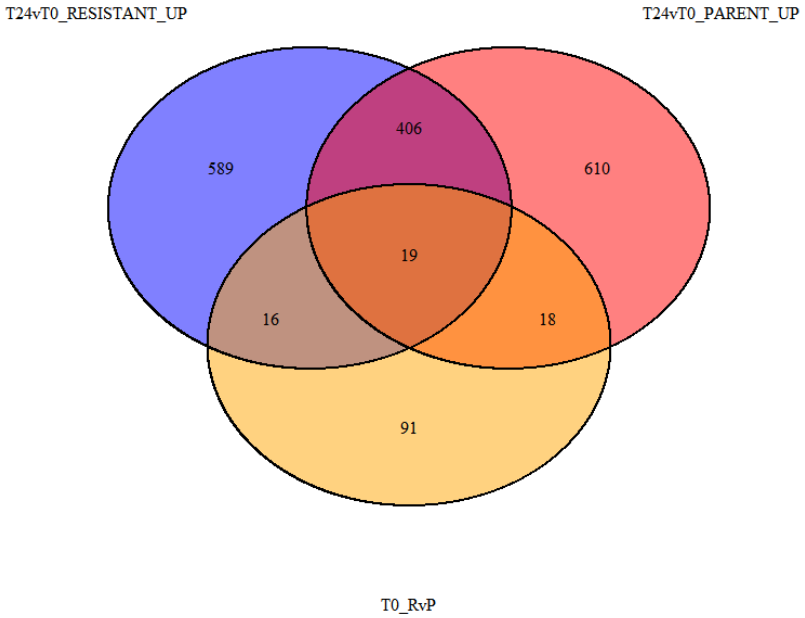
PIPELINE	Mean Unique Mapped Reads	% Uniquely Mapped	Samples <1million reads
STAR – no trim	906186	21.45325	42
STAR + trim	2888800	62.9194	6
Subread (no trim)	2014268	51.0631	9
Subread + trim	2819270	61.5717	9

Appendix Figure 2.1 Comparison of Differential Expression Techniques

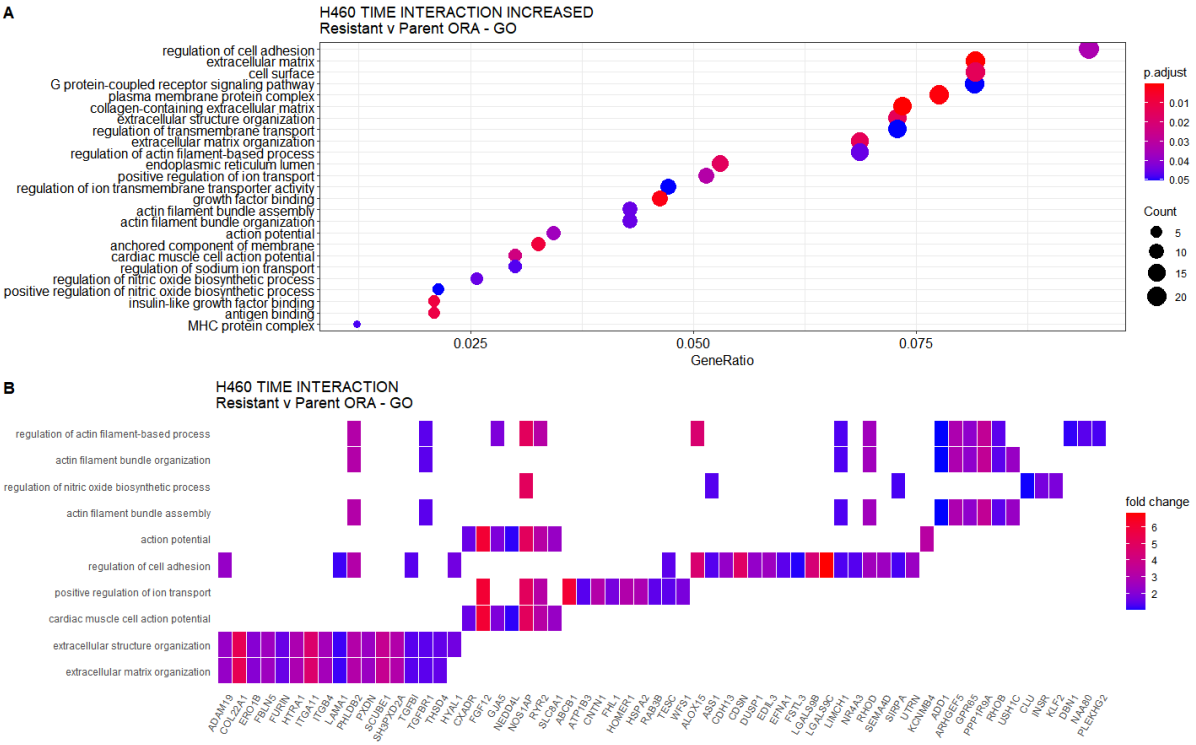


3 separate differential expression techniques were used to identify an overlapping set of genes that were differentially expressed between parental and radioresistant cell lines at time 0, i.e before radiation. The venn diagram illustrates the overlap between differential expression techniques with 1168 increased and 875 decreased, padj <0.05. Confirming that >80% of differentially expressed genes are concordant between DE analyses. DESEQ2 was used for the remainder of the analysis

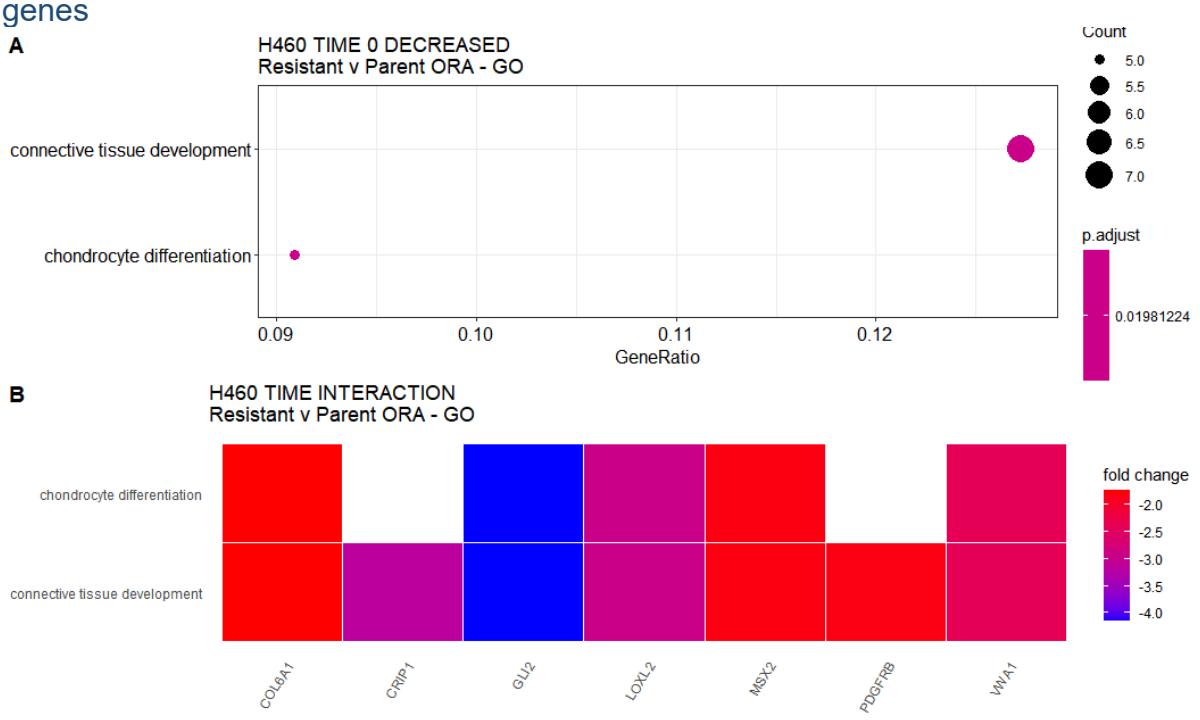
Appendix Figure 2.2 Overlap T24VT0 Resistant, T24vT0 Parent T0RvP



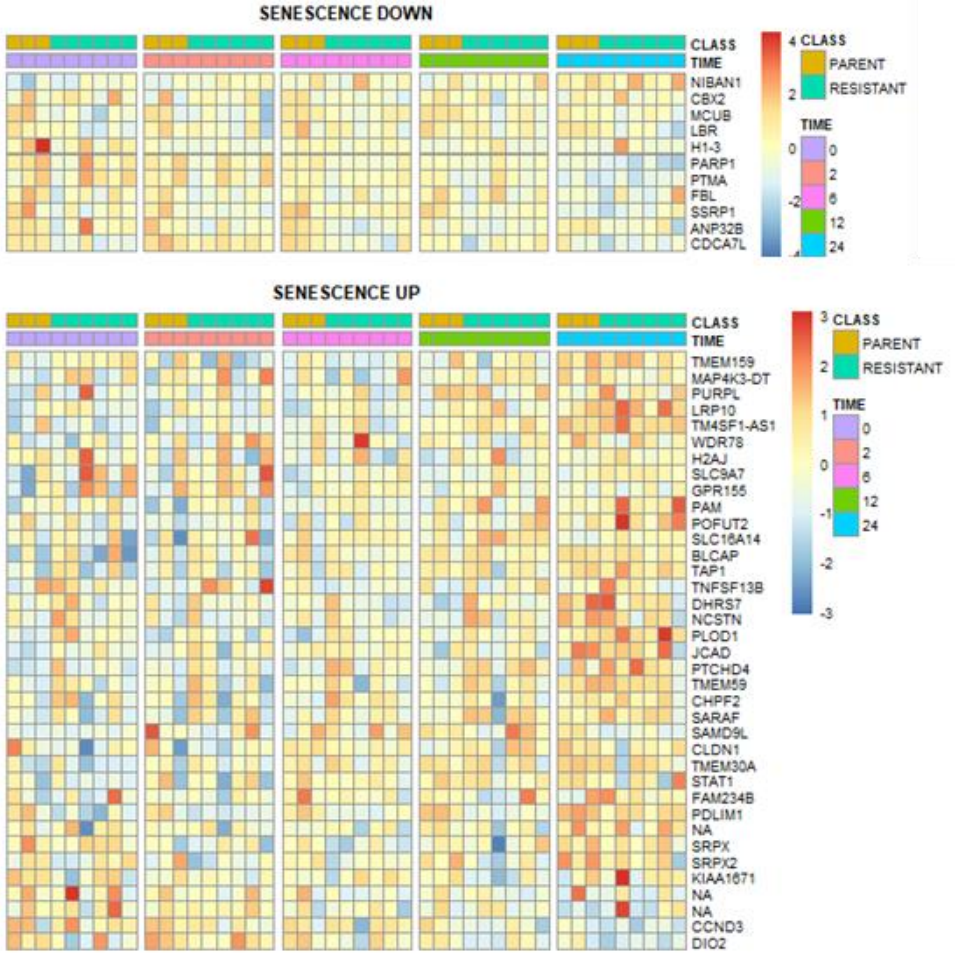
Appendix Figure 2.3 Dotplot and Heatplot ORA Time interaction Model Increased genes



Appendix Figure 2.4 Dotplot and Heatplot ORA Time interaction Model Increased genes



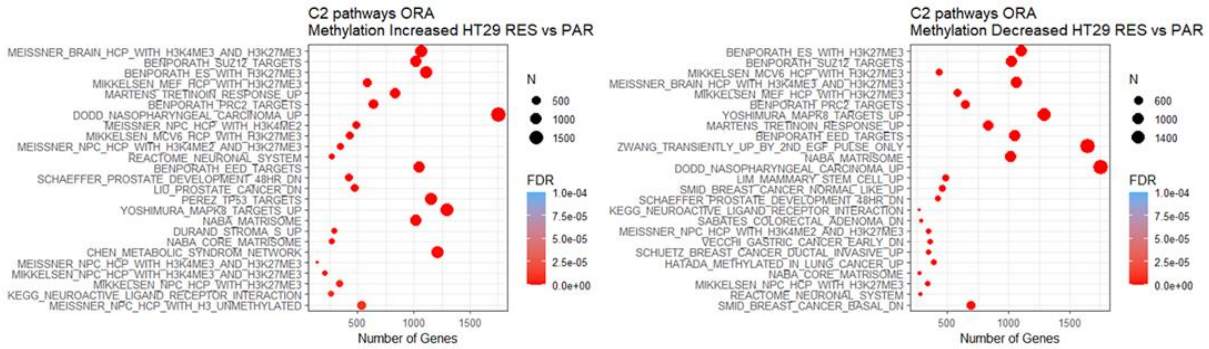
Appendix Figure 2.5 Heatmap of Senescence Signatures



Appendix Figure 2.6 Hallmark Pathways

Pathway	logFC	AveExpr	t	P.Value	adj.P.Val	B
HALLMARK_ANGIOGENESIS	0.047789	0.29235	7.799839	1.33E-09	6.64E-08	11.22294
HALLMARK_ESTROGEN_RESPONSE_EARLY	0.024251	0.171696	7.397556	4.79E-09	1.20E-07	9.934458
HALLMARK_NOTCH_SIGNALING	-0.03131	0.0718	-5.64004	1.44E-06	2.40E-05	4.232659
HALLMARK_UV_RESPONSE_DN	0.015938	0.063618	4.560278	4.62E-05	0.000577873	0.808893
HALLMARK_HYPOXIA	0.021467	0.240455	4.150154	0.000165	0.001403184	-0.43529
HALLMARK_IL6_JAK_STAT3_SIGNALING	-0.01262	0.098078	-4.14404	0.000168	0.001403184	-0.4535
HALLMARK_CHOLESTEROL_HOMEOSTASIS	0.017782	0.282448	3.843843	0.000417	0.002981198	-1.33342
HALLMARK_IL2_STAT5_SIGNALING	0.009682	0.150914	3.699415	0.00064	0.004001507	-1.74588
HALLMARK_G2M_CHECKPOINT	-0.02827	0.37274	-3.61814	0.000812	0.004512559	-1.97454
HALLMARK_SPERMATOGENESIS	-0.01571	0.126483	-3.50368	0.001132	0.005619943	-2.29213
HALLMARK_E2F_TARGETS	-0.03014	0.418974	-3.47282	0.001236	0.005619943	-2.37683
HALLMARK_MYOGENESIS	0.012286	0.146641	3.288976	0.002081	0.008672269	-2.87274
HALLMARK_MYC_TARGETS_V1	-0.01897	0.725349	-3.25987	0.002258	0.008683204	-2.94985
HALLMARK_TGF_BETA_SIGNALING	0.0125	0.274519	3.04829	0.004035	0.014411797	-3.49789
HALLMARK_ESTROGEN_RESPONSE_LATE	0.010974	0.208566	3.000258	0.004592	0.015306677	-3.61914
HALLMARK_MITOTIC_SPINDLE	-0.01648	0.260263	-2.91128	0.005819	0.01818282	-3.84047
HALLMARK_INTERFERON_GAMMA_RESPONSE	-0.00823	0.097396	-2.59132	0.013218	0.038876015	-4.59895

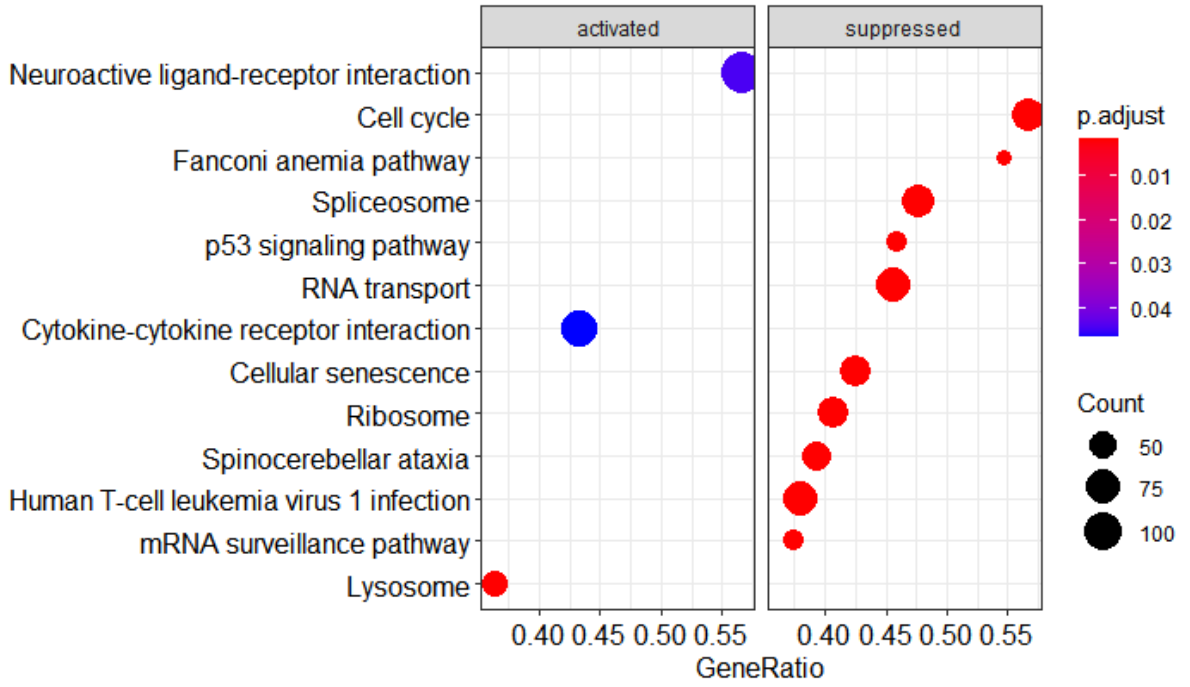
Appendix Figure 2.7 Methylation ORA C2 Pathways HT29



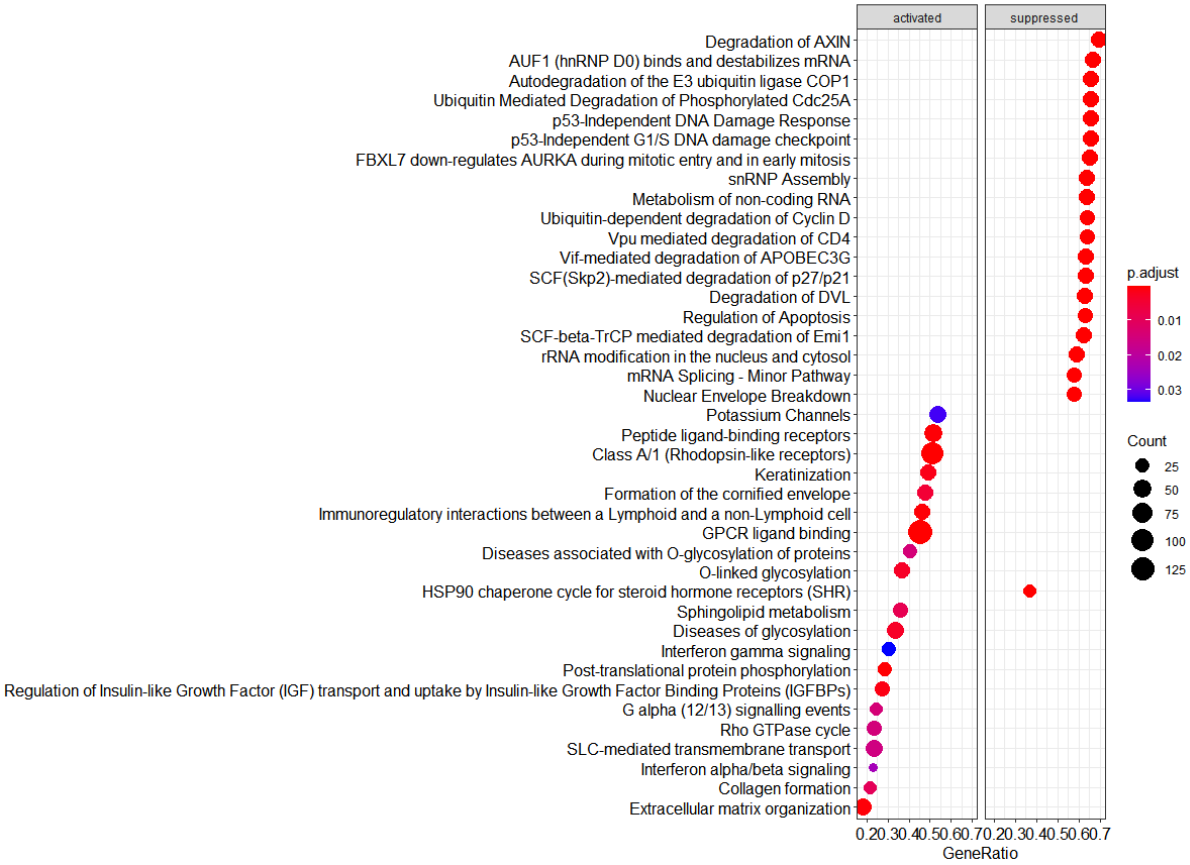
Appendix Figure 2.8 FAM83A Correlation (Spearman)

SYMBOL	ENTREZ	Correlation_Value	p_value
FAM83A	84985	1	0
COL22A1	169044	0.862582345	0
MGLL	11343	0.832806324	0
VWCE	220001	0.829512516	0
NCAM1	4684	0.82055336	0
DENND1A	57706	0.81541502	0
MYO1E	4643	0.808036891	0
NA	105375612	0.795388669	0
ABHD17C	58489	0.788142292	0
CALB2	794	0.77483531	2.42E-09
HYAL1	3373	0.7743083	2.88E-09
SLFN11	91607	-0.704874835	2.52E-07
BACE1	23621	-0.710408432	2.01E-07
AUTS2	26053	-0.726613966	1.04E-07
HK1	3098	-0.734387352	7.48E-08
ADAM12	8038	-0.747957839	3.96E-08
CD24	100133941	-0.751383399	3.30E-08
SERPINB2	5055	-0.756389987	2.45E-08
CTNND2	1501	-0.79314888	0
CDH2	1000	-0.811067194	0
GABBR2	9568	-0.843346509	0

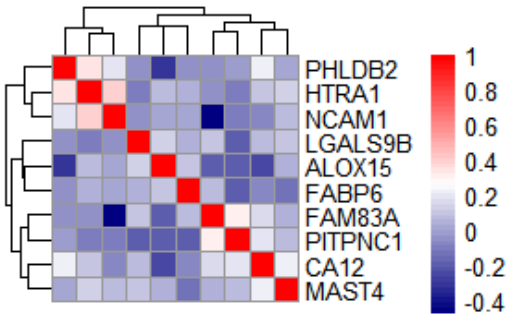
Appendix Figure 2.9 FAM83A Correlation GSEA KEGG



Appendix Figure 2.10 FAM83A Correlation GSEA Reactome



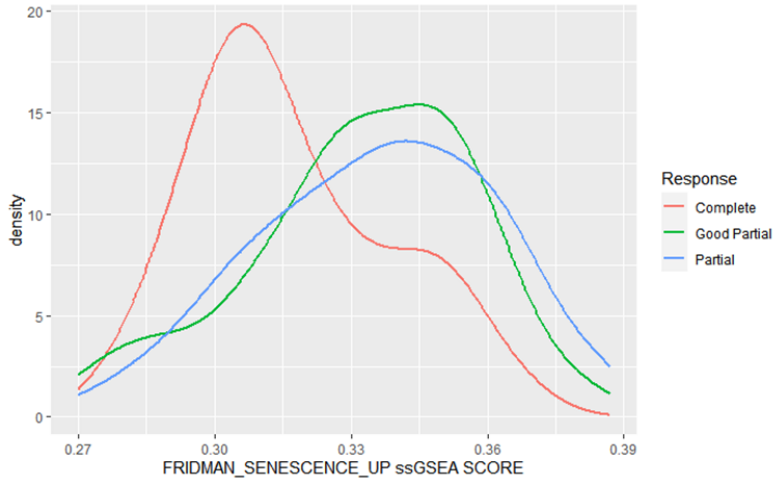
Appendix Figure 2.11 Correlation H460 Increased Resistant genes in TCGA



Appendix Figure 2.12 Top Differentially Methylated Regions

	seqnames	start	end	width	strand	no.cpgs	min_smoothed_fdr	Stouffer	HMFDR	Fisher	maxdiff	meandiff	overlapping.genes
1	chr8	1.24E+08	1.24E+08	5183	*	30	2.32E-302	1.61E-40	2.96E-05	2.54E-46	-0.38262	-0.0732	FAM83A, U3, RP11-539E17.5
2	chr2	1.45E+08	1.45E+08	6274	*	45	0	7.73E-30	1.19E-05	1.05E-45	0.483522	0.090771	ZEB2-AS1, ZEB2
3	chr1	24643779	24646392	2614	*	20	5.44E-185	1.89E-40	9.84E-05	5.32E-37	-0.38556	-0.19332	GRHL3, RP11-10N16.3
4	chr8	1.14E+08	1.14E+08	6782	*	36	1.82E-264	4.53E-29	2.94E-05	7.64E-36	0.280395	0.058559	CSMD3
5	chr3	1.96E+08	1.96E+08	1189	*	11	1.94E-296	1.72E-36	9.07E-06	5.57E-34	-0.24345	-0.15679	TM4SF19, TM4SF19
6	chr15	31194984	31196599	1616	*	14	2.57E-304	1.43E-32	9.76E-06	6.55E-34	-0.47938	-0.22685	FAN1
7	chr8	1.01E+08	1.01E+08	1341	*	11	0	2.75E-37	3.89E-06	6.59E-34	-0.30182	-0.14799	RNF19A
8	chr7	45959667	45963024	3358	*	39	4.65E-94	4.76E-34	0.000832	3.22E-33	0.297763	0.096712	IGFBP3
9	chr19	46282571	46286683	4113	*	19	0	5.54E-32	1.30E-06	1.07E-32	-0.503	-0.13782	DMPK, AC011530.4, DMWD
10	chr8	21988147	21994573	6427	*	19	4.64E-134	3.27E-28	3.49E-05	1.88E-32	0.269035	0.115608	HR
11	chr6	48035236	48037415	2180	*	17	1.22E-179	2.23E-34	5.92E-05	2.13E-32	0.252989	0.127928	PTCHD4
12	chr11	2016513	2020560	4048	*	58	1.14E-120	1.15E-21	0.000123	3.46E-31	-0.28112	-0.04851	H19
13	chr6	32054094	32058396	4303	*	45	9.08E-245	5.79E-23	7.88E-05	8.85E-31	-0.36281	-0.07104	TNXB
14	chr8	10585394	10591081	5688	*	26	3.06E-160	9.48E-22	0.000101	2.75E-30	0.334451	0.083517	CTD-2135J3.3, SOX7, SOX7
15	chr2	96808802	96813398	4597	*	17	3.79E-136	3.19E-27	2.96E-05	2.19E-29	-0.387	-0.13006	DUSP2, AC012307.2
16	chr9	85676347	85679000	2654	*	14	3.18E-229	6.92E-28	1.19E-05	3.46E-29	-0.44166	-0.102	RASEF
17	chr6	33128903	33138108	9206	*	85	4.37E-66	1.76E-25	0.001767	2.37E-27	0.196221	-0.01004	COL11A2
18	chr6	33156164	33167772	11609	*	101	6.04E-72	1.15E-17	0.000329	3.12E-27	-0.21234	-0.02326	RNY4P10, COL11A2, RXRB
19	chr12	48395868	48399493	3626	*	21	2.38E-180	1.07E-22	3.52E-05	3.46E-27	-0.47409	-0.08485	COL2A1, RP1-228P16.3
20	chr10	1514391	1518725	4335	*	11	9.14E-134	1.74E-27	1.84E-05	7.95E-27	-0.33447	-0.1914	ADARB2
21	chr4	4866800	4874172	7373	*	28	7.81E-81	8.74E-23	6.83E-05	1.28E-26	0.339119	-0.02874	NA
22	chr5	1.49E+08	1.49E+08	2684	*	8	1.93E-180	7.03E-27	5.27E-06	2.07E-26	-0.26071	-0.16672	PDE6A, Y_RNA
23	chr15	26107658	26111090	3433	*	34	6.37E-107	6.44E-21	0.000608	2.41E-26	0.198793	-0.02818	RP11-2C7.1, ATP10A
24	chr11	2158438	2163538	5101	*	66	3.50E-78	3.32E-22	0.000848	4.15E-26	0.385885	-0.02538	IGF2-AS, IGF2, INS-IGF2
25	chr3	49314155	49315853	1699	*	16	1.07E-221	7.59E-19	2.11E-05	1.47E-25	0.420759	0.161423	C3orf62, USP4

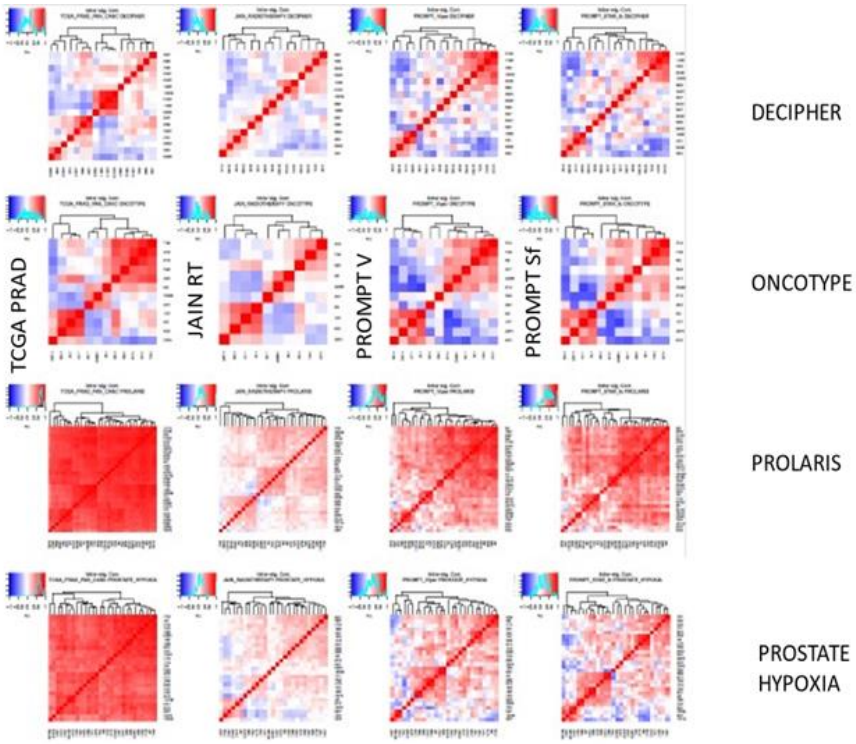
Appendix Figurex 2.13 Per Outcome Density Distribution Fridman_Senescence_Up



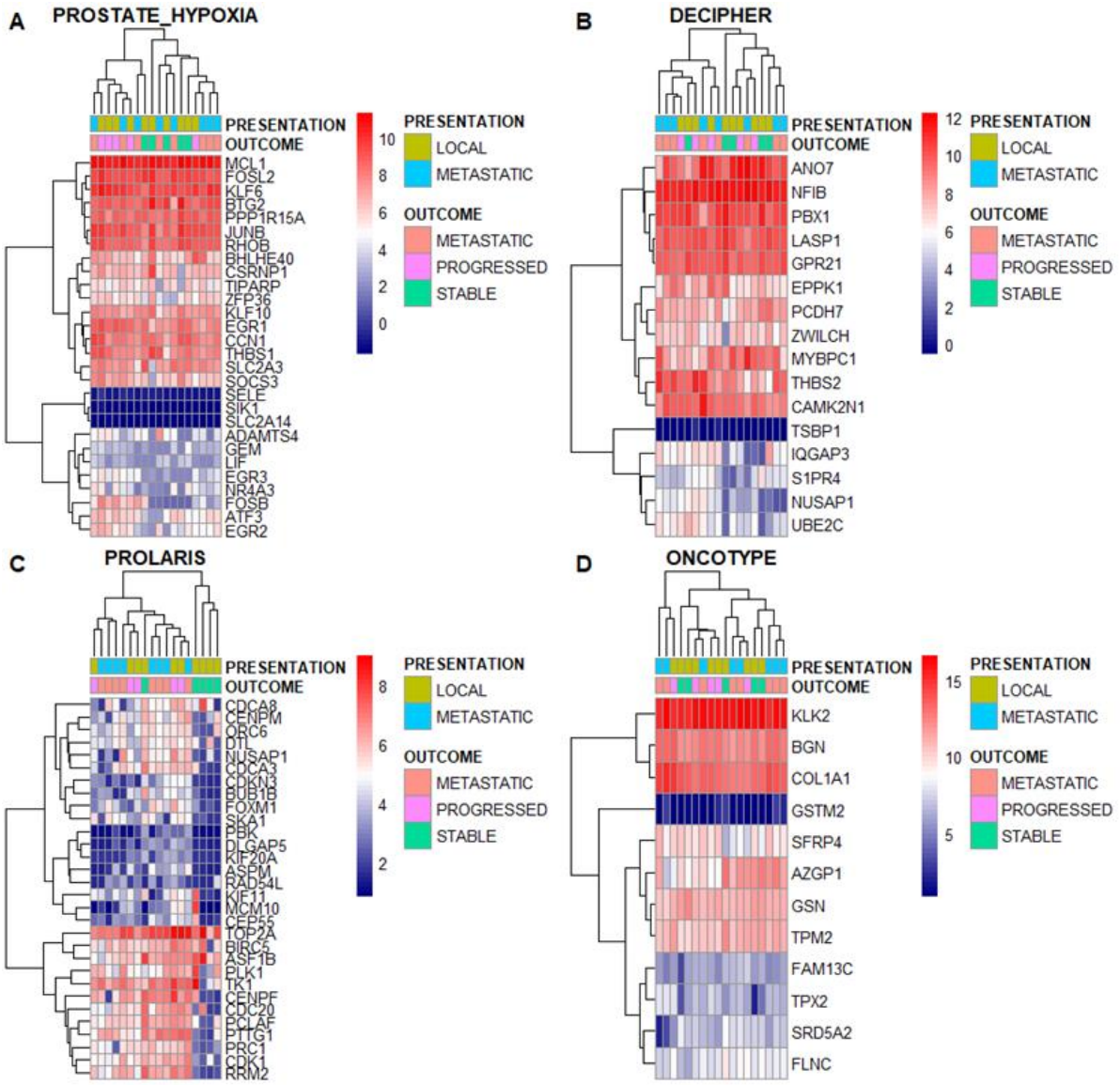
Appendix Table 3.1. Baseline clinicopathological characteristics of the cohort.

Group	T	N	M	Primary Gleason	Secondary Gleason	Gleason Sum	Grade Group	PSA (ng/ml)	Follow-up (years)	Risk Group/ Metastatic
Stable	T1c	N0	M0	3	3	6	1	7.2	8.17	Low
Stable	T2	N0	M0	4	5	9	5	8.7	6.75	High
Stable	T2c	N0	M0	3	4	7	2	1.2	6.75	High
Stable	T3b	N0	M0	4	3	7	3	11	8.17	High
Stable	T3b	N0	M0	3	4	7	2	80.3	8.25	High
Progressed	T2c	N0	M0	4	3	7	3	52	6.58	High
Progressed	T2a	N0	M0	5	4	9	5	4.7	3.75	High
Progressed	T2c	N0	M0	4	3	7	3	43.7	6.75	High
Progressed	T1c	N0	M0	4	3	7	3	2.5	4.33	Salvage RT
Progressed	T3b	N0	M0	4	3	7	3	6.4	3.75	High
Metastatic	T2c	N0	M1c	3	4	7	2	18.7	5.66	Metastatic
Metastatic	T4	N0	M1	4	5	9	5	52.4	5.08	Metastatic
Metastatic	T2c	N0	M1b	4	5	9	5	149.5	4.17	Metastatic
Metastatic	T3	N0	M1b	4	5	9	5	19.1	2.42	Metastatic
Metastatic	T4	N1	M1b	4	4	8	4	15856.6	2.33	Metastatic
Metastatic	T3b	N0	M1b	4	5	9	5	1524.6	1.42	Metastatic
Metastatic	T3b	N0	M1b	4	5	9	5	31.4	1.58	Metastatic
Metastatic	T3a	N1	M1b	4	4	8	4	513.4	1.0	Metastatic

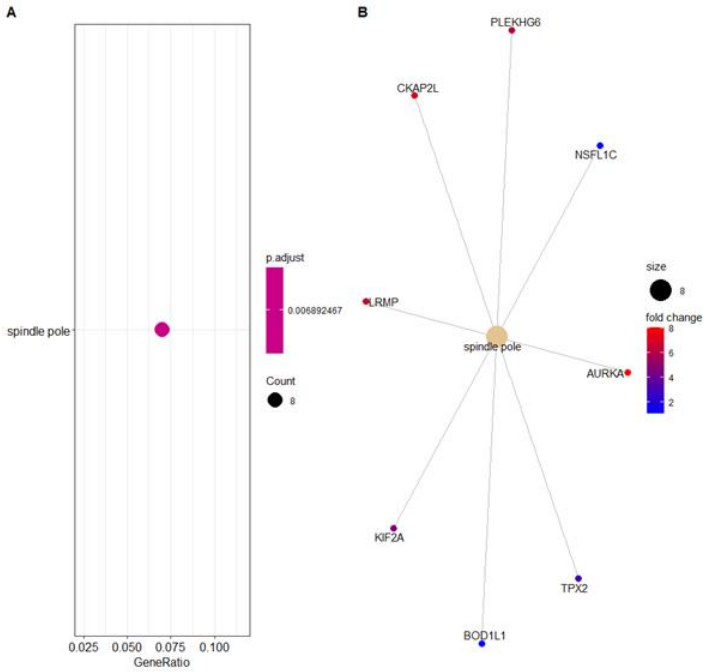
Appendix Figure 3.2. Intra-signature correlation generated by SigQC for previously validated signatures in prostate cancer (Prostate Hypoxia, Prolaris, Oncotype and Decipher) in the PROMPT dataset compared to TCGA Prostate PanCancer and Jain Radiotherapy datasets. Intra-Signature correlation is comparable in the PROMPT dataset compared to previously published datasets.



Appendix Figure 3.3. Signature expression heatmaps. Expression of previously validated prostate cancer signatures in PROMPT dataset; Prostate Hypoxia (A), Decipher (B), Prolaris (C), Oncotype (D).


A
PROSTATE_HYPOXIA
PRESENTATION
OUTCOME
MCL1
FOSL2
KLF6
BTG2
PPP1R15A
JUNB
RHOB
BHLHE40
CSRNP1
TIPARP
ZFP36
KLF10
EGR1
CCN1
THBS1
SLC2A3
SOCS3
SELE
SIK1
SLC2A14
ADAMTS4
GEM
LIF
EGR3
NR4A3
FOSB
ATF3
EGR2
10
8
6
4
2
0
PRESENTATION
LOCAL
METASTATIC
OUTCOME
METASTATIC
PROGRESSED
STABLE
B
DECIPHER
PRESENTATION
OUTCOME
ANO7
NFIB
PBX1
LASP1
GPR21
EPPK1
PCDH7
ZWILCH
MYBPC1
THBS2
CAMK2N1
TSBP1
IQGAP3
S1PR4
NUSAP1
UBE2C
12
10
8
6
4
2
0
PRESENTATION
LOCAL
METASTATIC
OUTCOME
METASTATIC
PROGRESSED
STABLE
C
PROLARIS
PRESENTATION
OUTCOME
CDCA8
CENPM
ORC6
DTL
NUSAP1
CDCA3
CDKN3
BUB1B
FOXM1
SKA1
PBK
DLGAP5
KIF20A
ASPM
RAD54L
KIF11
MCM10
CEP55
TOP2A
BIRC5
ASF1B
PLK1
TK1
CENPF
CDC20
PCLAF
PTTG1
PRC1
CDK1
RRM2
8
6
4
2
PRESENTATION
LOCAL
METASTATIC
OUTCOME
METASTATIC
PROGRESSED
STABLE
D
ONCOTYPE
PRESENTATION
OUTCOME
KLK2
BGN
COL1A1
GSTM2
SFRP4
AZGP1
GSN
TPM2
FAM13C
TPX2
SRD5A2
FLNC
15
10
5
PRESENTATION
LOCAL
METASTATIC
OUTCOME
METASTATIC
PROGRESSED
STABLE

Appendix Figure 3.4. Over-representation analysis of genes increased in progressed versus stable prostate cancer patients post-RRT (padj <0.01) (A). A network plot visualizing the enriched genes (B).



Appendix Figure 3.5. Over-representation analysis of genes increased in metastatic versus stable prostate cancer patients post-RRT (padj <0.05) (A). A network plot visualizing the enriched genes (B).

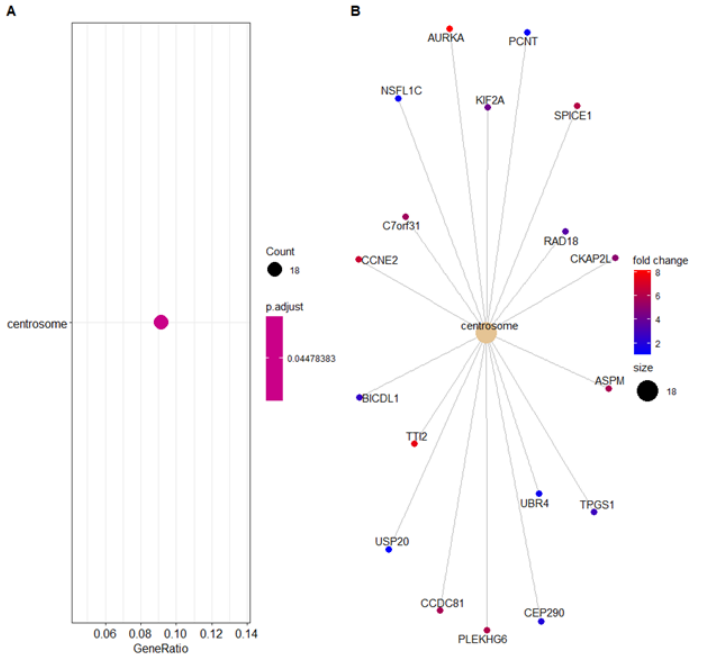
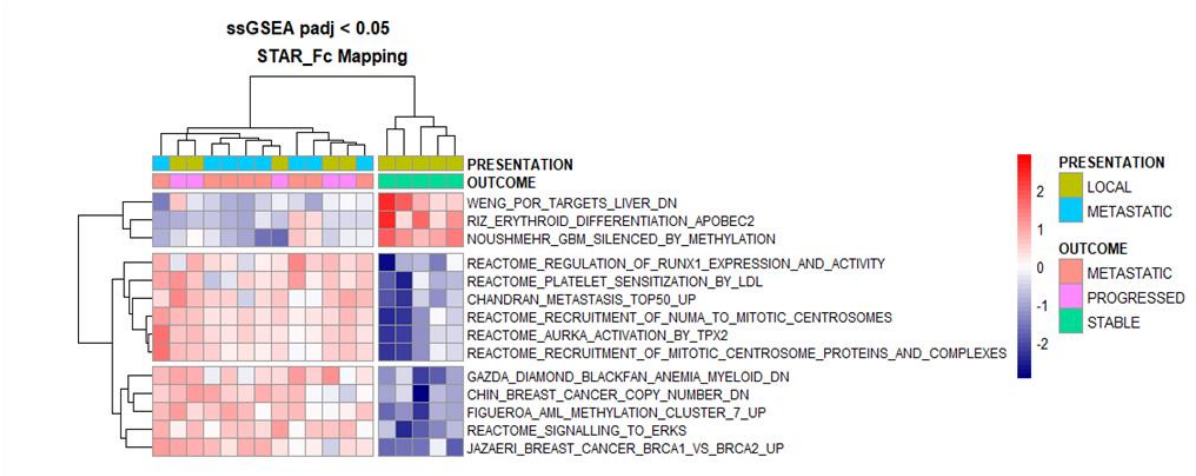


Table 3.6. Overlap of significantly differentially expressed genes between progressed and stable samples in both Quantseq (padj <0.2) and Nanostring (p<0.05) analysis. Genes increased in progressed versus stable cases in Quantseq and Nanostring include GNAS, ETV1, COL2A1. HDAC5 expression is decreased in progressed versus stable cases in Quantseq and Nanostring. LogFC - Log Fold Change; lfcSE - standard error; padj - adjusted p-value.

Quantseq													
	baseMean	logFC	lfcSE	stat	pvalue	padj	ENTREZ	SYMBOL					
GNAS	3424.709	0.657159	0.204451	3.214258	0.001308	0.039557	2778	GNAS					
HDAC5	566.0214	-1.28296	0.332002	-3.86431	0.000111	0.006872	10014	HDAC5					
ETV1	614.1381	2.874883	1.113311	2.582281	0.009815	0.142602	2115	ETV1					
COL2A1	505.9484	7.520908	1.50439	4.999306	5.75E-07	0.000125	1280	COL2A1					
NANOSTRING													
hgnc_sym	logFC	std.error..	Lower.cor	Upper.cor	Linear.fol	Lower.cor	Upper.cor	P.value	BY.p.value	method	Gene.sets	probe.ID	
GNAS	1.12	0.157	0.808	1.42	2.17	1.75	2.68	5.60E-05	0.0937	lm.nb	Driver Ge	NM_080425.1:1910	
HDAC5	-0.393	0.0998	-0.588	-0.197	0.762	0.665	0.872	0.00432	0.544	Wald	Chromatir	NM_005474.4:3160	
ETV1	3.6	1.1	1.45	5.75	12.1	2.74	53.8	0.00946	0.962	lm.nb	Transcript	NM_004956.4:1719	
COL2A1	2.24	1.38	-0.461	4.95	4.73	0.726	30.8	0.143	1	Wald	PI3K	NM_001844.4:4745	

Appendix Figure 3.7. Single sample GeneSet Enrichment Analysis (ssGSEA) pathways analysis of progressed & metastatic cases versus stable cases post-RRT. ssGSEA identifies multiple pathways from MSigDB C2: curated gene sets, significantly different between progressed and metastatic cases versus stable cases post-RRT. 3 Pathways are enriched in stable cases, 11 pathways are enriched in progressed and metastatic cases.



Appendix Figure 3.8 Overlap of genes in significant single sample GeneSet Enrichment Analysis (ssGSEA) pathways, comparing progressed and metastatic cases versus stable cases. Overlap is measured between 0 (0%) and 1 (100%).

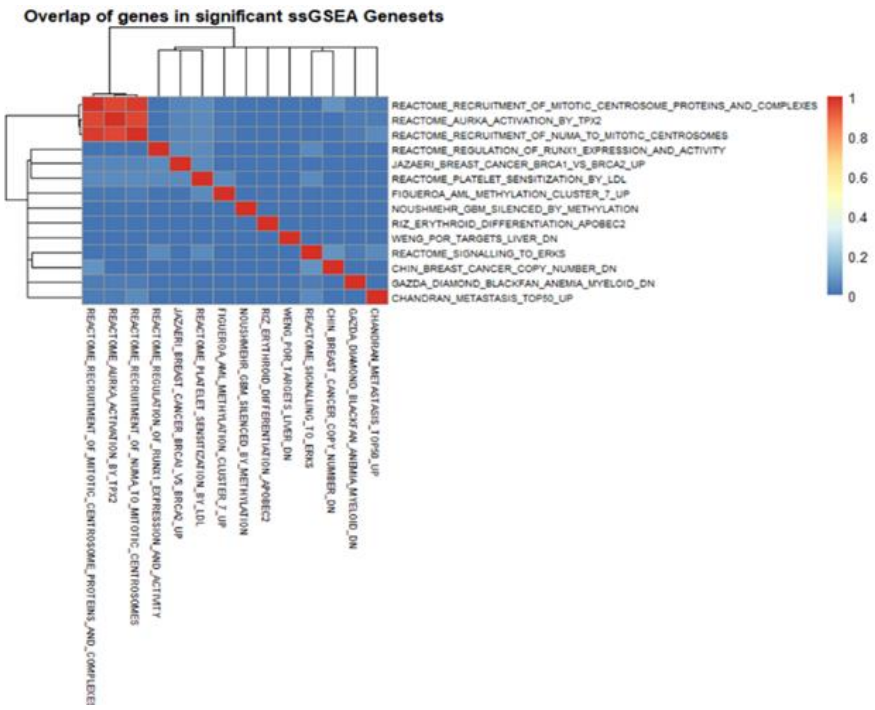
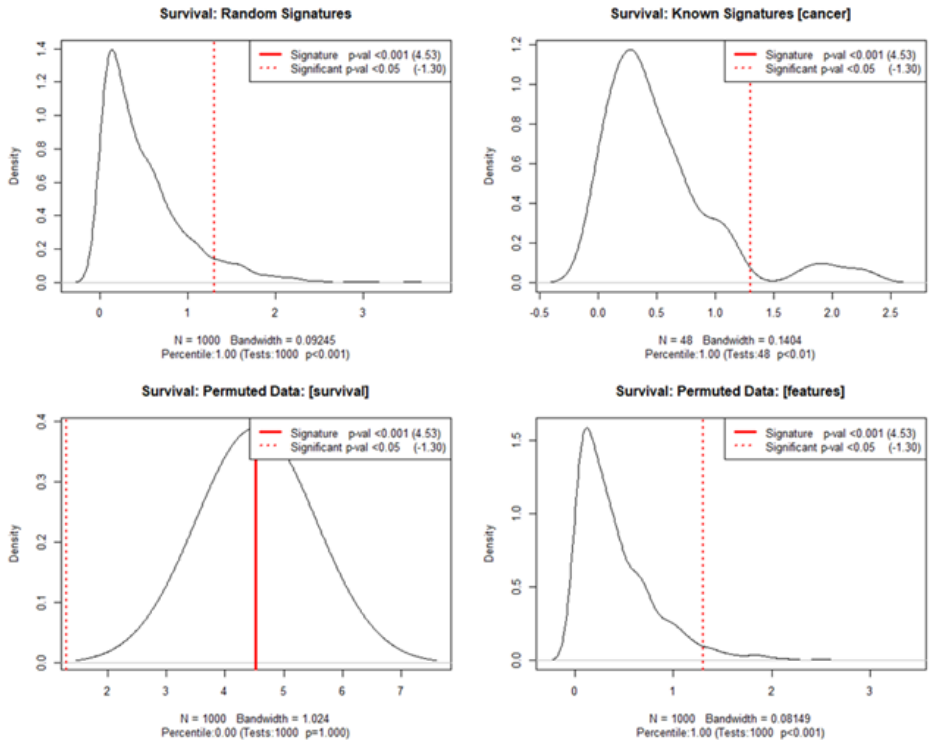


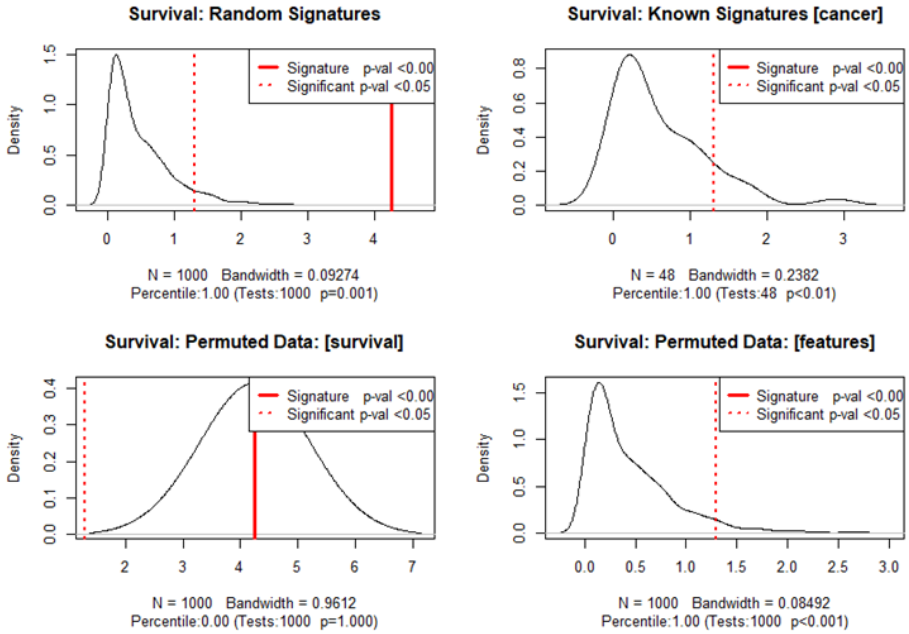
Table S3. Guided single sample GeneSet Enrichment Analysis (ssGSEA) of MSigDB C2: curated metastasis gene set signatures in the Jain dataset. LogFC - Log Fold Change; AveExpr - Average Expression; t = t statistic; P.Value - p-value; adj.P.Val - adjusted p-value; B - B statistic.

	logFC	AveExpr	t	P.Value	adj.P.Val	B
CHANDRAN_METASTASIS_TOP50_UP	0.0432	0.940626	5.081351	5.59E-05	0.003184	1.864197
BIDUS_METASTASIS_DN	0.042476	0.519153	3.188963	0.004582	0.082714	-2.3971
CHANDRAN_METASTASIS_UP	0.03397	0.634138	3.111228	0.00547	0.082714	-2.5649
TAVAZOIE_METASTASIS	-0.03039	0.024315	-3.08515	0.005804	0.082714	-2.62088
TOMIDA_METASTASIS_UP	0.033769	0.718665	2.921186	0.008403	0.095794	-2.96901
RAMASWAMY_METASTASIS_UP	0.036273	0.644774	2.750608	0.012278	0.099163	-3.32314
LIAO_METASTASIS	0.02087	0.497343	2.736128	0.012676	0.099163	-3.35278
PROVENZANI_METASTASIS_UP	0.022527	0.572919	2.596877	0.017182	0.099163	-3.63414
LEE_METASTASIS_AND_RNA_PROCESSING_UP	0.049333	0.680765	2.592328	0.017352	0.099163	-3.64321
PEDERSEN_METASTASIS_BY_ERBB2_ISOFORM_1	-0.05147	0.104123	-2.59114	0.017397	0.099163	-3.64558

Appendix Figure 3.8. SigCheck comparison of random signatures, cancer signatures, and survival and feature permutations for biochemical progression-free survival in the Jain dataset. The vertical red dotted line shows where a "significant" result (p=0.05) would lie relative to the background distribution.



Appendix Figure 3.9. SigCheck comparison of random signatures, cancer signatures, and survival and feature permutations for metastatic progression-free survival in the Jain dataset. The vertical red dotted line shows where a "significant" result (p=0.05) would lie relative to the background distribution.



Appendix Figure 3.10. SigCheck comparison of subset and all curated cancer signatures (MsigDB) compared to PROMPT signature for metastatic progression-free survival in the Jain dataset. The vertical red dotted line shows where a "significant" result (p=0.05) would lie relative to the background distribution.

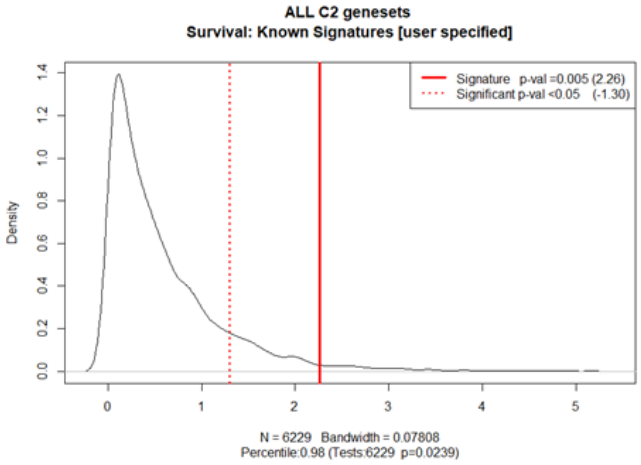
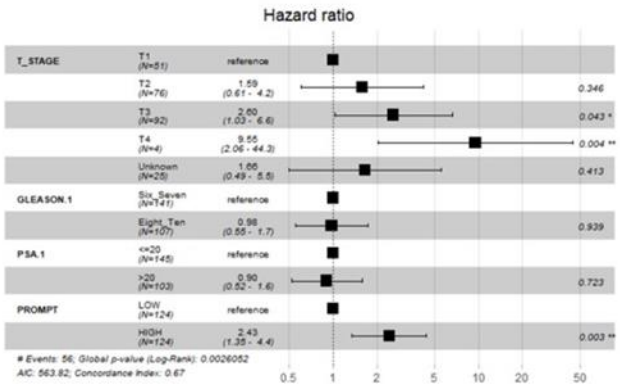


Table S4. Cox proportional hazard analysis of genes independently associated with biochemical progression-free survival.

Variable	Term	Beta	StandardE	Z	P	LRT	Wald	LogRank	HR	HRlower	HRupper	SYMBOL
ENSG00000164045	ENSG00000164045	0.956846	0.35338	2.707698	0.006775	0.009405	0.006775	0.006732	2.603472	1.302433	5.204156	CDC25A
ENSG00000173391	ENSG00000173391	0.406709	0.095872	4.242224	2.21E-05	8.78E-05	2.21E-05	1.34E-05	1.501867	1.244587	1.812332	OLR1
ENSG00000064309	ENSG00000064309	0.510918	0.198887	2.568882	0.010203	0.013811	0.010203	0.009752	1.666821	1.128745	2.461401	CDON
ENSG00000198563	ENSG00000198563	0.493169	0.246308	2.002248	0.045258	0.040567	0.045258	0.045329	1.637497	1.010469	2.653616	DDX39B
ENSG00000087586	ENSG00000087586	0.170078	0.095915	1.773217	0.076193	0.083299	0.076193	0.075295	1.185397	0.982248	1.430562	AURKA
ENSG00000131591	ENSG00000131591	0.685577	0.407456	1.68258	0.092456	0.10439	0.092456	0.093182	1.984917	0.893131	4.411329	C1orf159
ENSG00000055955	ENSG00000055955	0.56985	0.35549	1.603	0.108935	0.114623	0.108935	0.110542	1.768002	0.880824	3.548756	ITIH4
ENSG00000176125	ENSG00000176125	0.309713	0.234483	1.320831	0.186558	0.191315	0.186558	0.187351	1.363033	0.860823	2.158235	UFSP1
ENSG00000167261	ENSG00000167261	0.297207	0.234735	1.266138	0.205464	0.211412	0.205464	0.205149	1.346093	0.849706	2.132464	DPEP2
ENSG00000152133	ENSG00000152133	0.401069	0.300866	1.333049	0.182516	0.182156	0.182516	0.183227	1.493421	0.828103	2.69327	GPATCH11
ENSG00000180938	ENSG00000180938	0.367678	0.293749	1.251673	0.210689	0.223501	0.210689	0.210475	1.444376	0.812158	2.568739	ZNF572
ENSG00000070882	ENSG00000070882	0.085181	0.150925	0.56439	0.572489	0.573392	0.572489	0.572859	1.088914	0.810076	1.463731	OSBPL3
ENSG00000203668	ENSG00000203668	0.217485	0.228721	0.950874	0.341668	0.346582	0.341668	0.341392	1.242947	0.793898	1.945989	CHML
ENSG00000175445	ENSG00000175445	0.164628	0.212462	0.774859	0.438423	0.449456	0.438423	0.438464	1.178955	0.777408	1.787908	LPL
ENSG00000196605	ENSG00000196605	0.367586	0.325662	1.128734	0.25901	0.268968	0.25901	0.259827	1.444244	0.762845	2.734289	ZNF846
ENSG00000163507	ENSG00000163507	0.31309	0.300082	1.04335	0.296786	0.312492	0.296786	0.296327	1.367645	0.759527	2.462653	CIP2A
ENSG00000135114	ENSG00000135114	0.422475	0.362872	1.164254	0.244321	0.254959	0.244321	0.243761	1.525733	0.749206	3.107104	OASL
ENSG00000143333	ENSG00000143333	0.433914	0.371451	1.16816	0.242742	0.250308	0.242742	0.242671	1.543287	0.74519	3.196145	RGS16
ENSG00000197008	ENSG00000197008	0.059425	0.180504	0.329215	0.741993	0.744606	0.741993	0.741966	1.061226	0.74501	1.511658	ZNF138
ENSG00000198270	ENSG00000198270	0.098568	0.200504	0.491604	0.622999	0.629875	0.622999	0.623023	1.10359	0.74497	1.634846	TMEM116
ENSG00000164038	ENSG00000164038	0.492776	0.402184	1.225251	0.220481	0.231115	0.220481	0.220165	1.636854	0.744167	3.600387	SLC9B2
ENSG00000095485	ENSG00000095485	-0.05875	0.124751	-0.47092	0.637699	0.642893	0.637699	0.637582	0.942945	0.738412	1.204132	CWF19L1
ENSG00000219481	ENSG00000219481	0.129861	0.239663	0.541851	0.587921	0.588213	0.587921	0.587932	1.138671	0.711863	1.821376	NBPF1
ENSG00000136628	ENSG00000136628	0.051341	0.201431	0.254881	0.798815	0.797406	0.798815	0.79882	1.052682	0.709315	1.562266	EPRS1
ENSG00000028310	ENSG00000028310	0.424212	0.413123	1.026842	0.304495	0.305833	0.304495	0.30425	1.528385	0.680114	3.43466	BRD9
ENSG00000112137	ENSG00000112137	0.177612	0.312372	0.568593	0.569633	0.579817	0.569633	0.570043	1.194362	0.647507	2.203067	PHACTR1
ENSG00000121766	ENSG00000121766	0.11503	0.302857	0.379817	0.704081	0.704329	0.704081	0.704095	1.121907	0.619676	2.031183	ZCCHC17
ENSG00000197102	ENSG00000197102	0.331047	0.419195	0.789719	0.429692	0.424109	0.429692	0.429623	1.392425	0.612283	3.166588	DYNC1H1
ENSG00000054148	ENSG00000054148	0.066212	0.291141	0.227422	0.820096	0.820362	0.820096	0.820104	1.068453	0.603859	1.890493	PHPT1
ENSG00000172247	ENSG00000172247	0.151852	0.371989	0.408217	0.683114	0.684486	0.683114	0.683162	1.163988	0.56145	2.41316	C1QTNF4
ENSG00000129696	ENSG00000129696	-0.16879	0.223884	-0.75392	0.450897	0.454237	0.450897	0.450783	0.844686	0.544659	1.309983	TTI2
ENSG00000033122	ENSG00000033122	-0.0351	0.308316	-0.11386	0.909349	0.909014	0.909349	0.909349	0.965504	0.527612	1.766825	LRRC7
ENSG00000168062	ENSG00000168062	0.177936	0.441912	0.402651	0.687205	0.689084	0.687205	0.687219	1.194749	0.502482	2.840751	BATF2
ENSG00000198420	ENSG00000198420	-0.44449	0.232892	-1.90858	0.056316	0.06232	0.056316	0.056696	0.641149	0.406182	1.012038	TCAF1
ENSG00000105426	ENSG00000105426	-0.01774	0.512863	-0.03458	0.972414	0.972402	0.972414	0.972414	0.982421	0.359541	2.684402	PTPRS
ENSG00000114853	ENSG00000114853	-0.20754	0.518003	-0.40066	0.688668	0.688651	0.688668	0.688698	0.812577	0.294401	2.242792	ZBTB47

Appendix Figure 3.11. 4-gene (mean per gene - median cut-off for cohort) cox proportional hazards model analysis of biochemical progression-free survival. Forest plot of the cox proportional hazard model for biochemical progression-free survival, incorporating clinicopathological features (T-stage, Gleason score and PSA) using clinically relevant cut-points in the Jain radiotherapy cohort (A). Cox proportional hazard model incorporating clinicopathological features (T-stage, Gleason score and PSA) using clinically relevant cut-points in the Jain radiotherapy cohort (B). Univariate analysis of 4 overlapping genes (COL2A1, GNAS, ETV1 and HDAC5) between Quantseq and nanoString expression analysis identifies ETV1 and COL2A1 as having a lower HR value >1 in the Jain Radiotherapy Cohort.



Clinically relevant thresholds

GLEASON.1 dichotomised to <= 20/ >20

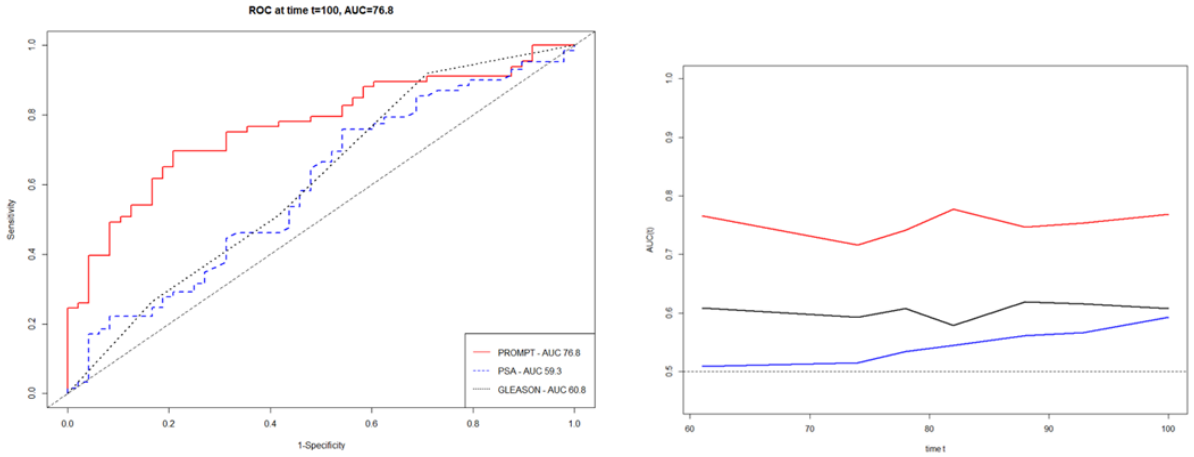
PSA.1 dichotomised to Six_Seven/ Eight_Ten

```
 Cox model: response is Surv(time, time_status)
Terms added sequentially (first to last)

           loglik   Chisq Df Pr(>|Chi|)
NULL      -285.88
T_STAGE   -279.72 12.3273  4   0.015077 *
GLEASON.1 -279.72  0.0027  1   0.958424
PSA       -279.22  0.9985  1   0.317680
PROMPT    -274.39  9.6450  1   0.001899 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Analysis of Deviance Table
 Cox model: response is Surv(time, time_status)
Terms added sequentially (first to last)

           loglik   Chisq Df Pr(>|Chi|)
NULL      -285.88
T_STAGE   -279.72 12.3273  4   0.015077 *
GLEASON.1 -279.72  0.0027  1   0.958424
PSA.1     -279.69  0.0413  1   0.839017
PROMPT    -274.91  9.5653  1   0.001983 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

PER GENE UNIVARIATE

Term	Beta	StandardE	Z	P	LRT	Wald	LogRank	HR	HRlower	HRupper	SYMBOL
ENSG00000139219	0.190489	0.088815	2.144795	0.031969	0.048235	0.031969	0.030404	1.209842	1.016551	1.439885	COL2A1
ENSG00000108840	0.298047	0.500966	0.594945	0.55188	0.555108	0.55188	0.551934	1.347225	0.504681	3.596361	HDAC5
ENSG00000087460	0.118691	0.319839	0.371096	0.710566	0.709384	0.710566	0.710664	1.126022	0.601588	2.107632	GNAS
ENSG00000006468	0.363658	0.076723	4.739879	2.14E-06	6.31E-05	2.14E-06	2.26E-07	1.438582	1.237736	1.67202	ETV1

Appendix Figure 3.12. 4-gene receiver operator curve for biochemical progression-free survival in the Jain radiotherapy dataset.



Appendix Figure 3.13. 4-gene receiver operator curve for metastatic progression-free survival in the Jain radiotherapy dataset.

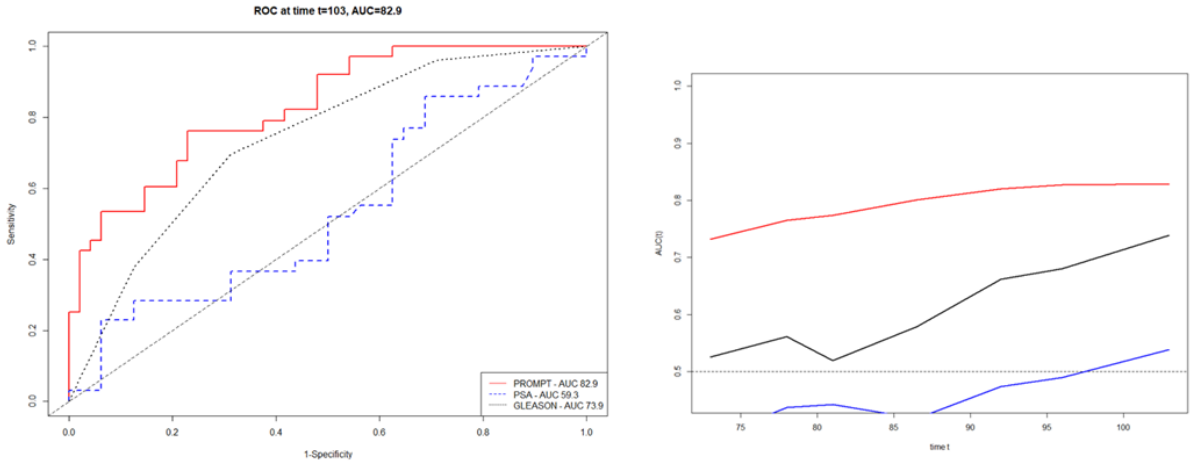
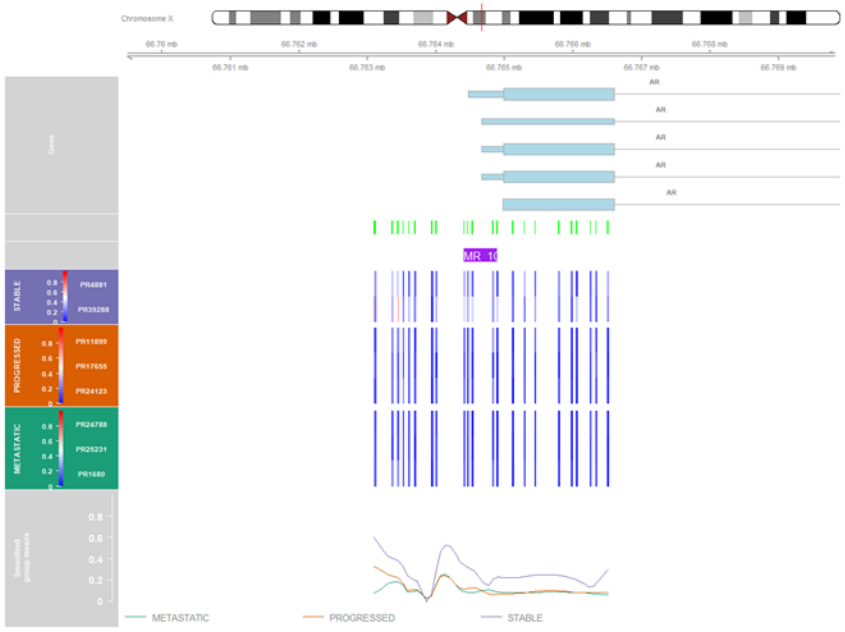


Table S5. 4 gene (mean per gene - median cut-off for cohort) cox proportional hazards and clinical variables.

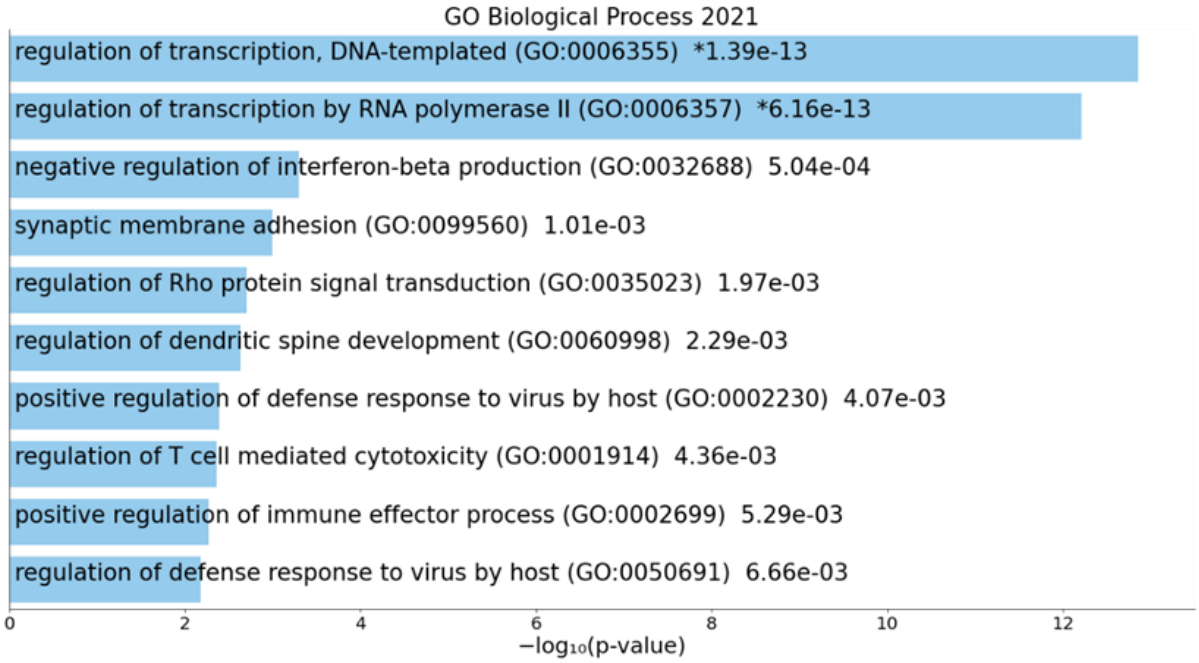
Variable	Term	Beta	StandardE	Z	P	LRT	Wald	LogRank	HR	HRlower	HRupper
DDX39B	T_STAGET2	0.370335	0.489449	0.756636	0.449268	0.007243	0.003301	0.000695	1.44822	0.554899	3.779676
DDX39B	T_STAGET3	1.081952	0.475729	2.274304	0.022948	0.007243	0.003301	0.000695	2.950432	1.161301	7.495949
DDX39B	T_STAGET4	2.645764	0.780381	3.390349	0.000698	0.007243	0.003301	0.000695	14.09421	3.053368	65.05827
DDX39B	T_STAGEUnknown	0.575797	0.611459	0.941678	0.346358	0.007243	0.003301	0.000695	1.778547	0.536526	5.89576
DDX39B	GLEASON.1Eight_Ten	0.027696	0.295617	0.09369	0.925356	0.007243	0.003301	0.000695	1.028083	0.575968	1.835093
DDX39B	PSA.1>20	-0.03394	0.285025	-0.11906	0.905225	0.007243	0.003301	0.000695	0.966633	0.552902	1.689955
DDX39B	DDX39B	0.388273	0.151545	2.562094	0.010404	0.007243	0.003301	0.000695	1.474432	1.095543	1.984359
CDC25A	T_STAGET2	0.334027	0.492335	0.678455	0.497483	0.009271	0.002886	0.000539	1.396581	0.532096	3.665581
CDC25A	T_STAGET3	0.872398	0.477962	1.825247	0.067964	0.009271	0.002886	0.000539	2.392642	0.93764	6.105474
CDC25A	T_STAGET4	2.401922	0.781721	3.07261	0.002122	0.009271	0.002886	0.000539	11.04439	2.386381	51.11441
CDC25A	T_STAGEUnknown	0.423098	0.626528	0.675306	0.499481	0.009271	0.002886	0.000539	1.526684	0.447144	5.212555
CDC25A	GLEASON.1Eight_Ten	0.102228	0.298354	0.34264	0.73187	0.009271	0.002886	0.000539	1.107636	0.617217	1.987725
CDC25A	PSA.1>20	-0.15416	0.287861	-0.53553	0.592286	0.009271	0.002886	0.000539	0.857138	0.487554	1.506877
CDC25A	CDC25A	0.336819	0.129711	2.596679	0.009413	0.009271	0.002886	0.000539	1.400486	1.086096	1.805882
OLR1	T_STAGET2	0.391187	0.491818	0.79539	0.426386	0.001669	0.000394	4.41E-05	1.478736	0.563968	3.877278
OLR1	T_STAGET3	0.883408	0.482636	1.830382	0.067193	0.001669	0.000394	4.41E-05	2.41913	0.939375	6.229879
OLR1	T_STAGET4	2.122276	0.794472	2.671304	0.007556	0.001669	0.000394	4.41E-05	8.350118	1.759694	39.62307
OLR1	T_STAGEUnknown	0.436512	0.624361	0.699135	0.484468	0.001669	0.000394	4.41E-05	1.547301	0.455112	5.260559
OLR1	GLEASON.1Eight_Ten	-0.29488	0.312719	-0.94294	0.345709	0.001669	0.000394	4.41E-05	0.744623	0.403413	1.374433
OLR1	PSA.1>20	-0.1739	0.292538	-0.59445	0.552213	0.001669	0.000394	4.41E-05	0.840382	0.473661	1.491027
OLR1	OLR1	0.443226	0.129258	3.429013	0.000606	0.001669	0.000394	4.41E-05	1.557724	1.209111	2.006849
CDON	T_STAGET2	0.343375	0.491894	0.698066	0.485136	0.032618	0.010957	0.002276	1.409697	0.537557	3.696809
CDON	T_STAGET3	0.827325	0.48198	1.716514	0.086068	0.032618	0.010957	0.002276	2.287192	0.889284	5.882534
CDON	T_STAGET4	2.358766	0.786396	2.999464	0.002705	0.032618	0.010957	0.002276	10.57789	2.264737	49.40606
CDON	T_STAGEUnknown	0.497061	0.617809	0.804554	0.421077	0.032618	0.010957	0.002276	1.643883	0.489768	5.517608
CDON	GLEASON.1Eight_Ten	-0.07263	0.295465	-0.24581	0.805826	0.032618	0.010957	0.002276	0.929945	0.521143	1.659427
CDON	PSA.1>20	-0.08058	0.285589	-0.28214	0.777836	0.032618	0.010957	0.002276	0.922585	0.527124	1.614731
CDON	CDON	0.219884	0.125762	1.74841	0.080393	0.032618	0.010957	0.002276	1.245932	0.973745	1.594201

Appendix Figure 3.15. Differentially Methylated Region plot for the Androgen Receptor gene, visualizing differential methylation between progressed and metastatic cases versus stable cases post-RRT. Genomic co-ordinates and proximal coding regions (top), and heatmap and mean methylation plots (bottom), are illustrated.

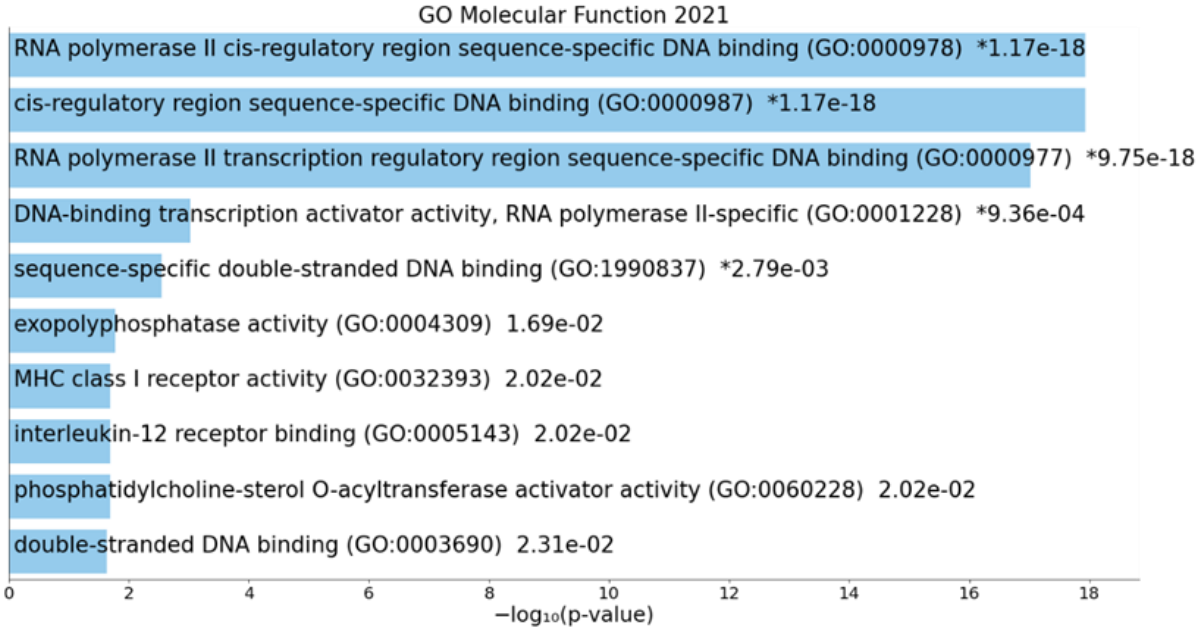


DMR	seqnames	start	end	width	strand	no.cpgs	min_smoothed_fdr	Stouffer	HMFDR	Fisher	maxdiff	meandiff	Gene
106	chrX	66764411	66764893	483	*	6	1.90E-35	0.008741	0.091384	0.021871	-0.24634	-0.16073	AR

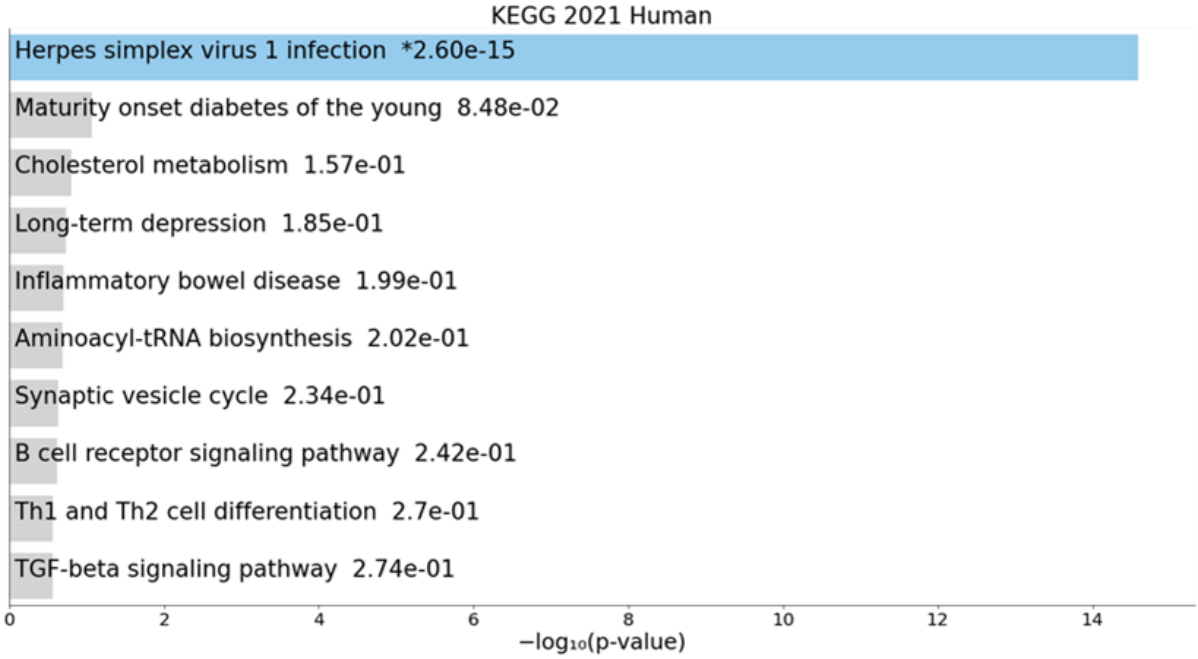
Appendix Figure 3.16. Bar-plot of over-representation analysis of Chromosome 19 (enrichr) GO Biological Process pathways, and Differentially Methylated Genes, between progressed and metastatic cases versus stable cases.



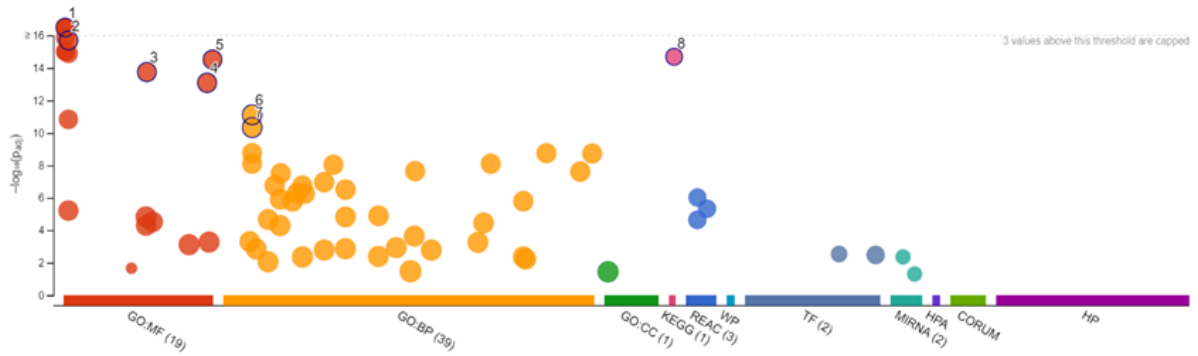
Appendix Figure 3.17. Bar-plot of over-representation analysis of Chromosome 19 (enrichr) GO Molecular Function pathways, and Differentially Methylated Genes, between progressed and metastatic cases versus stable cases.



Appendix Figure 3.18. Bar-plot of over-representation analysis of Chromosome 19 (enrichr) KEGG pathways, and Differentially Methylated Genes, between progressed and metastatic cases versus stable cases.



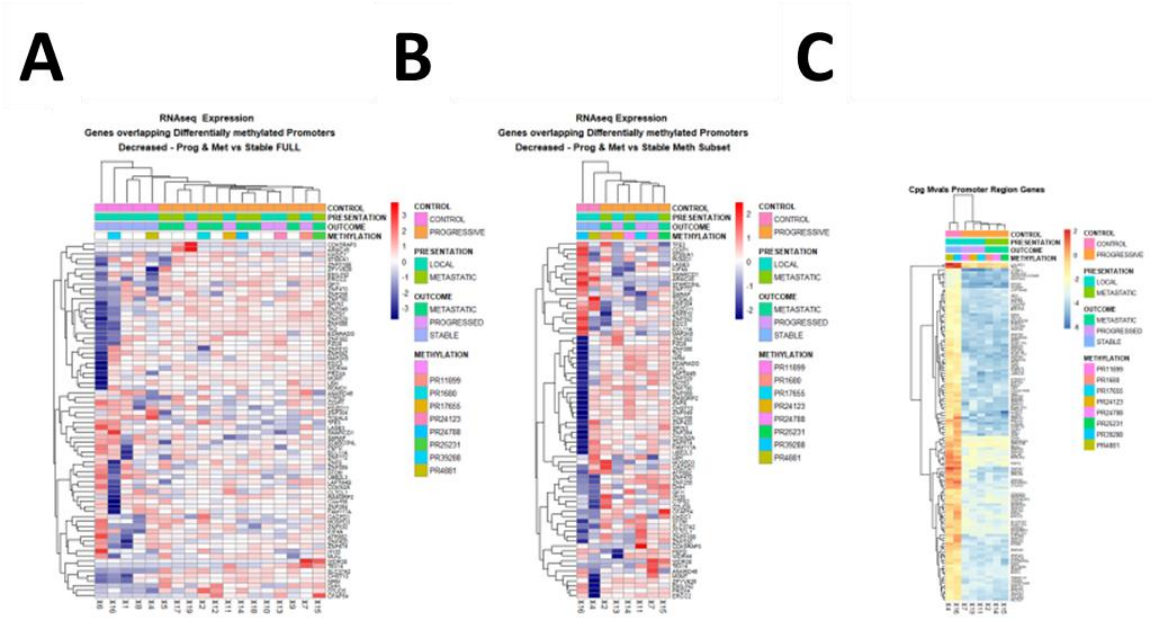
Appendix Figure 3.19. Over-representation analysis of all pathways for Chromosome 19 (gprofiler).



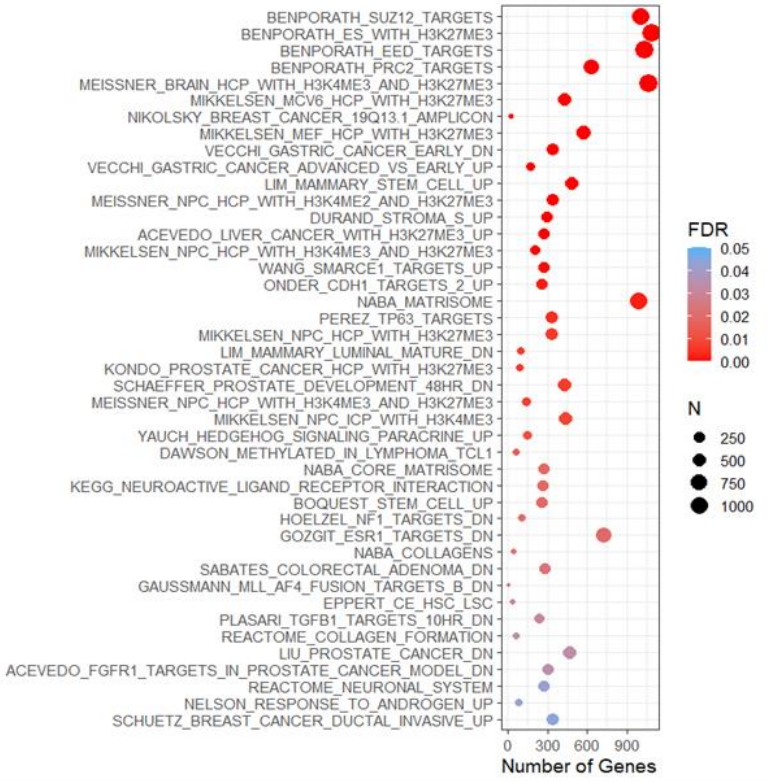
ID	Source	Term ID	Term Name	p_{adj} (query_1)
1	GO:MF	GO:0000978	RNA polymerase II cis-regulatory region sequence...	1.922×10^{-17}
2	GO:MF	GO:0003700	DNA-binding transcription factor activity	2.019×10^{-16}
3	GO:MF	GO:0043565	sequence-specific DNA binding	1.791×10^{-14}
4	GO:MF	GO:0140110	transcription regulator activity	8.155×10^{-14}
5	GO:MF	GO:1990837	sequence-specific double-stranded DNA binding	3.035×10^{-15}
6	GO:BP	GO:0006357	regulation of transcription by RNA polymerase II	7.650×10^{-12}
7	GO:BP	GO:0006366	transcription by RNA polymerase II	4.675×10^{-11}
8	KEGG	KEGG:05168	Herpes simplex virus 1 infection	2.017×10^{-15}

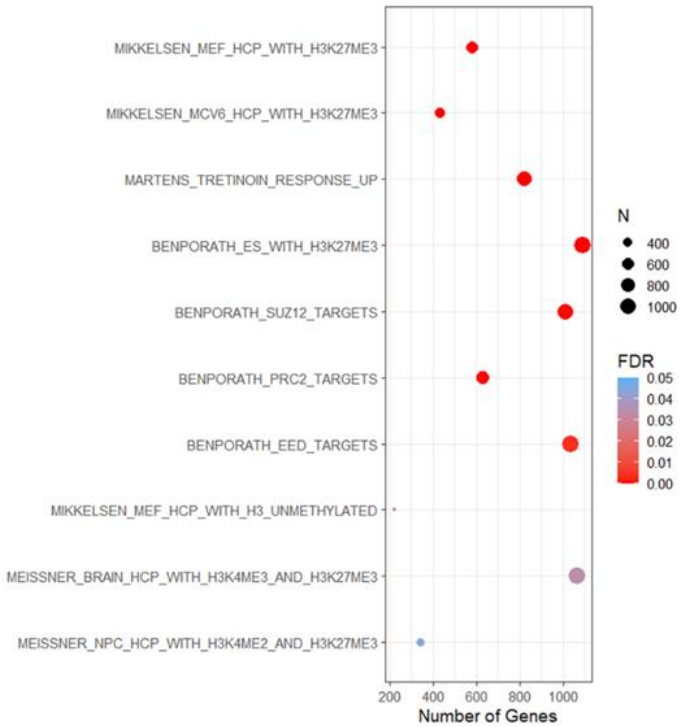
Appendix Figure 3.20. Heatmap analysis of RNA expression of differentially hypo-methylated genes at promoter in the full cohort, methylation subset, and cpg Mvals for methylation. Quantseq RNA expression analysis heatmap of genes with differentially hypo-methylated promoter regions between progressed and metastatic cases versus stable cases in the whole cohort (A). Quantseq RNA expression analysis heatmap of genes with differentially hypo-methylated promoter regions between progressed and metastatic cases versus stable cases in the methylation analysis cohort (B). Methylation (Mvals) heatmap of Cpgs in promoter regions differentially methylated between progressed and metastatic cases versus stable cases (C).

.

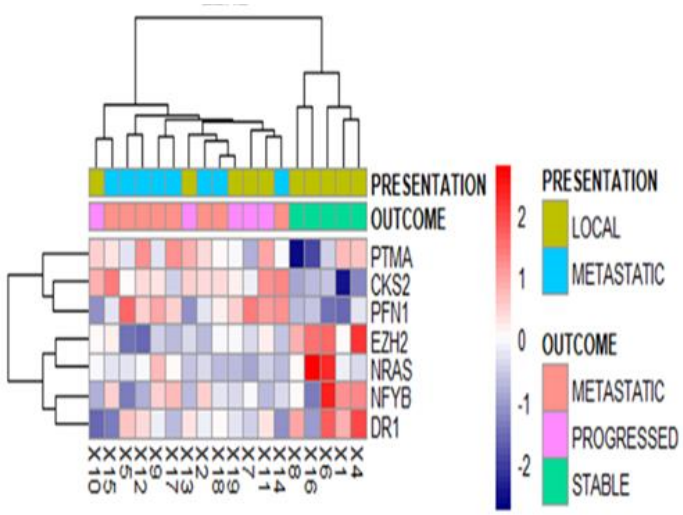


Appendix Figure 3.21. Dot-plot over-representation analysis of hypomethylated (A) and hypermethylated (B) C2 Curated Pathway (MSigDB) genes in progressed and metastatic cases versus stable cases post-RRT.





Appendix Figure S22. Heatmap of EZH2 expression, and expression of EZH2 target genes, from the Quantseq data, demonstrating significant differential expression between progressed and metastatic cases versus stable cases (padj<0.05).



Appendix Figure S23. Signature Genelist overlap.