Investigating disease and radiotherapy response associations with rectal tumour expression of the DNA damage response proteins, ATM, MRE11, NBS1 and RAD50

By
Vincent Ho MBBS, FRACP

Submitted in fulfilment of the requirements for the degree of

Doctor of Philosophy

Discipline of Pathology

School of Medicine

Western Sydney University

September 2018

Doctoral Supervisors: Professor Soon Lee and Professor Jim Yong

Table of Contents

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

Rectal cancers are thought to contribute approximately one third of all colorectal cancers worldwide, and are associated with considerable morbidity and mortality. Worryingly, the incidence of rectal cancer is increasing in developed economies, as genetics and environment converge to cause pathology in increasingly older populations.

Despite current detailed knowledge of various molecular mechanisms responsible for oncogenesis, in general, the precise sequence of events causing disease, or influencing prognosis, in a particular patient is not completely understood. This is unsurprising given the multifactorial nature of disease and treatment responses in human populations with highly variable clinical histories.

To overcome this knowledge gap, this thesis sought to further refine the understanding of the molecular mechanisms at work during rectal cancer development, and their effect on treatment responses and patient outcomes. Furthermore, by defining the molecular mechanisms of disease, biomarkers (single or multiple molecules whose expression levels serve to identify disease processes common to many patients with the same disease) can be developed and applied to clinical situations – helping to diagnose, prognosticate and determine treatment modalities, depending on the application in question.

To this end, and given the large heritage literature concerning DNA damage response proteins and cancer pathophysiology, the expression of four central DNA repair proteins (ATM, MRE11, NBS1 and RAD50) in rectal cancers has been quantified by immunohistochemistry. This will enable correlation of expression levels in different regions of the tumour with available clinicohistopathological variables, such as overall and disease-free survival.

Furthermore, although radiotherapy represents a first-line treatment for rectal cancer, highly variable treatment responses have been documented amongst patients. As not all patients stand to benefit from such treatments, the expression of the candidate proteins – central to repairing damaged DNA generated by radiotherapy – and the association with radiotherapy responses are investigated in rectal tumours.

Firstly, in the case of ATM, it was found that reduced expression in the growing edge of the tumour (tumour periphery) was associated with better responses to radiotherapy and improved disease-free survival. ATM expression in the tumour centre was also associated with disease-free survival by uni- and multi-variate analyses. Secondly, MRE11 expression was found to be predictive of patient outcomes, when patients were also scored positive for high-grade disease, metastasis positive, and showed perineural invasion. In contrast, NBS1 expression levels in rectal tumours were only found to have a marginal association with patient overall survival, necessitating additional studies of NBS1 in rectal cancer. Low RAD50 expression was associated with worse disease-free survival. RAD50 levels also proved to be useful to delineate low- and high-grade disease subgroups. Together, the analysis of these four markers individually, led to several novel associations with regards to rectal cancer – highlighting their 'biomarker' potential in this clinical context.

Furthermore, by combining expression of these proteins into combinatorial panels – made up of either ATM and MRE11, or MRE11, NBS1 and RAD50 – a greater predictive power of

their expression levels with respect to patient outcomes was demonstrated, and support the use of multiple markers to better understand disease in different patient groups.
Therefore, the utility of examining DDR proteins in the context of rectal cancer is demonstrated in this thesis, and the results provide evidence to support future studies investigating the roles of these proteins in larger rectal cancer patient cohorts and other cancers. Further studies and validation of the results of this thesis will help determine whether such proteins can serve as clinically-useful biomarkers for disease intervention.

STATEMENT OF ORIGINALITY

The work presented in this thesis was carried out in the Department of Anatomical Pathology at Liverpool Hospital, Sydney and Ingham Institute for Applied Medical Research. The material presented in this thesis is, to the best of my knowledge and belief, original except as acknowledged in the text. I hereby declare that I have not submitted this material, either in full or in part, for a degree at this or any other institution.

In agreement with the Western Sydney University policy on doctoral theses, the word count of the document extends to 83,203 words, and will be deposited open-access in the University's central library.

Student ID: 17546111

ACKNOWLEDGEMENTS

This thesis represents the contributions of many teachers, colleagues and friends over six years, and would not have been at all possible without their continued support, guidance and encouragement.

This long process certainly would not have been possible without the support, encouragement and advice of many persons whom I have had the great fortune to meet during my candidature. First of all, I would like to express sincere gratitude to Professor Cheok Soon Lee, my principal supervisor who has enabled me to train as an independent researcher under his supervision. I have found Professor Lee to be that extraordinary combination of clinician, academic and philosopher which is so inspiring and intensely motivating as I would like to follow in his footsteps someday. Professor Lee has not only provided the funding and laboratory to enable the research to be carried out but has also directly expressed his interest in my work despite his many competing interests, for which I am deeply grateful. Additionally, I would like to thank Professor Lee for the opportunity to be corresponding author on the 3 manuscripts that I have submitted (2 published and 1 under review). This opportunity has allowed me to engage directly on comments back from journal reviewers and gain good grasp of the publication review process, which will serve me well in the future.

I am most indebted to the wonderful support of Dr Liping Chung postdoctoral scientist in the Pathology Laboratory at the Ingham Institute (and now Clinical Research Coordinator) for the generous disposal of her time, expertise and resources which facilitated a great large portion of this thesis to be realised. Her knowledge of the basic science environment and the support that she provides budding new researchers is legendary and along with technical officer Mr Askar Abubakar have helped bring new invaluable insights into immunofluorescence work.

My deep thanks must go to Dr Joo Shin who helped to provide direct supervision of my work earlier on in my candidature. Dr Shin has an abundance of patience, support and good will and this has helped to give me the confidence and motivation to realise this doctoral research work. Professor Jim Yong who has kindly helped to supervise me during the latter part of my candidature I must thank for his gracious support and time.

My heartfelt thanks go to Dr Thein Ga Tut for lending an ear to my initial inquiriesand I greatly appreciate his advice on research direction from the time that I spent with him during his doctoral candidature in the Pathology Laboratory.

Dr Vivienne Lea and Dr Amandeep Singh as pathology registrars working in the laboratory I want to personally thank for their efforts with scoring specimens. For someone who is very interested in the clinical relevance of gastrointestinal pathology I have come to appreciate the often time consuming and exact way in which diagnostic histopathology is done.

Accessing the clinical information would only be possible with the generous support and time of medical oncologists Dr Stephanie Lim, Dr Weng Ng, Professor Paul de Souza, Dr Wei Chua and radiation oncologist Dr Mark Lee. All of them have not only been supportive of access to the Cancer Registry of South Western Sydney Local Health District (SWSLHD) but have gone out of their way to find any missing details of patient clinical outcomes.

Thanks go to Professor Annemarie Hennessy Dean of the Medical School and Dr David Mahns Higher Degree Coordinator for their support over the 6 years that I have worked on this research as a PhD candidate. It has been very challenging at many times to do this work whilst

balancing a clinical job but they have always been supportive and provided helpful advice and direction.
Finally none of this work would not have been at all possible were it were not for the very kind and loving support of my dear wife Cindy Luong. She has always provided a kind and sympathetic ear during the most challenging times of the candidature. And those challenging times with our very young and dearly loved children Olivia and Brandon have led to clarity of purpose and the strengthening of my resolve to complete this candidature.

PUBLICATIONS ARISING FROM THIS RESEARCH

1) Revoltar, M., Shin, J., Lim, S., Tut, T., Dissanayake, I., Descallar, J., **Ho, V**., Chua, W., Ng, W., Lee, M., et al. (2016). Early marker of DNA damage response, atm, as a predictor of clinical outcome following radiotherapy in rectal cancer patients. Pathology.*48*, s153.

2) Shin JS, Tut TG, **Ho V** and Lee CS. (2014). Predictive markers of radiotherapy-induced rectal cancer regression. J ClinPathol. *67*, 859–864.

3) **Ho, V**., Chung, L., Revoltar, M., Lim, S.H., Tut, T.G., Abubakar, A., Henderson, C.J., Chua, W., Ng, W., Lee, M., et al. (2016). MRE11 and ATM expression levels predict rectal cancer survival and their association with radiotherapy response. PLoS One. *11*, e0167675.

4) **Ho, V**., Chung, L., Singh, A., Lea, V., Revoltar, M., Lim, SH., Tut, TG., Ng, W., Lee, M., de Souza, P., Shin, J and Lee, CS. (2017). Early postoperative low expression of RAD50 in rectal cancer patients associates with disease-free survival. Cancers. *9*, 163.

5) **Ho, V**., Chung, L., Singh, A., Lea, V., Abubakar, A., Lim, SH., Ng, W., Lee, M., de Souza, P., Shin. J and Lee, CS. (2018). Overexpression of the MRE11-RAD50-NBS1 (MRN) complex in rectal cancer correlates with poor response to neoadjuvant radiotherapy and prognosis. BMC Cancer. *18*, 869.

ABBREVIATIONS

ACP	Australian clinico-pathological staging
AIHW	Australian Institute of Health and Welfare
AJCC	American Joint Committee on Cancer
APC	adenomatous polyposis coli
APR	abdomino-perineal resection
ATCC	American Tissue Culture Collection
ATM	ataxia telangiectasia mutated (also known as TEL1)
ATP	adenosine triphosphate
ATR	ataxia telangiectasia and RAD-3 related
AUC	area under curve
B/F/B	bridge/fusion/breakage
BRAF	B-Raf protooncogene
BRCA1	breast cancer susceptibility type 1
BRCT	BRCA1 C Terminus
CagA	cytotoxin-associated gene A
CD	cluster of differentiation
CDK	cyclin dependent kinase
CEA	carcinoembryonic antigen
CHK	checkpoint kinase
CI	chromosomal instability
Ct	cycle threshold
CRC	colorectal cancer
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
DDR	DNA damage response

DFS	disease free survival
DHFR	dihydrofolate reductase
DNA	deoxyribonucleic acid
cDNA	complementary DNA
DNAPKcs	DNA dependent protein kinase catalytic subunit
DSB	double strand break
DXT	radiotherapy
E	efficiency (of PCR)
EGFR	epidermal growth factor receptor
EXO1	exonuclease 1
FHA	forkhead-associated domain
FIT	faecal immunochemical test
FOBT	faecal occult blood test
FU	fluorouracil
G1	growth 1 (phase of cell cycle)
G2	growth 2 (phase of cell cycle)
GI	gastrointestinal
GR	grade of regression
Gy	Grays
H&E	haematoxylin and eosin
HNPCC	hereditary non-polyposis colorectal cancer
HR	hazard ratio
hTERT	human telomerase reverse transcriptase
hTR	human telomerase RNA
IGF2R	insulin like growth factor-2 receptor

IHC	immunohistochemistry / immunohistochemical
JNK	c-Jun N-terminal kinase
KM	Kaplan-Meier
LCR	long course regime
LOH	loss of heterozygosity
MMR	mismatch repair
MRFF	medical research future fund
mRNA	messenger RNA
MRN	Mre11-Rad50-Nbs1 complex
MSI	microsatellite instability
MSI-H	microsatellite instability – high
MSI-L	microsatellite instability - low
MSS	microsatellite stable
MTT	methyl thiazolyl tetrazolium
NBS1	Nijmegen breakage syndrome 1
NF-κB	nuclear factor kappa B
NHEJ	non-homologous end joining
NOD	non-obese diabetic
NSW	New South Wales
OS	overall survival
PBS	phosphate buffered saline
PCNA	proliferating cell nuclear antigen
PCR	polymerase chain reaction
PIKK	phosphatidylinositol 3-kinase-like
PML	Promyelocytic leukemia protein

PNI	perineural invasion
pRb	retinoblastoma protein
PRKDC	Protein Kinase, DNA-Activated, Catalytic Polypeptide
pTNM	pathological TNM staging
REST	relative expression software tool
RFS	relapse free survival
ROC	receiver operating characteristics
RT-PCR	reverse transcription polymerase chain reaction
S	synthesis (phase of cell cycle)
SCR	short course regime
SEER	Surveillance Epidemiology and End Results
SFx	surviving fraction at/after x Gy of ionizing radiation
SHP2	Src homology region 2-containing protein tyrosine phosphatase 2
SNPs	Single nucleotide polymorphisms
SSbs	sugar-sweetened beverages
SWSLHD	South Western Sydney local health district
TBP	TATA box binding protein
TBST	tris-buffered saline with tween 20
TC	tumour core
TGFβ	transforming growth factor-β
TGFβR	transforming growth factor-β receptor
TIMP1	metallopeptidase inhibitor 1
TMA	tissue microarray
TME	total mesorectal excision
TNF	tumour necrosis factor

TNFR	tumour necrosis factor receptor
TNM	tumour-node-metastasis
TP	tumour periphery
TRG	tumour regression grade
TSG	tumour suppressor genes
VEGF	vascular endothelial growth factor
WHO	World Health Organisation
95% CI	95 percent confidence interval

CHAPTER 1
INTRODUCTION

1.1 The basis of biological systems and human health in modernity

Without a grounding in the workings of our environment, cells, proteins, genes, and evolution, understanding cancer is difficult (Berg et al., 2013). Therefore, before considering the molecular mechanisms of rectal cancer, the DNA damage response, biomarkers, and how the pathology can be diagnosed and treated, it is useful to consider how biological systems first came into existence, and how, in general, we understand them.

1.1.1 Chemical beginnings

Life on Earth is believed to have originated from basic chemical reactions between atoms and molecules interacting in the complex soup of the early atmosphere – *abiogenesis* (Maher and Stevenson, 1988). On Earth, such reactions may have paved the way for the development of more complex molecules and networks of reactions (perhaps put in place by chemical catalysis already existing in the present matter) – like adenosine triphosphate (ATP), carbohydrates and lipids – which allowed for energy-dependent cellular processes to evolve and biological information to be stored. The triggers for such chemical activity are not known, although lightning and radiation are often primary candidate environmental cues (Martin et al., 2014).

The precise sequence of molecular events leading to complex organic life remain a matter of intense research and debate, although various studies have demonstrated that essential biological molecules, such as amino acids, can be formed from their constituent parts under conditions that resemble those of the young Earth (Hörst et al., 2012). As complex organic molecules have been found in space, it is difficult to rule out that they were the precursors to life; the Panspermia hypothesis (Gribbin, 1999).

Regardless where in the universe life may have started, science has shown us that fundamental elements exhibit properties allowing for life. For example, carbon and water - both of which are readily abundant. Carbon is unique in its ability to provide a stable scaffold for other elements, such as oxygen, allowing complex molecules to stably exist (i.e. CO_2). Water, on the other hand, is nature's great solvent, and by having both electrically negative and positive domains, it allows a wide range of molecules to dissolve in it (Clayden et al., 2001). Furthermore, as water is less dense when frozen compared to its liquid state, ice floats and cellular life can exist below it.

Based upon carbon and water, diverse lipids, carbohydrates, and nitrogenous bases can be generated. Nucleic acids, for example, are built from monosaccharides, nitrogenous bases and phosphate groups, and together encode deoxyribonucleic acid (DNA) and ribonucleic acid (RNA) recipes. A considerable portion of the biological community considers that life first emerged in an RNA world, in which proteins could be made from the RNA messages replicating themselves in the environment (Robertson and Joyce, 2012). From such collections of proteins, networks of interactions are predicted to have joined together to generate new molecules and carry out tasks that would presumably have benefitted that present in the environment. Nucleic acids, DNA and RNA store the information for these complex networks

to exist. DNA is as a linear molecule consisting of different bases – adenine, cytosine, guanine, thymine – joined together by a phosphate backbone.

In respect to living organisms, it's useful to remember that we (and all our cells) are all built from DNA-encoded messages, many of which are shared amongst ourselves, plants and microorganisms; systems are natural, related and dynamically adapt to their environment; cancer included. DNA is responsible for cancer, a disease in which a cell in a multi-cellular organism becomes rogue to its autologous neighbours and attempts clonal expansion and invasive tissue formation (neoplasia).

Zircon-preserved carbon suggests that biogenic life is at least 4.1 billion years old; whereas Earth formed circa 4.5 billion years ago (Bell et al., 2015). Putative fossilised organisms – dated to be circa 4.2 billion years old – have also been discovered in hydrothermal vents in the deep ocean, meaning that they would have come into existence shortly after the oceans themselves were formed, 4.4 billion years ago. These time scales provide ample time for the evolution of the systems that we are discussing. At present, scientific theory predicts that early life on Earth first consisted mainly of bacterial and archaeal biofilms which replicated asexually (Donlan, 2002). The later development of photosynthesis, approximately 3.5 billion years ago, led to a build-up of oxygen in the atmosphere, causing the great oxygenation event of approximately 2.4 billion years ago. Around 1.8 billion years ago, eukaryotic cells containing mitochondria (allowing for oxygen to be more efficiently used metabolically) emerged. Endosymbiotic theory holds that mitochondria and other free-living prokaryotes, such as chloroplasts, were joined to form more complex cells (Archibald, 2015). This ability to metabolise oxygen efficiently via the mitochondrion (to generate energy stored as ATP) is thought to have spurred the development of multi-cellular life – which is currently dated to have occurred approximately 1.7 billion ago – allowing individual cells within the same organism to become specialised and carry out unique functions; endowing organisms with increased functional capacities, and thus, abilities to exploit diverse ecological niches. Given the time at which all these early evolutionary events took place, it is perhaps unsurprising that the origins of sexual reproduction are also shrouded in mystery.

The remarkable similarity between different species (now known to be due to shared genetic mechanisms) was the defining feature of nature leading to Charles Darwin's Theory of Evolution, published in 1859. In it, Darwin suggests that natural selection of inherited genomic variants allows populations to better adapt to their environment; those variants ill-suited to the present are more likely to not be passed on to subsequent generations given that they encode a *reduced fitness* (Shanks and Pyles, 2007). When thinking about cancer, one can view the loss of regulatory genomic control in tumour cells as leading to a *survival of the fittest scenario*, in which the cancer outcompetes the normal cells to benefit its own survival and replication; that generated by the variants present in its genes, and as has been shown in cancer stem cell resistance to therapeutics (Eyler and Rich, 2008).

These evolutionary systems have led to an explosion of life, with life forms as diverse in appearance and function as imaginable. This diversity is genetically encoded in DNA. Yet, life

is also short-lived and dynamic, with new genes replacing new ones every generation and every cell division (Stratton et al., 2009). Sadly, we are also living through a mass extinction event driven by human population expansion, urbanisation, pollution and illegal activities, and it is estimated that natural causes have led to more than 99% of all species to have inhabited the Earth to be now extinct (Ceballos et al., 2015). Presently, approximately 14 million species are hypothesised to be free-living.

Most biologists now consider life to have a single common ancestor, as it is highly improbable that unique lineages evolved the same complex biochemical mechanisms common to all of them. With that in mind, much can be learnt about diseases, such as cancer, from studies in related animals that share the same proto-oncogenes, such as mice and dogs, and indeed we owe a great deal of debt to the discoveries made in animal models that have benefitted human health (MacEwen, 1990).
Remarkably, certain types of cancer have also been reported to be infectious and transmissible between members of the same species; such as canine transmissible venereal tumours and Tasmanian devil facial tumour disease (Murchison, 2008). Transmissible tumours are not thought to occur in humans. The use of animals in research has also recently shown RNA-lipoplexes are able to deliver neo-antigens to dendritic cells in murine lymph nodes, which then prime T cells to attack tumour cells around the body. These studies demonstrated a reduction an IFN-α-dependent T cell rejection of progressive tumours (Kranz et al., 2016).

The importance of our relatedness to other organisms cannot be overstated when it comes to defining molecular mechanism and understanding the evolution of life (Moulder, 1985).

1.1.2 Genomics, health and disease in the 21st century

Since the Human Genome Project completed its primary aim in 2003, our understanding of human biology and disease has been transformed (Siva, 2008). In this tractable system (and map) we can refine biological knowledge at the molecular level in various ways. In fact, it guides many of the hypotheses used to enquire about the natural world.

Principally, we use genetics in the same way (albeit at higher resolution) that Mendel learnt of segregating entities in the 1860s; we increasingly define the genes and genomic marks (such as methylation) associated with various phenotypes and conditions (Westra et al., 2013; Xu and Li, 2012). Understanding such relationships greatly refine our understanding of pathophysiological mechanisms, as we can correlate genetic activity to function and clinical pathology.

In many ways, the increased genomic resolution available to human genetics - largely discussed in terms of the advent of massively parallel DNA sequencing (Shendure and Ji, 2008) - coupled to our increasing mechanistic understanding of the genome, proteome and metabolome, allows us to see a sub-cellular world afflicting our health; much in the way that the microscopes of Van Leeuwenhoektransformed our understanding of the equally important microbial world. A

future in which a person's genome is known and professionally consulted upon by public healthcare services is only a matter of time (Kaye et al., 2010). Indeed, the National Health Service in England was reported to be considering plans to have whole genome sequences produced from all children born in NHS hospitals (Gray, 2013).

To date, the genomic architecture of a multitude of human traits have been defined, such as diverse cancers, height, intelligence, neurodegeneration, immune responses and obesity. In prostate cancer, nine loci have been shown to be associated with disease development (Thomas et al., 2008). This was facilitated by genome-wide association studies, which analysed single nucleotide polymorphisms segregated between cohorts of interest and controls. In this way, the *common-variant, common-disease mechanism* hypothesis came about, which states that common variants present in populations (some with very low minor allele frequencies) together account for the genetic component of the majority of common, complex human diseases (Frayling et al., 2007); non-communicable diseases have been particularly well defined. In essence, we are defining the genetic features constituting our functioning and how these contribute to diseases, such as cancer.

This mechanism for disease, however, sits alongside the examples of monogenic conditions (and cancer mutations) where variants are highly penetrant and drive a clear phenotype individually; a topic reviewed well elsewhere (Duncan et al., 2014). It suffices to say that not all mutations are equal, and genetic complexity in disease seems endless, with each patient and lesion unique.

As we will discuss, the environment is thought to provide the additional risk for disease development, but determining the factors remains a challenge logistically in the clinic (i.e. recruitment of sufficient samples sizes and co-variable reduction between outbred humans) and in the laboratory (i.e. difficult to experiment in human subjects) (Emanuel et al., 2003).

Cancer, in many ways, is a prototypical genomic disease – single base-pair changes amongst 3.3 billion can lead to oncogenesis, and somatic mutation accumulation is a cancer hallmark. How and why these mutations arise remains a highly active area of research. Yet, in this day, it is safe to say that cancer cells have a different genome to their neighbours – with this uniqueness targeted diagnostically and therapeutically (Schumacher and Schreiber, 2015).

Given the nature of cancer, it stands to benefit tremendously from the technological advancement currently underway of relevance to biological and medical sciences – the so-called *–omic age* - where multi-dimensional, high-throughput phenotyping technologies (genomic, proteomic, and metabolomic read-outs) are combined with microfluidics, nano-scale engineering, and advanced programming and machine learning, to generate vast archives of biological data, and characterise and intervene in biological systems (Becher et al., 2014; Bendall et al., 2011; Brodin et al., 2015; Price et al., 2012).

In the next 30 years, one will likely have access to one's whole genome sequence (and, therefore, knowledge about cancer predisposition). This will allow for genetic counselling,

screening, and crucially, prior knowledge before seeing a doctor with an acute ailment. During the consultation, the doctor will have a wider array of precise and multi-dimensional bed-side assays available to generate a more complete picture of your health (Ginsburg and Willard, 2009). Indeed, recent breakthroughs in liquid biopsy technology - based on cell-free cancer DNA present in the patient's plasma - are testament to the fact that ever smaller sample volumes and less invasive procedures can diagnose cancer more quickly than was previously possible (Diaz and Bardelli, 2014). Couple this to tumour-specific pharmacogenomics (Wheeler et al., 2012), and more patients can be saved.

Through such ambitious approaches, not only will more diverse types of cancer be identified more quickly and accurately, and treatments targeted, but we will have a better idea of organ function through time, helping us care for our bodies and further improve human life expectancy. What's more, in the future, artificial intelligence will be able to examine billions of data points from an individual to more accurately determine the nature and potential cause of disease. Indeed, IBM has recently rolled out the computer, Watson, to do that just in hospitals around the world (Chen et al., 2016). Watson's answers are compared to human decisions and sometimes used to guide treatment. As computers can perceive data in more dimensions than we can, it may be only a matter of time before cognitive computing is diagnosing disease.

Already, the application of genomic techniques has been hugely beneficial to the fight against cancer. For several years, it has now been possible across many countries to have your genome sequence checked for highly-penetrant risk variants for cancer; with the examples of *BRCA1* and *BRCA2* in breast cancer standing out as great success stories. Here, crucially, genetics enables prevention, as prophylactic mastectomy proves Such advancements would not have been possible without basic research studies in genetics and cell biology, allowing the defective genes to be identified, and robust tests developed.

Applications such as these clearly illustrate the power of genetics and molecular biology in cancer medicine, and defining and characterising the molecules involved in disease development, maintenance and progression (such as DNA damage response proteins), remains our best hope to prevent, treat and cure disease. As we learn more about the different molecular compositions of different tumours(Shipitsin et al., 2007), we will be better placed to understand the defective cellular mechanisms involved and identify individuals susceptible to, or suffering from, disease.

1.2 The aetiology and epidemiology of cancer

Given the exciting developments underway in biomedical research and clinical medicine, we stand poised to make breakthroughs in cancer medicine. Cancer refers to a group of diseases characterised by rapid, un-regulated cellular division and growth in which the abnormal cells acquire aggressive, functions and invade and occupy different tissues than those of origin (Cooper and Hausman, 2007). In all cases, cancer evades the cellular regulatory mechanisms

in place to maintain normal tissue homeostasis, remodelling and repair; discussed in section 1.3.

1.2.1 The history of cancer

The acknowledgement of cancer as a disease entity is as old as human history, with several sources from ancient civilisations describing in detail the phenotypes and consequences associated with the disease. For example, the Edwin Smith Papyrus, a medical text from ancient Egypt written in 1600 BC (considered to be amongst the first attempts to distinguish medicine from magic) describes cases of breast cancer, amongst 48 other common ailments (Atta, 1999). Later, at around 400 BC, Hippocrates used the Greek term *karkinos* (meaning crab) to describe cancer morphology based on the highly-vascularised nature of sectioned tumours; providing valuable insight into the angiogenic nature of abnormal tissue formation (Sudhakar, 2009).

Subsequently, in 40 AD, the encyclopaedist Celsus translated *karkinos* to Latin, giving us the contemporary name, *cancer*. The Roman also suggested for the first time that surgical resection may be a means to halt the deleterious effect of the condition, something which we still rely upon in the present day (Karamanou et al., 2009). Unfortunately, however, another prominent scientist of the second century AD, Galen, disagreed with the notion of surgical resection, and given his sphere of influence, meant that cancer was treated mostly with purgatives for the next 1,000 years. These observations illustrate the pervasive and common nature of the disease throughout evolutionary history, suggesting that if environmental triggers are required for cancer formation, they may have been present in ancient times also; indeed, excessive exposure to UV radiation is a well-known risk factor for melanoma (Berwick et al., 2005).

Despite our historical appreciation of the clinical problems associated with cancer, it was not until the era of the European renaissance and the scientific method – in the 15th, 16th and 17th centuries – that our understanding of cancer was furthered in considerable detail. During this time, clinicians began to attempt to explain the causes of cancer through post-mortem dissections and modern pathological techniques. These studies led to the hypotheses that acidic lymph fluid and milk clots were responsible for tumour formation in different anatomical areas. An important insight into the disease at the time came from Nicolaes Tulp, who posited that cancer is like a poison able to spread from one location to another (Blumenthal, 1996). Indeed, Campbell De Morgan, building on the observations of John Hilland Percivall Pott (who found associations between cancer and tobacco and chimney sweeping, respectively), later exploited advances in microscopy at the time (between 1871-1874) to describe metastases (or cancer spreading) of the primary tumour to secondary lymphodensities, helping cement our fundamental understanding of the disease (Sudhakar, 2009).

1.2.2 An introduction to cancer biology

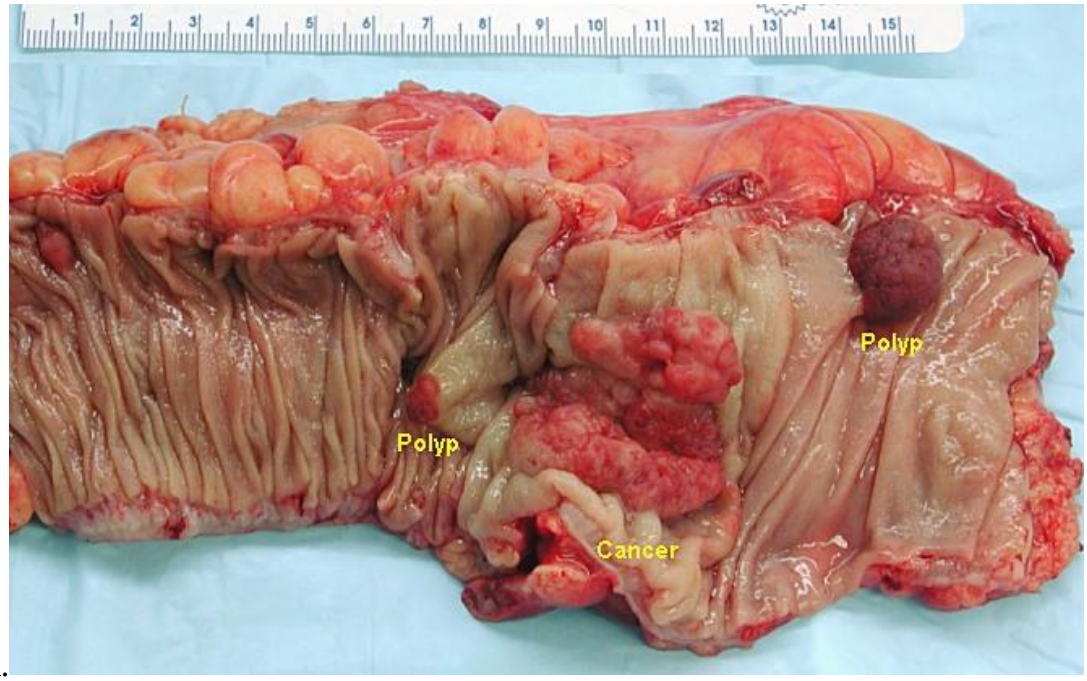
Wherever cancer cells are located, they are able to disrupt normal tissue organisation and function; often with fatal consequences. Figure 1.1 illustrates the nature of oncogenic lesions histologically, and in terms of their gross morphology. Such abnormal growths of cancerous

cells are referred to as tumours (and may be solid or diffuse), although it is important to remember that not all tumours are malignant (cancerous); with some being benign and lacking the ability to spread.

All cancers are a form of neoplasm (a new and abnormal growth of tissue), and according to present day clinical guidelines must adhere to the following to be considered malignant:

- Exhibit un-restrained cell growth and division in the absence of normal physiological stimuli.
- Exhibit continuous growth even in the presence of inhibitory signals.
- Exhibit a lack of programmed cell death pathway functions.
- Exhibit promotion of angiogenesis (for tumour nutrient supply).
- Exhibit limitless replication potential.
- Exhibit tissue invasion capacity, leading to the establishment of secondary tumours (metastases).

A.



B.

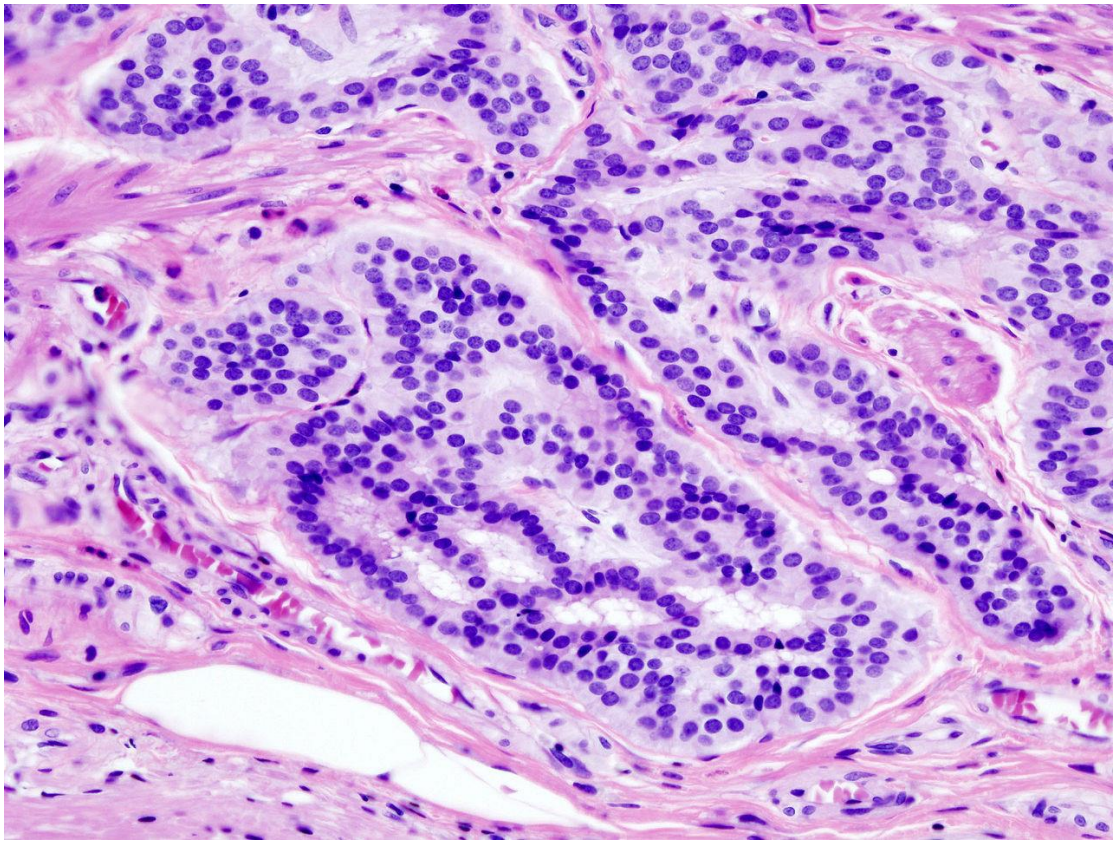


Figure 1-1Gross and histological pathology examples of cancer.

(A) Gross pathology specimen of a colectomy specimen showing an invasive colorectal carcinoma with its irregular disrupted form (B) Histological image of colonic carcinoid stained by haemotoxylin and eosin, with prominent mitotic figures seen

From: https://en.wikipedia.org/wiki/Adenocarcinoma

If we couple the destructive power of un-restrained cancer growth to a cancer's ability to spread and set-up secondary tumour sites, scenarios in which normal biological function is severely compromised are easy to appreciate. Early detection is key to catch the cancer before secondary metastases are formed. Depending on the nature of any particular tumour, its anatomical location, and metastatic potential, cancers can be more or less problematic clinically, and within different time frames; with some developing quickly and very deleteriously, and others taking years to progress to relevant phenotype. In the most extreme cases, the tumour burden of a tissue can lead to organ failure, and other life-threatening complications, such as peritonitis and sepsis arising from a loss of epithelial barrier integrity in the colon or rectum.

However, it is worth remembering that cancer cells do not only affect cells near to them. In many instances, paraneoplastic syndromes arise due to the systemic effects of tumour cells. Many paraneoplastic syndromes affect endocrine functions, due to cancer cell production/consumption of hormones, whilst other cancers are associated with neurological inflammation (Delellis and Xia, 2003). Overall, these changes lead to global dysregulation that further undermine health. Indeed, cancer always has a consequence on the body.

1.2.3 The biology of rectal and colorectal cancer

The focus of our investigations here presented is, specifically, rectal cancer – that is, malignant tumours affecting the most distal portion (bar the anus) of the human gastrointestinal tract. This point is emphasised as rectal cancer is often considered alongside (and the term is used synonymously with) colorectal or colon cancer when diagnosing and treating malignant disease in the distal intestine. This makes quantifying incidence and refining biological mechanism more challenging; as clinical heterogeneity abounds (Li and Lai, 2009). That said, current estimates suggest that approximately one third of colorectal cancer cases are comprised of tumours originating in the rectum.

Until the recent advent of advanced live surgical imaging and molecular genomic techniques (such as genome-wide mRNA transcription and DNA methylation profiling), which allow the tissue origin of biopsied tumours to be more clearly defined, the origin of the primary tumour – be it rectal, colonic or other segments of the distal bowel – was largely unknown, with clinicians relying upon exploratory surgery to determine the nature of the disease and treatment modalities.

Unfortunately, this lack of clinical stratification has hampered efforts to better understand disease aetiology and deliver targeted, effective treatments. For example, if we consider that a considerable portion of the rectum lies beyond the peritoneum and is associated with its own specialised lymphoid tissue to counter environmental challenge (Standring, 2008), the triggers required for malignancy in this location are likely often different to those driving oncogenesis in more proximal bowel segments, where dietary antigens may more strongly predispose to cellular transformation (Yu et al., 2001). In general, mechanistic differences between cancers originating in different anatomical locations impact not only how the cancer functions and

metastasises, but also how disease can be treated; as different cell lineages and tissues have specialised functions, the tissue of origin may endow important mechanistic properties to the malignancy. In future, with improved diagnostic procedures, cancers affecting different parts of the large bowel will be increasingly considered as separate diseases, removing some of the ambiguity arising from the clinical heterogeneity present between patients with various forms of large bowel cancer, thereby refining the molecular mechanisms at play. Such efforts will also be increasingly helped by the collection of human tissue biobanks (such as the Cambridge BioResource of the University of Cambridge; https://www.cambridgebioresource.group.cam.ac.uk), where clinically well-defined tissue samples from patients are deposited, processed and made available to basic researchers for discovery and drug development work.

Most tumours in these anatomical regions of the GI tract start off as polyps which protrude from the epithelial (inner) surface before becoming malignant (see Figure 1.2) (Subramaniam et al, 2016). Polyps may be pedunculated or sessile. Furthermore, although polyps are present at diverse locations in the GI tract, especially in older age, not all of them are cancerous, making delineating rectal and colon cancer difficult in many cases. Adenomas are more likely to become malignant (Brenner et al., 2007). Despite the difficulties associated with cancer location determination, research has shown that the clinical and biological features of gastrointestinal tumours, drug responses and patient prognoses depend upon tumour location, further suggestive of distinct aetiologies and functional consequences (Artinyan et al., 2008).

Further studies in different ethnic groups are required to determine the molecular mechanisms leading to disease in diverse populations, whilst patient out-reach programmes furthering medical education in particular communities is likely to be of broad benefit and can include modern approaches such as social media and community health groups to help alleviate the fears associated with cancer and the medical community in general.

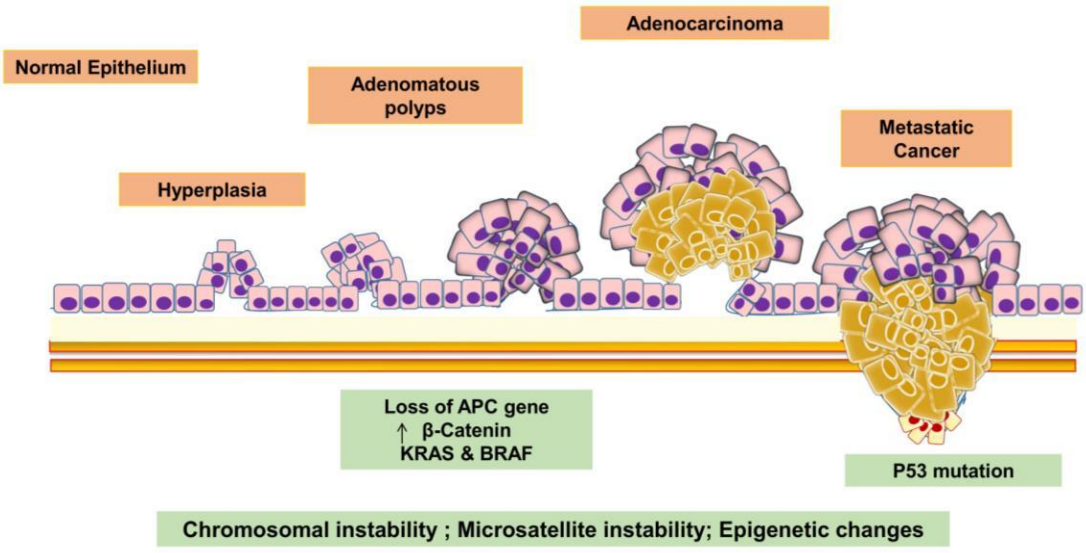


Figure 1-2 A schematic of colorectal cancer development.
From: Subramaniam et al, 2016.

Disease progresses from polyps in the intestinal/rectal epithelium to cancer that is able to invade the basement membrane and then metastasise. Candidate genetic triggers to oncogenesis are shown in the green boxes, and include the loss of DNA damage response protein functions (APC), oncogene activity (KRAS and BRAF) and loss of tumour suppressor gene function (p53), as we will discuss further. Importantly, the environmental triggers leading to poly generation and oncogenesis are not shown, but heavily impact upon genetic factors in cancer.

1.2.4 Cancer prevalence

At present, over 100 different types of cancer have been documented in humans, and according to the World Health Organisation, more than 90 million people had cancer during 2015, although diagnostics (especially relating to prostate cancer) are believed to account for a greater than 3% reduction in overall cancer incidence in males in the western world (Siegel et al., 2016). Fourteen million new cases are thought to arise every year. During 2015, cancer was estimated to be responsible for nearly 9 million deaths globally, or approximately 15% of all human deaths in that year; demonstrating the profound clinical and socio-economic burden of cancer worldwide.

Worryingly, these numbers trend upwards in many diverse populations (Jemal et al., 2011), especially as life expectancy increases worldwide. Indeed, the risk of developing cancer is strongly associated with age in the majority of cases affecting adults (White et al., 2014). One in 2 Australia men and women will be diagnosed with cancer by the age of 85 (AIHW, 2017a). The most commonly diagnosed cancer in Australia is prostate cancer, followed by colorectal cancer, breast cancer, melanoma of the skin and lung cancer. These cancers are estimated to account for about 60% of all cancers diagnosed (AIHW, 2017a).
The types of cancer most commonly causing death are lung, prostate and colorectal cancers in males and lung, breast and colorectal in women (AIHW, 2017a).
The most common cancers in children are acute lymphoblastic leukaemia and cerebral tumours, with approximately 150,000 children diagnosed with cancer worldwide per year (Pui et al., 2008). As we will discuss in Section 1.3, highly penetrant genomic lesions may predispose to cancer at younger ages, compared to older individuals, in which mutations and risk factors accumulate over time to cause pathology.

It is important to note that amongst all types of cancer, the 5-year survival rate stands at 68% of patients based in Australia, a world leader in cancer diagnostics and treatment (AIHW, 2017a). Of course, heterogeneous cancers are associated with different survival rates, and treatments can be more or less efficacious for different diseases, again highlighting the tremendous power of personalised genomic medicine to help target appropriate treatments to patients.

1.2.5 Colorectal cancer incidence and mortality

The Australian Institute of Health and Welfare (AIHW) is a national agency that collects data on vital health statistics including cancer. Cancer of the large intestine and cancer of the rectum are collectively known as colorectal cancer. The information provided below is extracted from their salient book on colorectal cancer (AIHW, 2017b).

Colorectal cancer was the third most commonly diagnosed cancer in Australia in 2013 however in 2017 it is thought that this will become the second most commonly diagnosed cancer. In 2013, there were 14,962 new cases of colorectal cancer diagnosed in Australia (8,214 males and 6,748 females). In 2017 this is projected to increase to an estimated 16,682 new cases of colorectal cancer Australia (9,127 males and 7,555 females). Additionally in 2017 the risk of a person being diagnosed with colorectal cancer by their 85th birthday will be 1 in 13 (1 in 11 males and 1 in 15 females).

The mortality data is striking. In 2014, colorectal cancer was the second leading cause of cancer deaths in Australia. This is projected to remain the second most common cause of cancer death in 2017. We know in 2014 that there were 4,071 deaths from colorectal cancer in Australia (2,236 males and 1,835 females). In 2017, it is estimated that this will increase to 4,114 deaths (2,136 males and 1,978 females).

Figure 1.3 shows the age-specific incidence and mortality rates for colorectal cancer, by sex in 2017.

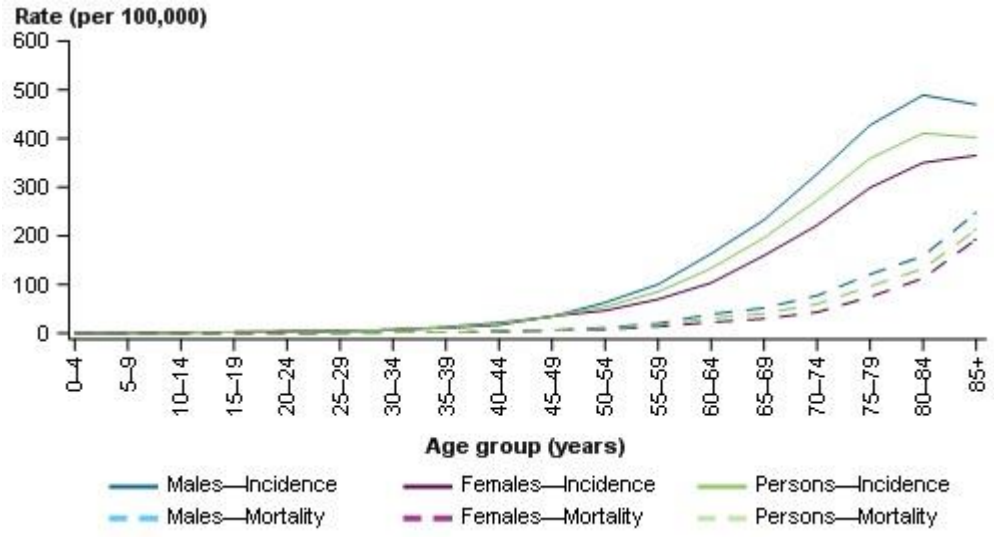


Figure 1-3 Estimated age-specific incidence and mortality rates for colorectal cancer, by sex, 2017 (AIHW, 2017b)

1.2.6 Cancer risk factors – genetics

When discussing specific risk factors, we will first consider genetics. Cancer is a loss of control over our genome. Estimates suggest that between 5-10% of cancers are due to heritable genetic

mechanisms (Garber and Offit, 2005). That does not include individuals who may be genetically predisposed to a particular phenotype (i.e. dysregulated inflammation) which interact with various environmental factors to accelerate cancer development. It is useful to consider genetics as what is underlying our ability to cope with the environment, to which we are all differently suited.

When considering cancer genes, two major types are well described – oncogenes (which lead to cell division and growth) and tumour suppressor genes (TSG - which curtail cancer by limiting cell lifespan and growth. Promotion of the former and suppression of the later are both deleterious, with many examples showing considered risk with heterozygosity. Oncogenes are formed by mutations to proto-oncogenes, which normally regulate the cell cycle. TSGs, on the other hand, are involved in apoptosis, DNA repair, and again, the cell cycle.

Well known examples of inherited cancer genes include *BRCA1* and *BRCA2*, discussed earlier, which are major contributors to breast and ovarian cancer development (Ford et al., 1998).

Further examples and mechanisms will be presented in Section 1.4, although other well-known cancer-associated genes worth considering (if only for their remarkable capacity to profoundly alter the cellular state) include *MYC*, which regulates cell cycle progression and apoptosis (Dang, 2012), and p53, the 'guardian of the genome', which inhibits angiogenesis, activates DNA repair proteins and arrests the cell cycle at the G1/S phase boundary (Ryan et al., 2001). Finally, the apoptosis regulator BCL-2 is also associated with many common translocation events (Czabotar et al., 2013); indeed, it was named B-cell lymphoma 2 due to the growth-promoting activity of variants in cancer cells.

As an aside: Perhaps disappointingly, given the amount of animals used in research, the majority of tumour models of disease focus on a limited subset of genes, such as *KRAS*, *MYC* and *JUN*, which are undoubtedly helpful for refining biological mechanisms, but they may have little similarity to cancers caused by other lesions, which far out-number those currently used. Despite their historic utility, murine models have recently come under criticism due to poor translation of medical findings to human subjects (Seok et al., 2013). Advances in induced pluripotent stem and the gene-editing technology, CRISPR, are also leading to increased impetus for the development of human *in vitro* models (Mali et al., 2013).

In colorectal cancer, more than 50% of patient tumours carry mutations in a kinase or phosphatase gene (such as *PIK3CA*) - families essential for positively and negatively regulating cellular function via post-translational signalling influencing differentiation and growth (Samuels et al, 2004). Given the erroneous activity of PIK3CA and its ability to drive oncogenesis, various drugs have been developed to target the enzyme and related factors, as we will discuss.

As well as these examples of penetrant causal lesions, polygenic susceptibility to cancer has also been described. Researchers found over 20 loci to contain SNPs predisposing to the five most common types of cancer (Easton and Eeles, 2008). In terms of colorectal cancer, the

authors found a SNP within the *EIF3H* to be associated with disease (minor allele frequency 0.77; odd's ratio per allele 1.25), supporting work by Zhang and colleagues who described the gene that promotes oncogenesis when the h-subunit is overly phosphorylated (Zhang et al., 2008). Additional variants predisposing to disease are likely to be identified in more powered studies benefitting from improved patient stratification. Crucially, it is not only the nuclear genome that is susceptible to mutation, with mtDNA coding and non-coding sequence variants being associated with diverse tumour types (Chatterjee et al., 2006).

Aside from inherited risk, cancer at the cellular level is a truly genetic disease, meaning that changes in the somatic genomic code lead to pathology. Some cells may be more predisposed to cell-cycle dysregulation according to their genetic background. When compared to harsh environmental pollutants that directly damage DNA and lead to oncogenesis, such interactions in day-to-day life may be less determining and more insidious in nature.

Further discussion of the molecular mechanisms of cancer is presented in Section 1.4.

1.2.7 Cancer risk factors – environment

However, the genetics is only one side of the story, as evidenced by the fact that monozygotic twins show discordance in cancer phenotypes (Castillo-Fernandez et al., 2014), and many environmental triggers have been defined. Indeed, environmental factors are thought to account for the majority of cancers – 90-95% of total cases (Anand et al., 2008). Viewed in this way, cancer can be seen as a disease arising due to DNA damage caused by the environment.

It is not only cancer that is proposed to be highly dependent on environmental triggers, but also our body in general. Autoimmune and inflammatory diseases, such as coeliac disease (Han et al., 2013), and the functional state of your immune system (Maecker, 2012), are also hypothesised to be more greatly affected by your environment than by your genes (Brodin et al., 2015); with migration studies showing a higher incidence of autoimmune diseases in migrants to a new country (Li et al., 2009).
Importantly, it is also widely accepted that cancer risk is higher in developed countries, where modifiable environmental risk factors (such as photoperiod modulation and western diets) and susceptible genetic variants may be more commonly found (Carrera-Bastos et al., 2011). In such complex environments, a trait for self-tolerance (i.e. to limit the deleterious effects of prolonged inflammation or autoimmunity directed against self-antigens) may limit the deleterious chronic inflammation associated with different complex polygenic diseases (Salmond et al., 2014), although predispose you to infectious disease otherwise covered by our immunological history.

Environmental risk factors for cancer are many and diverse, and include chemicals (or *carcinogens*), diet and physical activity, infections, ionising/non-ionising radiation, inert agents; and how these interact with under-pressure genomic variants in modern environments. One cannot overlook the fact that the majority of cancer predisposition comes from modifiable environmental exposures – suggesting it is preventable. However, it is a considerable challenge

to prevent cancer, given that the number of cancer-promoting compounds and entities is diverse, and expanding rapidly. Thankfully, my generation may see the end of the internal combustion engine in the personal automobile, which is thought to account for 40,000 deaths per year in the United Kingdom (Roberts, 2016).

Chemical carcinogens include substances like alcohol, tobacco, nitrosamines and polycyclic aromatic hydrocarbons, such as benzene, which can cause leukaemia (Miller and Miller, 1981). The list is large, and include natural and synthetic products in foods and our environments.

The importance of diet to health cannot be understated, especially in respect to gastrointestinal cancer. Estimates suggest that 10% of cancer deaths are due to obesity, excessive alcohol consumption, a lack of physical exercise, and a poor-quality diet (Lopez et al., 2006). In my opinion, given the strong associations between diet and metabolic syndrome and dysregulated inflammation in various tissues (Esposito et al., 2004), the number of cases affected by diet is far higher.

Although a myriad complex aetiologies lead to cancer, the known phrase, *you are what you eat*, really does ring true. It turns out that specialised lymphoid cells guarding the human gastrointestinal tract are only formed when the aryl-hydrocarbon receptor is properly stimulated, and the best described stimulants are organic compounds present in broccoli (Li et al., 2011); remembering that dysregulated inflammation is strongly associated with cancer. A balanced diet, rich in fresh produce (vegetables, fruits and limited amounts of animal products) providing carbohydrates, proteins, fats and vitamins), grains, and limited amounts of red meat is generally accepted to provide the body what is required for normal homeostasis (James et al., 2007). All in moderation, as toxicity occurs.

In cases where individuals do not care for their diet, and consume mostly heavily processed food ill-balanced in the aforementioned substances, the body does not receive the optimal fuel for function. Dietary deficiencies are tales as old as books (Sommer, 2008), and chimpanzees use their known medicinal plants to treat their bodies in times of illness (Huffman, 1997). As we increasingly consume food sourced from far and wide, and mostly from industrialised farming origins, the quality is likely to be far below optimum (Rostagno, 2009). Indeed, much needs to be done to improve animal welfare in such settings, which will also affect the chemical (i.e. hormonal) composition of meat, for example. Cortisol and steroidal hormones in general are associated with stress and they are potent regulators of cell fate and function - they have been reproducibly associated with cancer development; as shown by the driving role of oestrogens in prostate cancer (Nelles et al., 2011).

More specifically from whole food groups, a carcinogen is defined as a molecule that can promote carcinogenesis; through genomic mutation or metabolic dysregulation. Smoke from combustible organic matter contains many carcinogenic compounds, such as acrolein found in tobacco smoke, which permanently binds to guanine bases and induces mutations due to irreversible DNA alkylation (Alwis et al., 2015). Nicotine, also present in tobacco smoke, is not thought to be directly carcinogenic itself, although it does inhibit apoptosis, and thereby

may accelerate existing cancers in sub-clinical smokers (Cucina et al., 2012). The radioactive carcinogen ^{210}Pb is also present in tobacco smoke; equivalent to 0.1pCi per milligram of smoke. Tobacco, is probably the best know cancer-associated factor, yet remains popular he world over. The risk is known, although people enjoy the effects, which may also have shaped human evolution - after all, nicotine is a potent parasympathomimetic. Tobacco is associated with 22% of all cancer deaths, and given the information available on the subject I refer you accordingly (WHO, 2017).

Importantly, different chemicals can cause disease through different mechanisms. Alcohol, for example, is a known mutagen and the risk of developing cancer is higher in tissues directly exposed to it (Lachenmeier et al., 2009). Furthermore, acetaldehyde produced from alcohol metabolism is also carcinogenic. Thus, alcohol can predispose to cancer by inducing mutations in cells which do not die of exposure to high levels of it or its metabolites, or because loss of mature cells due to alcohol toxicity stimulates accelerated division of stem cells which accumulate potentially deleterious mutations normally associated with cell division; fewer divisions healthier stem cells.

Radiation is also considered carcinogenic as it can directly damage DNA, however, given its ability to damage DNA, it is also harnessed therapeutically to eradicate tumours – see radiotherapy, discussed below. Up to 10% of cancers are associated with ionising and non-ionising radiation, with the vast majority of non-invasive melanoma in humans arising from excessive ultraviolet radiation exposure (Anand et al., 2008). In terms of ionising radiation (which has the energy required to directly break molecular bonds, causing DSBs in DNA), models suggest that cancer risk increases in a linear fashion with radiation doses at a rate of 5.5% per Sievert (ICRP, 2007). Radiation-induced tumours are believed to be particularly insidious in nature, with solid tumours taking between 10-15 years to cause a noticeable phenotype (Braunstein and Nakamura, 2013).

Microorganisms are also important for cancer development as will be discussed in Section 1.4. In the developing world, approximately 10-20% of cancers are thought to be due to infectious disease (De Flora and La Maestra, 2014), such as those caused by hepatitis B and C viruses; which cause liver cancer. Epstein Barr virus infection is associated with non-Hodgkin lymphomas and nasopharyngeal cancer (Carbone et al., 2008), human immunodeficiency virus is associated with Kaposi sarcoma (Hoffmann et al., 2017), and papilloma viruses can cause cervical and penile tumours (Crosbie et al., 2013). The mechanisms used by viruses to transform human cells and dysregulate their function are diverse, and can afflict the genome, metabolome and proteome of the host cell.

Given the many environmental risk factors associated with the development of different forms of cancer, it is not surprising to learn that front-line medical advice for cancer avoidance includes: moderating alcohol intake; maintaining a healthy weight; getting vaccinated; maintaining good personal hygiene; avoiding tobacco consumption; and partaking in moderate and regular physical exercise.

1.2.8 Risk factors for colorectal cancer

Although the majority of rectal and colorectal cancers are thought to be predisposed to by an environmental trigger, several conditions of familial tumours have been described and the erroneous genes and syndromes identified. Furthermore, a familial history of cancer, inflammatory bowel disease, Crohn's disease and ulcerative colitis – all pathologies with a genetic component) are also associated with increased risk of rectal cancer (Eaden et al., 2001; Freeman, 2008; Kim and Chang, 2014); further suggesting that cancer arises from situations in which cell damage is increased.

Familial adenomatous polyposis, or Gardner syndrome, is caused by mutations in the adenomatous polyposis coli (*APC*) gene which mediates cellular adhesion; and is thought to contribute to a loss of barrier integrity in the intestine, necessitating increased cellular turnover and neoplasia (Galiatsatos and Foulkes, 2006). *APC* is defined as a tumour suppressor, and when its activity is reduced polyps form in the colon. Tumours can form directly in the rectum or arise there as a secondary site. Mutations in *STK11*, another tumour suppressor, and *MUTYH*, which makes the enzyme responsible for making the DNA repair protein MYH glycosylase, are also known genetic components in gastrointestinal cancers (Resta et al., 1998; Sampson et al., 2005).

High-fidelity DNA replication is essential to avoid cancer, and of relevance to our specific interests is what is known concerning Lynch syndrome – where mutations in several DNA mismatch repair enzymes have been described to lead to colorectal cancer (Vasen et al., 1999). These include *MLH1*, *MSH2*, *MLH3*, *MSH6*, *PMS1* and *PMS2*. In these cases, DNA replication is error prone and mutations accumulate.

1.2.9 Why is cancer increasing in prevalence?

Cancer is a non-communicable disease, at least in humans (so far), and as in all complex human diseases, genetics and environmental factors are known to cause disease. At present, we are living through the first period in human history in which more annual deaths are due to non-communicable diseases (such as cancer, obesity, metabolic diseases and autoimmune pathologies) than infectious disease (Boutayeb and Boutayeb, 2005); although considerable infectious disease burdens persist. As economic development progresses worldwide, this disparity is predicted to continue getting bigger. It suffices to say that cancer is getting more common all the time.

As we have discussed, increasing life expectancy and exogenous carcinogens are believed to be the direct causes of cancer, although how these interact with other physical features of our environment are largely unknown. For example, it can be argued that this shift in overall cause of death - occurring over a relatively short evolutionary time frame - is largely due to the discovery of microbes, and the discovery and development of antibiotics and vaccines, which drastically reduced the mortal burden from pathogens. The same pathogens which exerted

profound evolutionary selective pressure on the human genome, endowing it with the capacity to respond and exist in human environments of old and exist in harmony with our body (Okada et al., 2010). At the same time, we have completely redefined the human environment, bringing heating, fast/processed food, disinfectants and other harmful chemicals, pathogens and homeostasis-disrupting factors (such as artificial lighting, shift work and jet-lag; which disturb circadian rhythmicity and are associated with complex disease development, including cancer) into the fold (Wang et al., 2011). Therefore, determining the factors that interact with genes to cause diseases such as cancer is a daunting task.

Furthermore, although we rarely consider ourselves as vertebrates, we respond to many of the same environmental cues, and also exist symbiotically with diverse and common species; i.e. as with *Staphylococcus* and *Streptococcus* in our skin and gut microbiomes, which help defend us from opportunistic pathogens, as in dogs and cats (Cho and Blaser, 2012). Some have famously posited that an absence of microbes and immune stimulation is degenerative and predisposes to autoimmunity; with nothing to attack, the body attacks itself (Okada et al., 2010). Further along these lines, it is perhaps unsurprising that higher autoimmune disease incidence rates have been reported in individuals delivered by caesarean section, compared to natural births, where the mother's flora is predicted to engender normal immunity (Neu and Rushing, 2011). Such habits of modernity are, therefore, disease-associated, and are likely to influence cancer.

Furthermore, recent work from the United Kingdom has shown that humans across the world exhibit profound seasonal changes in gene expression and the cellular composition of organs; demonstrating the profound effect repeatable environmental changes, such as days/nights, months or years can have on how our body works (Dopico et al., 2015). In this study, the authors propose that seasonal changes may have evolved in-keeping with seasonal diets, pathogen burden, photoperiod and temperature; all of which we have largely removed from modern societies.

Thus, it can be argued that the gene-environment interactions taking place in the modern era are very different from those evolutionary ancient triggers. Accordingly, the change in the aetiology of human deaths requires clinical and basic re-conceptualisation of the roots and causes of human diseases, and disentangling new gene-environment interactions in health and disease is the major focus for clinical medicine in the 21st century. Nowhere is this truer than in cancer.

For many reasons, cancer is increasing, trends would need to be reversed. The global cancer burden is already considerable, and places significant socio-economic restrictions on societies worldwide (Jan et al., 2012). Even in high-income economies, such as the United Kingdom and Germany, public finances are not sufficient to cover the cost of certain novel anti-cancer treatments, such as anti-CTLA-4 (which costs hundreds of thousands of dollars per year), for the large number of patients who might benefit (Goldstein et al., 2016).

1.3 Screening, diagnostics and treatments for cancer

Many treatment options exist for cancer, with available options dependent on the cancer type, tumour grade, location, and the other clinically-relevant features of the patient, such as age and immune status.

Since we have shared knowledge on cancer, various people have made crucial leaps for our understanding of the disease and how to manage it. Hippocrates described cancer as a disease arising from a humoral imbalance, a view which prevailed for over 1,300 years. When various people saw tumour formation from isolates of lymphatic fluid in the 17th century, this view was refined (Sudhakar, 2009). However, Muller then demonstrated that lymph was not required, and his student Virchow showed that all cancer cells were derived from each other (Balkwill and Mantovani, 2001). Another considerable group of individuals maintained that parasitic and other infections caused cancer; and in 1911, Rous sarcoma virus was described to cause cancer in avian species (Weiss and Vogt, 2011). These principled understandings of the disease are remarkably accurate. However, between the 1800-1920s, the view that cancer arose due to trauma gained support, although this was poorly evidenced in the animal models at the time and subsequently.

As discussed, various triggers can cause rectal cancer, and oncogenesis in general. This heterogeneity requires innovative technological methods to achieve greater resolution when analysing the genome, metabolome and proteome. System biology analyses could be used to define networks of interacting factors in dimensions beyond our easy comprehending (Werner et al., 2014). Indeed, such approaches have already yielded success in the cancer field. Nano-sensors have a sensitivity of approximately 1 x 10^{-16} g ml^{-1} of protein in human serum samples (Kosaka et al., 2014). When this technology is used to detect the presence of cancer biomarkers, such as prostate specific antigen, it greatly reduces the amount of sample required, as well as allowing for cancer to be detected at a much earlier stage, when fewer malignant cells release their cancer-antigen contents into the blood (Kosaka et al., 2014). Such techniques could also be used to monitor the effects of treatment.

As we introduced earlier on, such high-throughput and hypothesis-free (*iterative*) methods promise to greatly refine clinical medicine. For example, a cancer's molecular profile can be generated before treatment is considered. Further, the fusion of such technology with biologically identified and verifiable biomarkers also promises to transform not just cancer detection, but for tissue-function profiling in general.

1.3.1 Cancer screening and early detection

Prior to reviewing the major means by which cancer is treated once diagnosed, we shall consider how we identify cancer. Indeed, a major limitation to successfully treating cancer is the time from disease origin to clinical phenotype (Ellis and Vandermeer, 2011). In many cases,

by the time a patient presents to his/her clinician for treatment, the disease is already advanced, making treatment more problematic and less likely to succeed.

There are many reasons for why this interval is difficult to close. Firstly, in the patient, cancer is synonymous with morbidity and mortality, as demonstrated by the widespread fear associated with these pathologies in the wider public. Such emotions are problematic for cancer diagnoses; as patients will often put-off physician consultations for fear of the C-word. Remembering that a patient may not even notice or deal with his/her tumour, and serendipitous discoveries are the exception not the norm. Hopefully, these perceptions will change, especially as the internet and public healthcare education initiatives help individuals better understand their health and what can be done in the case of disease (Barros et al., 2014). Through education, important risk factors can be avoided.

Unfortunately, it is also still difficult to collect a patient sample that can inform upon cancer status without an invasive biopsy. Ideally, cancer should be detected before a visible mass is present. Indeed, advances in cell capture, microfluidics and nanotechnology are transforming how we detect pathology.

At present, screening mostly takes place at the genetic level and protein level in patient samples (Andriole et al., 2009). *BRCA1* and *BRCA2*, both involved in breast, ovarian and pancreatic cancer development are routinely analysed for deleterious variants in those at increased cancer risk; i.e. those with a familial history. The genes encoding the DNA mismatch repair genes, *MLH1*, *MSH2*, *MSH6*, *PSM1* and *PSM2* are also typed with genetic tests given the increased relative risk for colorectal, uterine and stomach cancer associated with variants in these factors (Vasen et al., 1999).

Other screening approaches of note include the screening test for human papilloma virus and mammography, which represents a physical method to analyse abnormal tissue morphology *in situ* (Gyllensten et al., 2012). Such approaches allow cancer to be identified and treatments commenced – the sooner identification takes place, the better. As we will discuss, novel methods are slashing diagnosis intervals.

In respect to our studies, when a cancer is found, prognostic markers of tumour aggressiveness, and likely survival pre- and post-intervention can not only help the patient contextualise his/her situation, but also allows for treatments to be based on mechanism and the individual from the very beginning and improve patient quality of life. Reporting such biomarkers of patient outcomes following treatment not only refines disease mechanism but also take the patient to a different level of understanding his/her treatment. Thus, to improve cancer survival, biomarkers associated with the earliest events in cancer development need to be identified and rolled-out in the clinic.

The old adage *Prevention is better than cure* is true for colorectal cancer. Removal of an adenoma, a precursor lesion to colorectal cancer, has been shown to prevent cases of colorectal cancer occurring (Winawer et al, 2011).

Blood originating from these precancerous polyps lining the rectum and colon can be detected in minute concentrations by a high-sensitivity faecal occult blood test (Clavarino et al., 2004). There are two main types of faecal occult blood tests (FOBT). The Guaiac FOBT is used to detect trace amounts of haem, whilst the faecal immunochemical test (FIT) uses antibodies to detect haemoglobin directly. It is important to remember that blood in stools may not be associated with cancer, and indeed, both tests can be associated with false positives.

The Guaiac test has shown to be incredibly useful, reducing the number of deaths due to colorectal cancer by up to 33% when performed bi-annually in those aged between 50 and 80 (Ouyang et al., 2005). Additionally, nucleic acids present in the stool can be analysed to determine whether a patient suffers from colorectal cancer (Imperiale et al., 2004). Recently, expression of SEPT9 mRNA in the blood has been used as an FDA-approved biomarker for colorectal cancer, although it is no more sensitive than the FIT test and is yet to be proven to reduce cancer deaths (Johnson et al., 2014). Nevertheless, it represents another way in which genomic techniques can be used to detect cancer. Multiplexing such assays in a quantitative manner is likely to further the definition of cancer detected. Protein biomarkers used to diagnose rectal cancer include carcinoembryonic antigen, or CEA (Markman et al., 2010).

In these examples, sample collection is minimally invasive. Both require individuals to mail faecal samples to the analysis laboratory, with results being available in a matter of days. Subsequently, based on these results, patients can be referred for colonoscopy often through direct access clinics.

We are fortunate in that the Australian government has funded the National Bowel Cancer Screening Program which provides faecal occult blood testing for free to all eligible Australians. The Program is currently expanding and by 2020 all Australians aged between 50 and 74 years will be offered free screening every two years (CCA and ACN, 2005). The faecal immunochemical test has been selected as the preferred testing method.

Colonoscopy – one of the most sensitive tests available for colorectal cancers – involves the use of the flexible colonoscope, a flexible tube containing a lens and tissue dissection tool. Patients need to maintain dietary restriction prior to the procedure, and risks of tears to the epithelium are a risk. A smaller alternative to colonoscopy that is particularly useful for rectal cancers (lower down the GI tract) is sigmoidoscopy, in which a light tube with a viewing lens is used to visualise tumours that can be simultaneously removed. Inert gases are commonly used to expand the colon to make visualisation easier higher up the GI tract. Regular screening with a sigmoidoscope and colonoscope is thought to account for a 60-70% decrease in colon cancer death in those over 50, with a single visit alone having a considerable beneficial effect (Atkin et al., 2010). Furthermore, as many polyps are benign, sigmoidoscopy allows biopsy samples to be collected. Subsequently, biopsies can be analysed histologically or using other molecular methods such as PCR to define carcinogenesis. Computed tomography can also be used as a virtual colonoscopy, where 3D reconstructions of the colon are generated by several

x-rays taken from outside the body (US Preventive Services Taks Force, 2016). Indeed, conventional X-rays on the lower abdomen are useful to identify cancer lesions.

1.3.2 Rectal cancer clinical features

Symptoms of rectal cancer include a change in bowel habits, such as constipation, unproductive straining, diarrhoea and abnormally-shaped faeces (Marley and Nan, 2016). General malaise and nausea is also commonly present, the appetite is decreased, energy levels are low, and cachexia arises. Rectal cancer is treated using a combination of surgical excision, radiotherapy and chemotherapy (Wolpin and Mayer, 2008).

1.3.3 Surgery

Surgery is very often the first-line treatment for a solid tumour, especially for colorectal neoplasia. Surgery involves resection of the tumour, and aims to leave no diseased cells present in the adjacent environment (Phang et al., 2002). Furthermore, surgery must avoid excessive tissue scarring which is known to contribute to the later successful seeding of any remaining tumour clones. For this reason, surgery is almost always applied with simultaneous chemo and radiotherapy, pre- and post-operatively.

The difficulties associated with surgery, are again, well appreciated. For many thousands of years, cancer has been known to reoccur after resection of the visible mass. Surgery entered a new phase when anaesthesia became available in the 1840s, allowing surgeons at the time to explore adjacent body areas to where tumours may have spread. This lead to the common practice of removing the draining lymph nodes of the tumour; as lymph carries solutes and cells from diverse tissues in the environment for profiling by the immune system, and is a major channel through which body cells move (Sudhakar, 2009). The practise also helped cement the notion that the periphery of the tumour environment is of primary importance to clinical outcome.

Nowadays, we rely less on exploratory surgery, and more on advanced medical imaging techniques to first characterise tumour size, number and plan surgical approaches (Frangioni, 2008). Imaging modalities of relevance to cancer surgery include, magnetic resonance imaging (MRI), ultrasound, computed tomography (CT; X-rays) and positron emission tomography (PET). As different modalities have different suitability to different tissues, they are used as appropriate. Unfortunately, the costs associated with some of these techniques often preclude thorough cancer diagnoses, even in wealthy countries.

Tumours can be removed in increasingly less-invasive ways. For example, endoscopic techniques and high-resolution cameras allow for tumours to be directly aspirated from an internal surface, which will considerably reduce the tumour burden and facilitate concomitant chemotherapy. Tumours are also resected with lasers in more modern approaches, which allows for fine boundary control, and *in situ* destruction using liquid nitrogen has also been

successfully deployed (Sudhakar, 2009). And high-standard storage and processing of surgical samples is invaluable for understanding the basis of disease on an *ad hoc* basis.

In order for colorectal cancer to be cured the malignant lesion needs to be totally resected including draining lymphatics. Smaller colorectal cancers that have yet to invade into adjacent structures can be treated by anterior resection (AR) with partial mesorectal excision (PME) or total mesorectal excision (TME) (Steele 1999), abdomino-perineal resection (APR) and transanal excision.

For colorectal cancers invading deeply into the wall that are impossible to dissect by total mesorectal excision, abdomino-perineal resection (APR) is the treatment of choice. However this does mean that patients end up with a permanent colostomy (Dorudi et al, 2002). More superficial colorectal cancers (cancers not spread beyond submucosa) can be removed by transanal excision. Although postoperative outcomes are usually better a major limitation of transanal excision is incomplete removal of tumour and associated regional lymph nodes that can increase the risk of local recurrence (Endreseth et al, 2005). For rectal cancer the combination of chemoradiotherapy with transanal excision improves outcomes (Steele et al, 1999; Duek et al, 2008).

1.3.4 Chemotherapy

Chemotherapy refers to the treatment of cancer (or alleviation of symptoms) using chemical entities that target rapidly dividing cells. Today, chemotherapy is more commonly used to describe non-selective intracellular poisons, most commonly targeting the mitosis inhibition. This is to separate it from targeted therapies that inhibit specific growth-promoting pathways, such as tyrosine receptor inhibitors and hormonal treatments. Chemotherapy can be given as an adjuvant therapy after surgery to kill any remaining cancer cells, and it can also be delivered as a neoadjuvant, to shrink the tumour prior to surgery – often employed in rectal cancer (Tubiana, 1987).

Chemotherapy is described as being cytotoxic and anti-neoplastic, and is associated with a significant burden (Lung Cancer Guide Book, 2013). Indeed, chemotherapeutics are normally systemically administered and have the potential to affect many cells across the body. In addition, they are often combined with local treatments (such as radiotherapy) to ensure that the most rapidly-dividing cells succumb to their toxic effects. The most rapidly dividing cells in the body – those in the intestine, hair follicles and bone marrow – are those most commonly affected by such anti-mitotic drugs as they have a high turnover rate; leading to myelosuppression, alopecia and mucositis. Furthermore, different compounds are associated with additional specific side effects, as shown by the neuropathy caused by oxaliplatin (Saif and Reardon, 2005).

Along these lines, many chemotherapeutic drugs are available to treat cancer, and they may be given in a palliative manner or with curative intent. The major classes are: alkylating agents;

antimetabolites; microtubule inhibitors; topoisomerase inhibitors; and cytotoxic antibiotics (Tol et al., 2009). Alkylating agents are derived from mustard gas, and crosslink DNA to generate DSBs (Fu et al., 2012). Cisplatin and carboplatin are known examples.

Antimetabolites take a different approach to reduce the cancer burden. This class of chemotherapeutics resemble either a nucleoside or nucleobase, although they are specifically altered (Parker, 2009). Their modifications can, for example, prevent DNA replication after they have been incorporated into the genome of new daughter cancer clones. Other anti-metabolites inhibit enzymes important for DNA synthesis to halt mitosis. Anti-folates such as methotrexate block the activity of dihydrofolate reductase (DHFR), reducing purine synthesis, and therefore, DNA-based cell division (Mikkelsen and Thorn, 2011).

Chemotherapeutics have also been designed that target microtubules in cells to stop alpha- and beta-tubulin participating in cytokinesis, genome segregation and division. Vinca alkaloids (such as those from *Catharanthusroseus*) and taxanes are the two most known subgroups of these drugs (Crown and O'Leary, 2000). The alkaloids prevent microtubule formation, whilst the taxanes prevent their disassembly. In both cases, the cells cannot reorganise their intra-cellular structure and die.

Yet another mechanism is used by topoisomerase inhibitors. Opening (unwinding) the DNA helix during replication and transcription, leads to increased tension in the preceding unwound DNA yet to be transcribed/replicated. This tension can lead to breaks in the DNA (cueing apoptosis if the burden is too high) and is relieved by topoisomerases I and II (Hande, 2008). Inhibition of these enzymes therefore increased the mutation rate in cells, leading to increased cell death in rapidly dividing cells. Consequently, compounds like irinotecan and topotecan have been widely used in many cancers, including lung and ovarian cancer.

Details on small molecule inhibitors of kinases and phosphatases are beyond the scope of this thesis and are covered by excellent reviews by others (Hoelder et al., 2012). The final example of a chemotherapeutic to consider is cytotoxic antibiotics, such as mitomycin C, doxorubicin, and actinomycin – all of which again interfere with cell division through various means (Denard et al., 2012). Many such compounds were isolated from bacteria, which presumably also wanted to halt the replication of DNA perhaps foreign to themselves. These compounds are often combined with more targeted treatments aiming to inhibit specific pathways at work in the cancer.

Indeed, many agents are commonly used in combination, as this not only provides additive effects of different tumour-inhibitory drugs, but also reduces the toxicity associated with metabolic degradation of each component. Chemotherapeutics are also widely used to maintain remission for a period after surgery, often at lower doses to negate the side effects as much as possible.

The most commonly used adjunct chemotherapeutics include fluorouracil and capecitabine (Douillard et al., 2000). Fluorouracil is most commonly delivered intravenously one day prior

to radiotherapy, and is thought to act as a thymidylate synthase inhibitor, reducing the levels of nucleosides required for DNA replication. Calcium folinate is also commonly used to stabilise fluorouracil and increased its potency *in vivo*. Capecitabine belongs to the same family as fluorouracil, and is on the WHOs list of Essential Medicines (Hogerzeil, 2004). Indeed, it is metabolised to fluorouracil. Other chemotherapy protocols used in the management of colorectal tumours are described by Kelly and Cassidy (2007), and is dependent upon tumour grade, metastases and other clinical variables.

Given the toxicity associated with chemotherapy, regional applications are sometimes employed, as shown when hepatic artery infusion is used to treat disease that has metastasised to the liver (Karanicolas et al., 2014). Furthermore, chemotherapy is delivered in treatment cycles typically lasting no more than a few weeks to allow patients to recover. Normally, several cycles of difficult chemotherapy are required.

1.3.5 Radiotherapy

Shortly after they were discovered, X-rays were already in the cancer clinic, in what must be one of the fastest times from discovery to widespread medical application in history (Teldo-Pereyra, 2009). It was an important finding. X-rays can directly ionise atoms in DNA (which need to be repaired to prevent apoptosis), and also generate hydroxyl radicals by ionising water that also damage DNA.

Radiotherapy (or radiation therapy) is now known to both cause cancer as well as kill cancer cells. In the most common treatment setting, ionising radiation is delivered by a linear accelerator to kill rapidly dividing cells, and is also commonly used as an adjuvant post-therapy to prevent remaining malignant cells from colonising the niche (Baskar et al., 2012). In brachytherapy, the radiation source is placed inside or adjacent to the area to be treated, and is useful for breast and prostate cancer. As ionising radiation can damage DNA in normal cells as well, the beam is focused to intercept the tumour from various angles to more selectively kill the tumour and not the healthy tissue. Where that boundary begins remains the matter of much debate.

Interestingly, depending on the type of radiation used, the induction of DNA damage occurs by different mechanisms. This is important for treating tumours with a hypoxic centre which are more resistant to photon radiation therapy which works predominantly through hydroxyl radicals (Baskar et al., 2014). Other radiation sources, such as boron, carbon, lead and neon usually cause double strand DNA breaks (DSBs) through direct energy transfer, and are thus not affected by the low oxygen environment.

Unsurprisingly, radiotherapy is associated with many negative and considerable side effects, although low dose palliative care is well tolerated (Lutz et al., 2014). The higher the dose, the higher the risk of unwanted side effects. Doses up to 80 Gy are commonly used, although oncologists consider various factors to determine the dose. Advances in the field seek to refine

the uniformity and accuracy of radiation delivery, and include methods such as intensity-modulated radiation therapy, which allows concave tumour shapes to be followed and targeted.

In terms of patient side-effects, oedema first leads to acute pain in the treated and damaged area. Systemically, radiotherapy also causes nausea and vomiting, especially if head, neck or inner ears are treated (Standring, 2008). Infertility is also commonly reported during and after the treatment of testicular and prostate cancer, and the hypothalamic-pituitary axis is also commonly dysregulated by radiotherapy (Stubbe and Valero, 2013). In rectal cancer patients receiving radiotherapy, diarrhoea is common due to the epithelial and endothelial damage generated (Österlund et al., 2007) – and is treated with *Lactobacillus* supplementation. Damaged epithelial surfaces typically become fibrotic and exhibit dryness, and in some cases, cancers arise due to the radiation-induced mutations of treatment.

However, although there are risks involved, X-ray therapy is still highly applicable to cancer therapy. In a study by the Early Breast Cancer Trialists' Collaborative Group (EBCTCG), radiotherapy was found to reduce 5-year relapse rate by up to 19% (Darby et al., 2011), echoed by several specialised protocols and oncogenic pathologies. More commonly than not, radiotherapy is combined with chemotherapy to deliver better outcome measures for patients, as both approaches exploit different and combinatorial mechanisms useful to reduce cancer cell growth (Tubiana, 1987).

For rectal cancer pre-operative radiotherapy has been shown to reduce local recurrence of rectal cancer after surgery (Colorectal Cancer Collaborative 2001). There are 2 preoperative radiotherapy regimens (Glimelius et al, 2003) available for the treatment of rectal cancer.
The long course regimen delivers 1.8 to 2.0 Grays (Gy) of ionising radiation to the tumour site for at least 5 days per week. The accumulated total dosage is around 45 – 50 Gy over a typical 5 week period. This may sometimes be augmented with chemotherapy. A 5-fluorouracil (5FU) infusion is provided over the first and last weeks of the long course. Surgery is carried out at least 4 weeks after the radiotherapy completion with that time delineated as a reasonable amount of time for any post-radiation oedema to subside.

The short course regimen on the other hand delivers a larger radiation fraction at 5.0 Gy per day for 5 days. The short course regimen does not incur the oedema that is seen with the long course and thus surgical treatment of the rectal cancer can be carried out the following week (Glimelius and Isacsson 2001). The short course does have some advantages over the long course such as reducing the duration of the radiation course for patients and attendant reductions in overall financial burden. Although the total cumulative dosage (25Gy) of the short course is only half that of the long course (50Gy) the biological effective dosages of irradiation delivered by the two regimens are overall comparable (Glimelius and Isacsson 2001).

1.3.6 Immunotherapy

Recent publicity has led to the suggestion that immunotherapy is the biggest breakthrough in cancer care since the 1940s (Gattinoni et al., 2006). Indeed, cancers shown to be resistant to radiotherapy and chemotherapy have been shown to be susceptible to immunotherapeutic approaches, highlighting additional scope for treating clinically-resistant disease. The immune system is highly adapted not only to fight pathogens, but also in surveillance and remodelling of normal, healthy tissue, and has potent anti-tumour functions. Tumour antigen-specific T cells are known to directly destroy tumour cells, whilst antibodies against tumour antigen also aid in clearing oncogenic cells (Stockert et al., 1998).

Amazingly, the tumour microenvironment expands and becomes highly specialised, as demonstrated through by active dampening of the immune response based on chemokine antagonism designed to inhibit cells responding to cancer antigens (Coussens and Werb, 2002; Johanna and Fearon, 2015; Kraman et al., 2010). Cancer immunotherapy refers to the use of the host immune system to reverse this imbalance and treat cancer. The immune system is ideally suited to this task, being able to selectively eradicate specific cells based on their molecular profile; compared to systemic chemotherapy which damages many cell types. This accounts for reduced levels of tissue destruction due to immunotherapy compared to systemic chemotherapy and errant radiation. As always, the difficulties lie in properly harnessing the power of the system in a safe manner; too much or too little inflammation is deleterious.

Cancer immunotherapy is classified as active, passive or hybrid, and can be achieved with the use of different immunological components such as antibodies, T cell infusions and vaccines designed to engender responses to tumour-associated antigens (Voena and Chiarle, 2016). Active therapies dictate the immune system to destroy cancer cells based on the expression of tumour-associated antigens. Passive immunotherapies take a more global approach, and attempt to aid the efficiency of natural anti-tumour immune mechanisms.

To date, despite the promise, no T cell infusion protocol has been approved for clinical use, although various means to clonally expand tumour-reactive autologous cells have been developed, and genetic engineering approaches allow for specific anti-cancer T cell receptors and co-stimulatory molecules to be introduced. The fate of the adoptively transferred cells remains a concern, with their potential to attack other self-antigens and potential to form cancer themselves often quoted as precautions (Kalos and June, 2013).

A well-known target for immunotherapy is CD20, expressed on the surface of B cells and regulates B cell division and maturation towards plasma cells. An anti-CD20 monoclonal antibody, Ofatumumab, is FDA approved for the treatment of chronic lymphocytic leukaemia (Reagan and Castillo, 2014). Interestingly, monoclonal reagents directed against CD20 are also effective at alleviating autoimmune diabetes mellitus in mice, although the reason for this are still unclear (Barr et al., 2012). Blocking the T cell inhibitory molecule PD-L1 using a monoclonal antibody has also been reported to benefit different cancers (Festino et al., 2016). Furthermore, high doses of the potent T cell mitogen, IL-2, is used to treat melanoma and other

types of cancer (Oppenheim, 2007). Such examples help illustrate the central role of the immune system in surveying the body, and how this can be exploited to destroy cancer. Indeed, we must not forget that conserved, innate, anti-tumour immune mechanisms have been under evolutionary pressure for considerable time.

In addition to these recent approaches, vaccines that prevent infections with oncogenic viruses have also been developed, and represent another means the immune system can be harnessed to fight cancer. Vaccines against human papilloma virus (HPV) decrease the risk of cervical cancer, and Hepatitis B virus vaccines help reduce liver fibrosis and cellular transformation associated with viral spread (Schlom, 2012).

There is an emerging role for the use of immunotherapy for colorectal cancer. It is exciting to note that there has been recent successful use of checkpoint inhibitors in those persons with mismatch repair colorectal cancer that have a relatively high mutational burden (Lynch and Murphy, 2016).

1.3.7 Palliative care

The administration of non-curative treatments to reduce suffering and alleviate disease symptoms is very important for maintaining dignity and quality of life when suffering from cancer. In general, palliative care refers to any procedure in which the pain of disease or side-effects is mitigated as much as possible. At the same time, palliative care aims to improve upon the psychological and social problems that cancer entails, and the use of community-based health workers able to provide spiritual and holistic support to patients should not be overlooked as a mechanism to improve health (Mariano, 2007). Palliative care also helps patients deal with many of the practical life issues during a very difficult time, such as rising financial and legal complications.

Palliative care is given at any time during cancer treatment, often continuing for many years post-surgery, and increasingly, palliative care is delivered by specialised, multi-disciplinary teams able to appropriately deal with the patient's emotional state and presenting symptoms (Jocham et al., 2006). These treatments can be based in patient groups, in hospices and other community health centres. The role of the healthcare professional in making a patient feel secure and confident to talk about their condition is of primary importance in medicine, and professionals should make every effort to foster open lines of communication to generate understanding and benefit health management and treatment compliance.

Common palliative treatments include anti-inflammatory medications and pain killers, although cancer-effective agents such as radiation, chemotherapy and hormone therapy can be used to reduce the tumour burden, although treatment does not have curative intent.

Specifically for colorectal cancer there are palliative measures that can be carried out to overcome bowel obstruction such as colonic resection, creation of a stoma or stenting (Costi et

al, 2014). Other types of palliative measures such as radiotherapy, laser therapy and transanal procedures exist for the management of problems such as bleeding and pain.

1.3.8 Prognostic markers

Prognosis is the major focus of our investigations, and again dates its origins to Hippocrates, who wrote about it 400 BC. Determining the likely course of a medical condition is crucial for patient welfare and disease management. For example, notifying a patient that they are likely to respond to their current treatment and live another 2 years, is considerably different to a situation in which a patient does not respond. Nevertheless, prognosticating is a difficult task, often with insidious disease like cancer, which can affect diverse tissues in an occult manner. Furthermore, when large sample sizes are used and read-outs highly quantitative, prognostic indicators can be very useful, although when sample sizes are small and changes highly dynamic, prognosis is more error prone. Prognosis is also variable through time, and patients must be repeatedly monitored to evaluate the disease state (Armitage and Southam, 2016).

Following diagnosis, prognosis is impacted upon by a range of features in a patient. For example, disease-free survival time is known to be affected by:

- The patient's overall state of health, age and other co-morbidities, such as viral infection.
- The stage of the tumour at diagnosis; i.e. whether a rectal tumour has progressed through the wall of the rectum, or whether it has spread to one or more lymph nodes or other tissues. Metastases are negatively correlated with survival from colorectal cancer.
- The location of the tumour and burden across the affected tissue; i.e. does the entire rectum need to be removed?
- Whether the scenario represents a relapse or a first event; where relapses are associated with a worse prognosis.

It is also worth noting that although cancer incidence is lower in the developing world, cancer survival rates in the developing world are also lower (Kanavos, 2006); as often the resources required to effectively diagnose and treat the cancer are lacking in these environments. This necessitates the continual search for single panel biomarkers that can be effectively employed with minimal accessory equipment in isolated locations or where hospital infrastructure is lacking.

Several prognostic protocols are already in the clinic for diverse cancers. The Manchester Score, for example, evaluates prognosis in small-cell lung cancer (Herbst et al., 2008). The Manchester Score is a combination of scores in 6 different biochemical and physical categories: serum lactadehydrogenase, serum sodium, serum alkaline phosphatase, bicarbonate, Karnofsky performance status, and the stage of disease. The International Prognostic Index is used to predict outcome in non-Hodgkin lymphoma in a similar manner (Solal-Céligny et al., 2004). In many ways, these tests represent measures of dysregulated systemic function, and are not highly cancer specific.

Prognosis estimates are likely to further improve, however, as the molecular aetiology and nature of tumours is defined in real-time, and biomarkers of treatment effectiveness emerge in greater numbers. Furthermore, integration of vast –omic data with detailed clinical histories in cancer will also help predict those more likely to survive harsh treatment regimens.

Biomarkers are highly useful prognostic markers, being representative of physiological states. The levels of blood pressure, CRP, low-density lipoprotein and p53 are well-known examples (Karadag et al., 2008; Lutz and Nowakowska-Swirta, 2002). Their utility before and after treatment is of particular interest, giving molecular details of on-going internal processes in minimally invasive manners; i.e. from blood, saliva or urine. Biomarkers are also required to be highly sensitive and be highly specific for their target to have a clinical utility.

Many cancer biomarkers have been found to have prognostic value. For example, metallopeptidase inhibitor 1 (TIMP1) expression is associated with more aggressive forms of multiple myeloma (Guedez and Stetler-Stevenson, 2010), and a mutation in exon 11 of proto-oncogene c-Kit predicts responsiveness to the biologic imatinib, a tyrosine kinase inhibitor (Growney et al., 2005). Recently, a Chinese group elegantly showed that hydroxy-2′-deoxyguanosine (8-OHdG), the most common by-product of oxidative DNA damage, could be more easily detected in the urine of colorectal cancer patients with metastatic disease (Guo et al., 2016). Thus, although such a biomarker may be good for identifying advanced disease in a non-invasive manner, its ability to detect cancers at an earlier stage remains unknown.

Further defining the multitude of molecular mechanisms at work in tumours of the lower GI tract is likely to allow for prognostic and treatment avenues to be pursued. However, the difficulty in defining a novel prognostic or treatment option is challenging. For a test/biomarker to have useful prognostic value, the mechanism must be common to all or a subset of individuals in the cohort. Given the genetic and cellular variability of cancer, identifying such biomarkers can be challenging. Often, biomarkers identify hallmark processes in the organism. As demonstrated by the clinically well-established increase of the inflammatory C-reactive protein (to measure inflammation) and interleukin-6, which mediate increased risk for cardiovascular disease (Dopico et al., 2015); another major non-communicable disease killer that can be largely avoided by reducing one's exposure to the detrimental factors here discussed.

As we evaluate the role of biomarkers, it is worth considering the avenues that could be taken with respect to the DNA damage response proteins here investigated (see Discussion section).

1.3.9 Current state-of-the-art in cancer medicine

Before moving on, it is worth summarising some of the most interesting developments in cancer medicine. In general, as we move towards personalised medicine, targeted and individually-tailored treatments are sought, to more specifically target tumour cells and

improve treatment efficacy based on the molecular mechanisms of disease. Advances in screening and patient care are equally important.

Between 2015 and 2016, the USA Federal Drug Administration approved eight new cancer drugs for the US market, and allowed an additional 12 to be used in other oncology therapeutic areas (Buffery, 2015). Many of these represented advances against difficult to treat cancers, such as chronic lymphocytic leukaemia. This is the result of the vast amount of basic and drug development effort directed towards cancer. The National Cancer Institute in at the National Institutes of Health, USA, benefits from an approximate 5.4 billion dollar budget per year (https://www.cancer.gov). In Australia given the constraints of the current NHMRC grant funding system we look at the Medical Research Future Fund (MRFF) with a current pool of around $1 billion to see if this can assist towards the development of new cancer based therapies.

In basic research, recent success is shown by efforts to define the earliest genes mutated along the path to melanoma from precursor lesions, including *BRAF* V600E mutations, *NRAS*, *CDKN2A* and *PTEN*, at the same time helping define the role of UV radiation in pathophysiology (Shain et al., 2015). Furthermore, when talking about the bench-bed transition, few avenues have been as successful as liquid biopsy screening. For example, in lung cancer cases where resistance to epidermal growth factor-related drugs develops (due to mutation in the receptor), the presence of the EGFR T790M variant can be screened for in the blood, helping refine treatments more likely to be efficacious. The ability of liquid biopsy to predict colorectal cancer recurrence highlights its relevance to our field (Tie et al., 2016), where biomarkers of the DNA damage response in peripheral fluids may also have clinical utility.

As discussed previously, immunotherapy is providing considerable advances in cancer medicine, not least in unleashing T cells to more effectively target malignant cells. These studies are well complemented by consortia-based projects detailing genetic, epigenetic and transcriptional maps in diverse cancer cells and primary tumours (Baylin and Jones, 2011; Weinstein et al., 2013). Furthermore, novel monoclonal antibody-based reagents that target a range of proteins (such as the tyrosine kinases AXL, VEGFR2 and MET) *in vivo* have been shown to further extend survival in kidney cancer (Rini, 2005). VEGFR inhibitors have also been developed for the treatment of kidney cancer, and have recently been shown to be of benefit in clinical trials. Topical nicotinamide for the prevention of skin cancer also holds promise (Chen et al., 2015).

Advancements in surgical techniques, such as laparoscopic-assisted resection, have also been used to improve outcomes in stage II and III rectal cancer (Fleshman et al., 2015). With regards to improving patient care and quality of life, the drug Olanzapine has been found useful in reducing the nausea and vomiting associated with chemotherapy (Navari et al., 2016).

Furthermore, increased focus on public education programmes will assist individuals in caring for the health more effectively and managing pathology when it arises. Patient non-compliance and self-harm contribute a major obstacle to human disease management, and methods to

overcome these barriers will be helped by more effective communication, which emphasises the positives of treatment (Cantor, 2007).

1.4 Molecular mechanisms in cancer: focus on colorectal cancer

1.4.1 Cancer development

Cancer-associated mutations can be present in the germ-line, or be induced by exogenous factors, such as the carcinogenic compounds generated during cooking red meat which have convincingly been associated with bowel cancer risk (Wang et al., 2012).

As we know, cancer arises due to cells accumulating mutations that allow them to escape normal regulatory mechanisms curtailing undesirable cell growth. These mutations are often so pervasive that cancer cells also demonstrate density-independent growth *in vitro* - unlike normal cells which halt the cell cycle by entering phase G_0 when too many of them exist in the environment (Cooper and Hausman, 2007). Cancer-derived cell lines like the T cell Jurkat line exhibit continual, un-restrained growth; reminiscent of the massively enlarged thymus sometimes seen in immune-compromised mice (such as the NOD.*scid*; a model of autoimmunity), where thymomas occur more frequently due to abnormal immune homeostasis in the periphery (Huang et al., 2011). Indeed, cancer requires the loss of genomic control of systems that control cell division; proto-oncogenes.

Theoretically, cancer can affect any cell type of the body, and tumours are classified according to their tissue and cellular origin. Carcinomas refers to cancer of epithelial origin, whilst sarcomas are malignancies of connective tissue, such as cartilage, muscle, bone and fibrous tissue (Cooper and Hausman, 2007). Lymphomas and leukaemias are those cancers arising from blood cell lineages and account for 8% of all cancers, perhaps due to their high turnover rate. Further classification of tumours is then dependent on the specific cell type involved.

Wherever we are in the body, although some deleterious variants can cause cancer directly, given their central roles in guarding the cell cycle and signal transduction, for example, often, somatic mutations accumulating over time in different loci are responsible for oncogenesis; reflecting a gradual, more insidious loss of cellular control. This fits with our emerging, predominating view for complex diseases – that genes, environment, and stochastic factors lead to disease, but the loci, triggers and events are unique to the individual, as discussed. When aiming to eliminate cancer, it is perhaps not surprising that so much effort is directed towards personalised medicine, where genomes, and cancer transcriptomics, proteomics and metabolomics, can be together interrogated to determine better aetiology and treat the malignancy.

On the whole, cancer is considered to be a clonal disease, where an individual cell with the required mutations for oncogenic transformation is the parent of all subsequently generated

cancer cells (Greaves and Maley, 2012). However, whilst cancer is thought to represent a lineage, it does not imply that the progenitor was a cancer cell from the very beginning, becoming progressively cancerous as it accumulates lesions. Importantly, once the cancer lineage is established, additional mutations can be acquired, driving diversification of the tumour microenvironment and expanding the genomic mechanisms contributing to pathology.

As cells begin to lose genomic and regulatory integrity, mutations favouring proliferation, growth and signal transduction are selected for, the cells transformed into neoplastic cells and acquire the ability to proliferate. Support for this multi-step development model for cancer also stems from the age association with cancer development, although as we know, the molecular landscape is more complicated (DeGregori, 2013).

Further support for this *developmental* model was also shown in early studies of colon carcinomas. Initially, abnormal proliferation of epithelial cells leads to polyp formation – at that time, a benign neoplasm. Additional mutation accumulation was then shown to transform such lesions into malignant ones, allowing for invasion of the basement membrane and metastases (Allen, 1995). However, alternative theories to explain cancer also merit attention. The cancer stem cell hypothesis posits that stem cells harbour the mutations required to drive oncogenesis in daughter cells, such that continual tissue replenishment is associated with a cancer predisposition; explaining why tumours return in the same patients. Presumably, such mutations affecting stem cells are encoded in the germ-line, or are acquired during early embryonic development. Regardless, cancer stem cells have been documented in the blood, prostate, breast, lung, liver, pancreas and colon (Rahman et al., 2011).

Regardless of the mechanisms giving rise to the cancer in the first place, it is very often the case that cancer cells go on to develop autocrine growth signalling (as shown for IL-6 and WNT) and other features that allow them to expand independently of other cells (Grivennikov and Karin, 2008). Indeed, cancers do various things to modify their environment to favour their growth. A classic example is how tumours promote angiogenesis to divert nutrients to the tumour and allow mass increase (Weis and Cheresh, 2011). Another example is shown by the reduced expression of adhesion molecules in cancer cells, helping cancer cells escape their tissue of origin, as seen in colorectal cancer (Paschos et al., 2009).

Cancer cells also normally lose contact inhibition of growth (as demonstrated by the cessation of fibroblast growth upon encountering a neighbouring cell), and secrete protease enzymes that digest the extracellular matrix to favour metastasis (Yin et al., 2009). Given these differences that favour the rapid expansion of cells above other considerations, it is perhaps not surprising that cancer cells do not differentiate along a given lineage in the same manner as normal cells; and thus, exhibit unique properties which may be important for disease management (Jögi et al., 2012).

As we will discuss in the subsequent section, the DNA damage response is associated with cancer development (O'Connor, 2015). Normal cells with damaged DNA undergo apoptosis, to avoid erroneous function and deleterious variant propagation. However, in many cancers,

cells with acquired mutation do not undergo programmed cell death, a feature that may be due to variation in DNA damage response genes or their ability to correctly function. Furthermore, activation of the DNA damage response (including genes such as *ATM* and *CHK2*) has also been shown to be induced (in diverse cancers, including lung, colon, breast and bladder) early during tumorigenesis, prior to genomic instability and cellular transformation, and is hypothesised to restrain cancer progression (Bartkova et al., 2005). Thus, the ability of the DNA damage response to mediate errors in DNA also influences cancer development.

A brief outline some of the best-described molecular mechanisms leading to tumorigenesis is presented below.

1.4.2 Oncogenes

Oncogenes were first discovered in the 1970s, with *RAS* being the first described (Fernandez-Medarde and Santos, 2011). They have the potential to cause cancer (due to their normal roles in primordial cell functions) and are frequently found in mutated tumours. Mechanistically, oncogenes mediate cellular survival in the face of programmed apoptosis, allowing cancer cells to escape the Hayflick limit (Shay and Wright, 2000).

A gene involved in such fundamental processes, such as cell division, that acquires a mutation leading to constitutive activation, will thereby become an oncogene; these are termed proto-oncogenes. Examples of proto-oncogenes in humans include, *RAS*, *WNT*, *ERK* and *TRK*, and others described previously. All of these examples represent genes involved in signal transduction pathways. Tyrosine kinase members of the Src- and BTK-families have been associated with colorectal cancer (Grávalos et al., 2007). Signal transducing molecules and cell cycle control genes are not the only oncogenes, however. Many oncogenes produce hormones at high levels, cueing mitosis (Delellis and Xia, 2003; Schuchard et al., 1993). Such effects can be helped by additional mutations favouring intra-cellular signalling pathways. In general, different pathways can be affected by such highly penetrant variants.

Another well-known example of a translocation event leading to oncogenesis (in this case mostly chronic myeloid leukaemia) in diverse populations is the Philadelphia chromosome, which results in the formation of the oncogenic tyrosine kinase *BCR-abl* (a fusion gene) between segments of chromosomes 9 and 22 (Ren, 2005). The generation of this protein leads to myelogenous leukaemia as it shows restricted activity when phosphorylating a diverse range of substrates and aiding cell division. This translocation is also found in approximately 20% of acute lymphoblastic leukaemia cases, and sometimes found in acute myelogenous leukaemia) (Talpaz et al., 2006). Given the activity of the fusion gene, this cancer is treated with tyrosine kinase inhibitors, again highlighting how the mechanism drives treatment. Bone marrow transplantation is also an option.

1.4.3 Tumour Suppressor Genes (TSG)

When a cell is under stress, and has suffered DNA damage, it employs mechanisms to halt growth and cue repair. These represent evolutionarily conserved anti-cancer mechanisms vital to genome integrity and multi-cellular life (Casás-Selves and DeGregori, 2011). In most cases, such TSG are transcription factors, which then activate/repress relevant genes. TSG may need to be turned on in the first place, which can be achieved by pattern recognition receptors in the nucleus and cytoplasm which sense, for example, degraded and free DNA.

Amazingly, the prototypical tumour suppressor, p53 (*TP53I*), is estimated to be mutated in approximately half of all cancers and has various functions in regulating apoptosis, the cell cycle and metabolism (Soussi and Wiman, 2007). Indeed, p53 is hypothesised to control the switch from aerobic respiration to anaerobic glycolysis, the latter of which cancer cells use for growth. *Rb* is another tumour suppressor gene involved in cell cycle progression. *BRCA1* and *BRCA2*, are also TSGs.

1.4.4 Genomics of colorectal cancer

The two major subtypes of colorectal cancer identified to date are defined according to the nature of genomic instability present in the tumour. Tumours with chromosomal instability (CIN) are most often found to carry mutations in *APC*, which regulates cell adhesion (Arends, 2013). Tumours with microsatellite instability (MSI; 15% of cases) are often found to have defective DNA mismatch repair pathway. A CpG island hypermethylator phenotype (CIMP) in colorectal cancer has also been described, where several DNA repair genes are suppressed by promoter methylation (Arends, 2013).

Overall, four different subtypes of colorectal cancer have been described based on their MSI and CIMP status, with differences in their incidences between populations. It is accepted that these different sub-types are associated with unique clinicopathologic and molecular features, meaning a different aetiology, and different treatment requirements (Kang, 2011); as demonstrated by the differing susceptibility of these subtypes to adjuvant therapy (Bae et al., 2016).

Genome, transcriptome, proteome and metabolome analyses in cancer cells will likely help define new subsets of colorectal cancer, as shown for breast cancer (Curtis et al., 2012), and delineating rectal and colonic tumours biologically. Early studies in colorectal tumours have shown that it is not only the cancer transcriptome that is altered and determines cancer functional capacity, the presence of a cancer-associated fibroblast (making up the tumour stroma) transcriptional signature has also been associated with worse prognosis (Isella et al., 2015). Transcriptional studies in colon cancer cell lines have also helped cement the role of insulin-like growth factor I signalling in disease (Diehl et al., 2005). Future work will expand upon these studies in a greater number of patients with more sensitive methods will advance prognosis and treatment options.

1.4.5 Carcinogens in colorectal cancer

Links between colonic and rectal neoplasia and chemical compounds is abundant in the literature (Oddone et al., 2014). Historically, associations between colon cancer and asbestos were widely reported in dockyard workers who were also likely exposed to welding smoke and aromatic hydrocarbons (Puntoni et al., 1977). Higher incidences of colorectal cancer were also found in petrochemical industry workers (Rodu et al., 2001). It is not difficult to imagine a situation in such environments when abundant carcinogens become ingested.

The presence of carcinogens in food and their association to colon cancer is most often reported in terms of red meat, now backed-up by the WHO (http://www.who.int/features/qa/cancer-red-meat/en/). Nitrosamines are also present in smoked fish, dairy products and salted meats. Several other carcinogens in food also have the potential to drive disease, and many natural carcinogens are known to exist in diverse foods, such as tannins, hydrazines, safrole and flavonoids (Koriech, 1994). Heterocyclic amines are a good example of synthetic compounds driving oncogenesis, and 2-amino-1-methyl-6-phenylimidazo[4,5-b]pyridine has been shown to induce colonic tumours in rats independently of p53 involvements (Nagao and Sugimura, 1993).

Thankfully, advances in health and environmental policy and food production have led to fewer of such chemicals being present in our food – although new putative carcinogens are discovered on a regular basis (Unicomb, 2009). Indeed, the synthetic content of our food is likely to increase in future – with lab grown burgers already on the menu (Post, 2012). Continual monitoring of environmental chemicals and atmospheric composition, as well as our practices, will be required to identify new harmful entities and define which compounds are associated with different cancers. Documenting incidence and accurate clinical classification will be required to elucidate new understanding (Baxevanis and Bateman, 2015).

1.4.6 How many mutations does it take, and where?

It has been widely accepted for the past three decades that you require two-hits (Knudson's Two-Hit hypothesis), in TSG to drive cancer. This was supported by the fact that most TSG are usually recessive, and encode loss-of-function mutations, and oncogenes are usually dominant, and encode gain-of-function mutations (Cooper and Hausman, 2007). Activation of one allele (cancer mutation) in a proto-oncogenes is sufficient to drive a dose-response, and both alleles need to be removed in the TSG to see an effect. However, as TSG mutations are recessive and can be passed in the germ-line, a single hit later in life is the second. However, many examples of haplo-insufficiency variation according to genetic background/environmental trigger exist (Rose and Bhattacharya, 2016). Furthermore, many TSGs have related proteins that carry out similar functions in their absence, meaning that loss of one allele may depend on the state of accessory factors.

High-resolution cancer genome sequencing will further our understanding of the genetic bottlenecks and regulatory processes in cancer. However, it is not as if we know less about

cancer genetics now than we did previously. The work of the 1970s-90s revealed a world of conserved mutations (sometimes called driver mutations) and structural genomic events to be associated with diverse forms of cancer. This conservation of cancer elements between people suggests that the genes responsible for cancer are under selective pressure and are highly important to other biological processes; otherwise, these genes would have found regions DNA less susceptible to large-scale DNA changes such as chromosomal losses and translocations.

Small base-pair insertions, substitutions and deletions, amplifications of short DNA stretches, and microsatellite variations have all been found to aid tumour progression, including in colorectal cancer. Any type of mutation has the possibility to affect gene expression, function and/or regulation; demonstrating the many genetic pathways to cancer. Given the somatic heterogeneity observed between different somatic cells across the human body, mutations will also have cell type-specific effects (MacPherson et al., 2004). Considering the burden of reactive oxygen species in our cells, and the amount of carcinogens afflicting us daily, our body does a good job at preventing neoplasia in general. There is a lot of cellular turnover.

The number of the essential cancer cell mutations and the rate, varies between tumours and individuals. However, commonality between patients does exist, as shown by the work of Laura Wood and colleagues, who analysed exonic regions in colorectal tumours, and found that one group of genes were commonly mutated, and another group were also mutated, but at lower frequency (Wood et al., 2007). The authors called this pattern, *mountains and hills*, referring to driver and passenger mutations in the cancer; of which 15 and 60 were described, respectively. When such observations are coupled to the common structural changes described (i.e. *BCR-abl*), it tells us that some mutations are more oncogenic than others; and their frequency is also constrained by genetics and environment.

On top of such base-altering mutations are epigenetic marks, such as cytosine methylation, which adds a functional group to the DNA sequence that can regulate gene expression and DNA replication. Indeed, various methylation marks have been associated with cancer development. For example, hypermethylation of different tumour suppressor genes (or intergenic regions) leads to their inactivation, and cancer development (Kulis et al., 2013). Thus, the identification of endogenous genes and proteins and exogenous factors mediating cancer-associated methylation changes require further exploration; as do other epigenetic marks present in cancers, like the increased deacetylation of histones H2 and H3 that predispose to tumorigenesis. Another, partly epigenetic mechanism involved in tumorigenesis involved microRNA silencing. In humans and other vertebrates, microRNAs are potent regulators of gene expression, and many important regulatory microRNAs are found to be heavily methylated and silenced in cancers – allowing erroneous gene expression to proceed (Garzon et al., 2009).

Importantly, in cells, the effects of many diverse forms of genomic lesions are often self-amplifying and additive. For example, mutations in DNA proof-reading enzymes often lead to mutation accumulation with normal division (with more mutations accumulating with every cellular generation), and accumulating additional mutations in putative oncogenic regions, can

both further a tumour's capacity to rapidly divide. This can be thought of as a chain reaction, in which one error leads to ever more errors, and disease. When cancer cell clonal expansion is considered, the proportion of cells dying due to the accumulation of deleterious variants is unknown. It is likely that along the oncogenic path, the mutation burden can also terminate cancer.

The mutations themselves, however, are not the *be all* and *end all* of tumour development. Other factors are often able to propagate or inhibit cancer progression in tandem with deleterious variants. In many scenarios, mutated cells require the action of growth promoters to cue proliferation. Oestrogens are well known mediators of breast and endometrial cancer in women (Travis and Key, 2003). Progesterone can be used to antagonise the effects of oestrogen clinically. It all depends on how the cancer cell can exploit the environment in which it finds itself. Ovarian cancers might arise more frequently when mutations in endometrial cells lead to constitutive steroid hormone receptor expression, driving cellular growth and proliferation. The gene-environment interaction again rears its head.

1.4.7 The role of infections in colorectal cancer development

The significance of the gene-environment interaction is highlighted by the role of infectious pathogens in carcinogenesis. Indeed, the vast evolutionary pressure put upon anti-pathogen genes during human evolution has been often cited as a genetic arms race (Ingle et al., 2006); where the pathogen alters its genome slightly to evade the pattern recognition receptors of the immunoregulatorysystem of the human body.

The remarkable selective pressure exerted by pathogens is evident in malaria. People with a single mutant copy of the haemoglobin S gene are highly protected from malaria, whilst those with two mutant alleles have severe anaemia, or alpha-thalassaemia (Mockenhaupt et al., 2004). In regions endemic for the parasite, up to 40% of individuals carry the pathogen-protective allele; balancing a foe with our capacity to transport oxygen. If we consider this, pathogens are likely highly important in many tumours, and indeed several viruses, helminths, bacteria, protozoa and plants have been associated with diverse forms of cancer.

We live life surrounded by microorganisms, and the toll and benefits they exert on us is only beginning to be understood, in part by multi-dimensional analyses of the human microbiome. There are more bacterial genes in your intestine than there are in all of your body cells combined – 150-fold (Proctor, 2011). Shen and colleagues have recently shown that the content of adherent bacterial populations in the mucosa of patients with colorectal adenomas is different to healthy controls (Shen et al., 2010). Higher numbers of *Proteobacteria* and lower numbers of *Bacteroides* species were observed in cases. Cases were also found to have a wider diversity of bacterial species present, which may allow for opportunistic pathogens to find a vacant niche.

Helicobacter pylori (a class I carcinogen) represents one such organism, and it has been repeatedly associated with rectal cancer (Burnett-Hartman and Newcomb, 2008). However,

other studies have shown it not to be associated (Moss et al., 1995); again the mechanisms are likely in different individuals. The bacterium is present ubiquitously in human environments, with some estimates suggesting that some Japanese populations are 90% positive. Undoubtedly, we interact strongly with this pathogen globally (The EUROGAST Study Group, 1993).

Helicobacter adheres to gastric epithelial cells and is responsible for stomach ulcers and neoplasia. Here, the bacterium uses the protein CagA to activate human SHP2 and drive continual signalling in the phosphatase – driving cell cycle progression. A meta-analysis of studies carried out between 1991 and 2002 showed a significant association between colorectal cancer (of which rectal cancer accounts for one third) and *Helicobacter* infection (Zumkeller et al., 2006). These results are still disputed, however (Sung et al., 2005). Future epidemiological studies are likely to refine this view.

Numerous studies have also linked *Streptococcus bovis* to intestinal polyp formation and neoplasia (Tjalsma et al., 2012). This bacterium is a normal commensal, which accounts for between 4-8% of the human intestine bacterial population, although its dysregulation could trigger neoplasia, and it has been shown to potently dysregulate inflammation, which will influence the epithelial barrier and integrity. Serology from colorectal cancer patients have shown higher IgG antibody titres against the pathogen compared to controls.

Bacteria aside, viruses have also been associated with colorectal cancer, with JC human polyomavirus being the best example. JCV infection was first implicated as a potential risk factor for colorectal cancer with the observation by Laghi and colleagues (1999) that 96% of the 24 colorectal cancer tissues they examined contained JCV DNA sequences. This virus inhibits p53 via its large T antigen (Staib et al., 1996), reducing TSG function. The large T antigen has also been shown to induce chromosome instability in colorectal cancer cell lines (Ricciardiello et al., 2003). Despite this evidence, the epidemiology does not yet back up the proposition, making further definition of disease aetiology in this respect difficult.

Human papilloma virus represents another well-discussed risk factor for colorectal cancer. Indeed, given its ability to infect genital epithelial cells and cause cervical cancer, much has been defined about the cancer-causing capacity of this family of viruses. Types 16 and 18 are those most commonly associated with cervical cancers (70% of cases) (Muñoz et al., 2003). The viral protein E7 again inhibits p53 and Rb, and induces telomerase, to avoid chromosomal shortening with each replication (Stirdivant et al., 1992); which was shown in cells of the anogenital tract. To date, however, conclusive epidemiology is not available, highlighting the difficulties associated with determining viral (or bacterial) presence, potential pathogenicity, and logistics associated with such studies; especially when dealing with organisms present widely in the environment with many sub-clinical infections. Again, given the genetic background and the incidence of other environmental factors, the ability of a virus to trigger cancer could be affected.

Before concluding this section, I would like the draw the reader's attention to an insightful schematic illustrating the mechanisms in colorectal cancer development and progression (Figure 1.4).

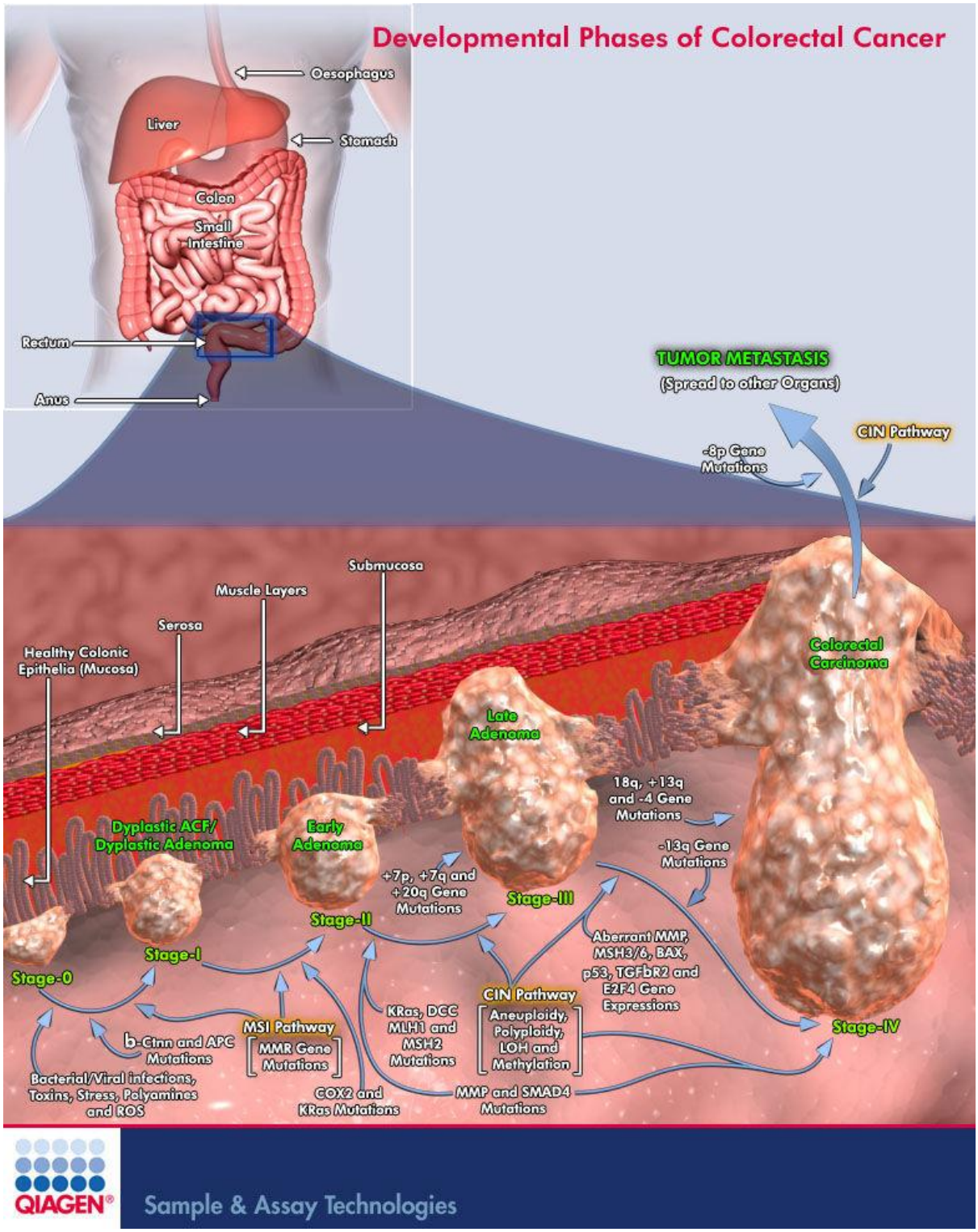


Figure 1-4Molecular mechanisms in colorectal cancer development and progression.

With kind permisswion from Qiagen: Website link:
https://www.qiagen.com/cn/shop/genes-and-pathways/pathway-details/?pwid=133&task=show&action=Accept

1.4.8 The DNA damage response and cancer

As the focus of our work concerns the role of DNA damage response proteins in cancer pathophysiology, we will now review the evidence implicating such processes with cancer.

The specific genes of interest in our investigations were, *MRE11, ATM, RAD50* and *NSB1*. All five are well-characterised members of the DNA damage response and have been implicated in cancer development previously (Lord and Ashworth, 2012). To further explore the role of our candidate factors in rectal cancer, we carried out histological analyses on primary patient tumour samples and defined new expression patterns predictive of radiotherapy response and prognosis. Supporting the importance of the DNA damage response in cancer, and its utility to evaluate treatments and derive prognostic information. Furthermore, as we focus on a limited number of genes, we hope to further refine the capacity of such elements as biomarkers common to diverse patient groups.

For the reader's convenience, a more detailed background to each candidate gene and the evidence for its association with the cancer DNA damage response, will be presented at the start of each results chapter; one chapter per gene. Below, I will introduce the damage response in a broader context.

1.4.8.1 Molecular pathways of the DNA damage response

The first evidence that cancer arose from damaged DNA was provided by researchers working at the The Institutes of Cancer Research in the United Kingdom in the 1950s (Venitt and Phillips, 2012). Until such studies came to light, cancer was thought to arise by carcinogens inhibiting/potentiating the effects of proteins, and not DNA. Brookes and Lawley (1964) showed that poly-aromatic hydrocarbons, such as those found in tobacco smoke, were able to directly bind DNA, and not proteins, changing the genomic code. This was a turning point in cancer research and medicine, also explaining why some cancers ran in families; allowing the hunt for cancer genes to commence.

The DNA damage response is a collective term for the many pathways and factors involved in maintaining genome integrity in the face of normally-occurring errors in DNA replication. To date, more than 450 factors have been identified as being involved in these repairs (Pearl et al., 2015); some of which may be drug-able candidates involved in tumour progression. Figure 1.4 illustrates the complexity of the mechanisms at work. The 2015 Nobel Prize in Chemistry was awarded to Tomas Lindahl, Paul Modrich and Aziz Sancar for their contribution to defining these factors and how they work.

How much DNA repair occurs depends on several factors, with metabolic and other chemical factors and biological processes impinging on the mutation burden. For example, older cells have accumulated more mutations, and may exhibit increased damage repair. Many DNA damage response genes are associated with human lifespan (Browner et al., 2004). When DNA damage response mechanism cannot repair the DNA accurately enough, three things can

happen: cell cycle arrest; apoptosis; cancer. Delineating the mechanisms underlying these fates will likely help develop new cancer treatments.

When DNA damage is sensed by the cell, a variety of things can happen to prevent the cell entering a defective functional mode (due to such errors) and pass on the mutated genes. These mechanisms include halting DNA replication, inducing cell-cycle arrest and/or apoptosis; depending on the extent of the damage encountered (O'Connor, 2015). Given the diversity of factors involved in DNA damage repair, different types of damage (such as single-strand breaks, SSBs, DSBs, and bulky adducts) are repaired by different mechanisms, as illustrated in Figure 1.5 – with redundancy occurring to a differing degree between the pathways.

The most common type of lesion arising in cells are SSBs, which occur approximately 20,000 times per day (Lindahl et al., 1995). These lesions are repaired by the base excision repair pathway (Caldecott, 2014). Base excision repair deals with the majority of small-scale changes to the DNA code, and uses the proteins poly(ADP-ribose) polymerase 1 and 2 to sense DNA damage and signal additional repair proteins to the lesion. The damaged DNA is excised and a newly synthesised template is used to correct the gap.

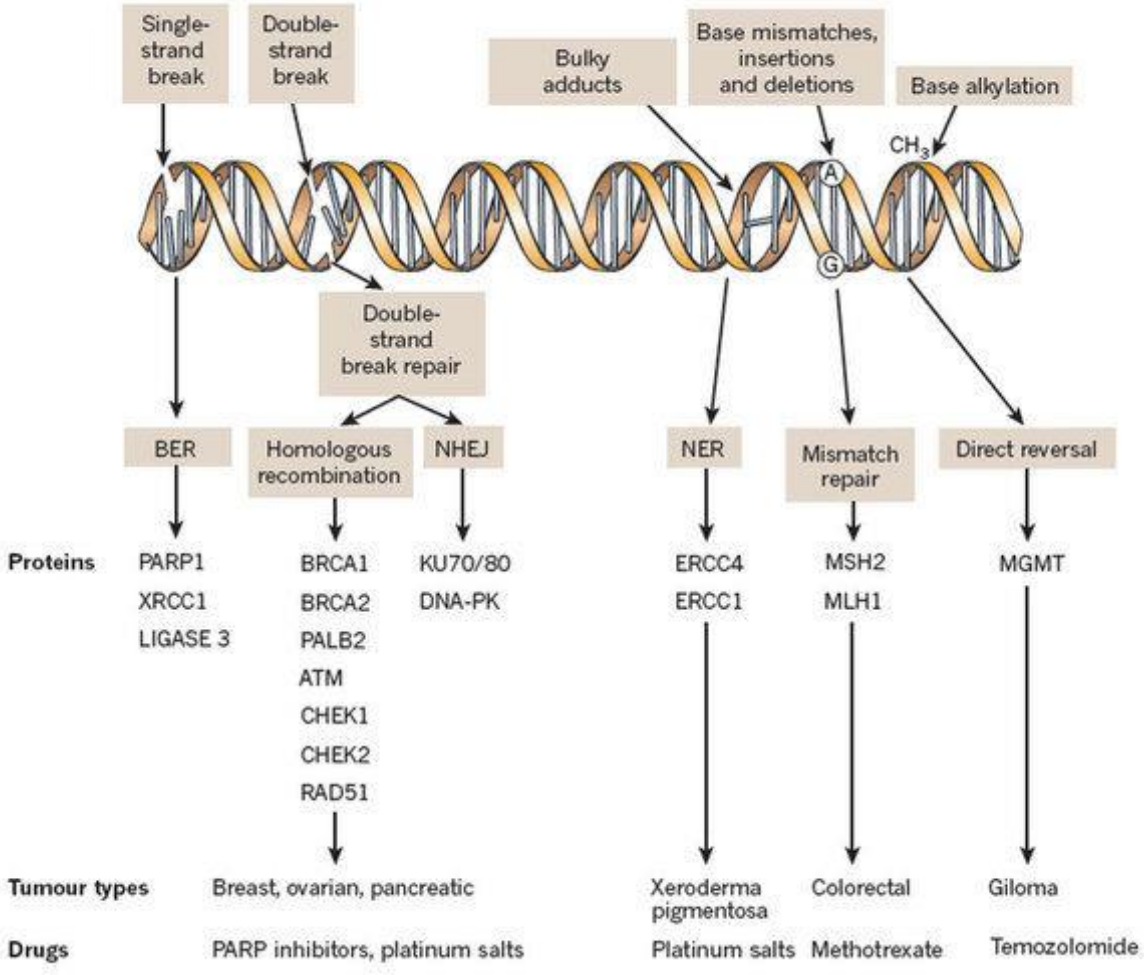


Figure 1-5 The myriad of genes and proteins modulate the DNA damage response and maintain genomic integrity.

In the image taken with permission from a review article by Lord and Ashworth (2012), different proteins involved in repairing different types of DNA damage are shown, although more examples abound. Furthermore, the tumour types usually affected by different DDR pathways, and the drug classes used to treat tumours with these defects, in shown. *BER, base excision repair; NER, nucleotide excision repair; NHEJ, non-homologous end-joining.*

Double-strand breaks, which are more genotoxic, are repaired by either non-homologous end joining (NHEJ), or homologous recombination, although these mechanisms often result in imperfect repair; especially NHEJ, which directly ligates the two ends of the DSB, which often leads to in-frame mutation generation. Mismatch repair (carried out by enzymes such as MSH2 and MLH1) deals with single nucleotide insertions and substitutions, whilst nucleotide excision repair handles nucleotides distorting the three-dimensional structure of the DNA helix (O'Connor, 2015). In general, all repair pathways follow the same three basic steps to repair: (i) DNA damage detection; (ii) accumulation of repair factors at the lesion; (iii) physical lesion repair (Lord and Ashworth, 2012).

It is worth also considering that translesion synthesis and template switching – which allow DNA replication to progress whilst lesions are being repaired – are also part of the DNA damage response (Jansen et al., 2015). Indeed, as DNA replication progresses, template strands can be switched and high-fidelity polymerases transiently replaced with less-stringent ones when correcting the lesion, before switching back (Lord and Ashworth, 2012). Furthermore, when repairing DNA, it is important to note that the stage of the cell cycle at the time will determine part of the success of repair. For example, the absence of sister chromatids during G_1 phase means that lesions occurring during this time rely on NHEJ, a pathway more error-prone than template-dependent homologous recombination. Exogenous factors affecting the DNA damage response, as well as germ-line mutations in any auxiliary factors, will also influence repair success. As discussed, although in many instances single mutation/structural events lead to cancer, cancer can also be viewed as a polygenic disease, where variations between individuals can predispose to pathology.

As discussed, the nature of the base-pair sequence (and its epigenetic state) determines whether cancer develops. Accordingly, mechanisms to maintain high fidelity DNA usage are abundant and widespread in the cell. When DNA is damaged, or mutated, replication fork stalling and adduct formation trigger the cellular DNA damage response. For example, after a DSB, chromatin remodelling carried out by histone modifying enzymes (responding to phosphorylated H2AX) causes relaxation at the target site. JNK then phosphorylates SIRT6, which recruits PARP1 to the damaged DNA (Van Meter et al., 2016). At the lesion, PARP1 produces poly-ADP-ribose sensed by ALC1. This allows the DNA repair enzyme, MRE11, to get to work.

The DNA damage response also needs to maintain the cell in cycle arrest. Key sensors of DNA damage that activate the G_1 restriction check point are the kinases ATM (Ataxia telangiectasia

mutated) and ATR (Ataxia telangiectasia and Rad3 related). These kinases phosphorylate the cell-cycle check point kinases Chk1 and Chk2, which facilitate the degradation of Cdc25A, required to drive exit from G_1. A loss of ATM is highly deleterious, as evidenced by the development of lymphoma in absence of its function (Schaffner et al., 2000). At the same time, Chk1 and Chk2 phosphorylate residues on the TSG p53, thereby stabilising it. Active p53 is then able to work as a transcription factor to drive the expression of genes involved in DNA damage repair and maintaining the cell in a non-replicative state.

Different mechanisms are at work depending on the stage of the cell cycle DNA damage occurs, and the metabolic state of the organism. After DNA has been replicated, during the S phase, the cell undergoes growth (G2). Prior to entering the proliferative mitotic (M) phase, the cell again pauses to survey DNA damage. ATM and ATR are again involved, as are cyclin-Cdk complexes, and is reviewed elsewhere (Branzei and Foiani, 2008).

1.4.8.2 The DNA damage response in colorectal cancer

As the DNA damage is response is central to cancer, much is known about how the function of these pathways differs between malignant and healthy cells. In fact, the expression of DNA repair proteins is used in screens testing potentially carcinogenic compounds in cell lines (Pool-Zobel and Leucht, 1997). Besides what may occur in somatic cells, inherited mutations in the DNA mismatch repair enzymes, such as *MSH2* and *MLH1*, have been associated with increased risk of colorectal cancer development (Gille et al., 2002).

Erroneous mismatch repair is believed to account for the changes in microsatellite length associated with up to 15% of sporadic colorectal malignancies. In hereditary cases, mutations in the mismatch repair enzyme *MLH1* are associated with polyp generation and colorectal cancer; overall, 13% of colorectal cancer suffer mutations in mismatch repair enzymes (Truninger et al., 2005). Whether these variable length repeats contribute to cancer development directly remains to be fully determined, although they do highlight the difficulties faced by colorectal cancers in maintaining genomic integrity and stability (Jiricny, 2006). Markers of these processes, therefore are likely to be highly informative in patient cells when attempting to diagnose cancer or evaluate its response to treatment.

Furthermore, the very early stages of tumorigenesis are typified by the constitutive activation of several DNA damage response proteins, such as ATM, CHK2, H2AX and p53, which is thought to help limit cancer growth (Bartek et al., 2007; Oka et al., 2010). This can be viewed as a protective response, which may be stronger in some individuals over others. In cells progressing to malignant phenotypes, such control barriers have often been depleted or lost. Along these lines, most cancers have been found to suffer from an absence of one or more DNA damage response proteins – which help cancer develop, and places additional strain on various compensatory pathways, as illustrated previously. That said, many cancer do not have observable genetic defects in DNA damage response genes, although they do have links to DNA damage response dysfunction; again, epistasis between genes is different between

tumours and healthy tissue, adding to the complexity of defining factors contributing to individual pathologies (Wang et al., 2014).

In contrasting and interesting recent studies, the essential anti-viral gene, *STING*, was found to be suppressed in colorectal tumours. Loss of this protein led to a reduced DNA damage response as well as a reduction in the cytokine IFN-gamma which recruits tumour-specific T cells and allows for tissue remodelling by other phagocytes of the immune system (Xia et al., 2016). Thus, these results are highly predictive of STING being a controller of oncolytic DNA viruses.

Together these studies demonstrate the importance of the DNA damage response to colorectal cancer, something which we hope to further refine through our investigations here presented.

1.4.8.3 Towards elucidating biomarkers of radiotherapy sensitivity

Before concluding this introductory chapter and describing our experimental approach, I would like to put our studies in a more specific clinical context. Something that we touched upon previously, in the section considering prognostic and clinically-informative biomarkers. In other words, what are we trying to do, why is what we here present important, what new knowledge do we hope to provide, and where can this work be used to benefit patients?

At present, patients with advanced rectal or colorectal cancers receive pre-operative radiotherapy as a first line treatment prior to the surgical resection of tissue (Kye and Cho, 2014). Thus, when the physical and observable tumour burden is reduced during surgery, radiation has already cued apoptosis and necrosis in rapidly-dividing cells in the vicinity of the lesion (Hellevik and Martinez-Zubiaurre, 2014); remembering that rapidly-diving cells, such as tumour cells are more sensitive than quiescent cells to radiotherapy for the various reasons, as discussed. It is also possible that the physical destruction of tumour tissue using radiation leads to a heightened immune response against the tumour (Park et al., 2014). Indeed, conserved anti-tumour T- and B-cell epitopes have been widely reported (Carmi et al., 2015; Han et al., 2014; Li et al., 2016; Shalapour et al., 2015); although the quantitative benefit of an immune response against the tumour *in vivo* is difficult to evaluate and report clinically.

Regardless, radiotherapy aims to downstage the tumour prior to surgery, which can be monitored through histological tumour regression; described in Chapter Two. Post-surgery, radiotherapy is often continued (Glimelius, 2002), given the risk to recurrence. Knowing who responded to pre-operative therapy would help avoid unnecessary treatment, help radiation dose tailoring, and help a patient manage his/her condition (Butow et al., 1997).

In all cases, radiotherapy aims to deliver a dose of radiation appropriate to the size and pathology of the tumour and the patient, noting that healthy cells have their own radio-sensitivity threshold; which can be different between people.

Radiotherapy is especially important in the context of rectal cancer, where confined space within the pelvis places rectal tumours in close proximity to additional viscera that may serve as sites for metastasis (de Wilt et al., 2007). Furthermore, the distal rectum does not benefit from an enveloping free serosa, potentiating metastatic potential. With this in mind, the use of non-invasive radiotherapy (and chemotherapy) becomes more important as a means for curing the disease. Survival times have been found to be lower for rectal cancer patients compared to colonic cancer patients, although this has been disputed (Lee et al., 2013); the importance to both is not contested. As we have discussed, and will do, a number of clinical, research and lifestyle variables could explain such discordancy.

Thus, radiotherapeutic treatments in rectal cancer are widely adopted for the beneficial effects on a patient's overall survival time, as well as for reducing disease recurrence. Indeed, during surgery, the likelihood of spreading tumour cells to new areas within the body (where they can seed new tumours or be transported further afield) increases (Supriya et al., 2008), and pre-operative radiotherapy is thought to hamper metastatic spread by killing escaping clones; remembering that a single dose of radiation post-surgery had no effect on reducing the occurrence of metastasis in malignant mesothelioma patients (Bydder et al., 2004). Regardless of whether chemotherapy is used or not, radiotherapy aims for tumour regression independently. Too much radiation can lead to nausea, incontinence, diarrhoea, thromboembolic disease, and other co-morbidities that reduce a patient's quality of life (Standring, 2008).

To date, although radiotherapy in rectal cancer has been associated with histological regression (Suzuki et al., 2014; Wheeler et al., 2002), the effects on long-term patient survival remains controversial (Colorectal Cancer Collaborative Group, 2001, Glimelius et al., 2003) It is important to remember that the radiation beam is focused upon the primary tumour and does not penetrate draining lymph node or associated tissues that may harbour cancer cells.

Despite the benefits associated with radiotherapy (which can be curative (Gérard et al., 2003)), there is a highly variable response to it between patients (tumour regression in the resected bowel) (Camma et al., 2000), and given the toxicity of ionising radiation it must be used sparingly and efficaciously. For instance, patients can either show a very good level of TRG, or a very low one. Now despite this apparent dichotomy the TRG score itself is not an ideal score and can lead to a lack of conformity and reproducibility (Kim et al., 2016). Being based on a human-derived score on a 1-5 scale, scope for variability between researchers and practitioners exists with regards to TRG scoring. That said, many studies in other diseases do show concordance between TRG and overall survival (Minsky and Rodel, 2014). Again, in the future, mRNA, miRNA and protein signatures from tumour tissue will help build a fuller picture of the tumour and its microenvironment.

Leaving aside the variation between laboratories and practices, the problem is that at present, we do not know which patients stand to benefit from the treatment the most, and which will not respond at all. This is a complex issue. Indeed, response (i.e. tumour regression) to preoperative radiotherapy is plausibly associated with several confounders, such as: age; sex;

disease stage; genetics; lifestyle factors/environment; and management and treatment. But before we get too distracted with the additive effects of our patient's clinical history and genetics, let's consider molecularly how our cells respond to radiation - something touched upon previously - which will help explain our reasons for investigating DNA DSB repair proteins.

When a patient receives ionising radiation, the incident energy generates oxygen-derived free radicals in the exposed cells (Cadet and Wagner, 2013). These radicals exceed the cell's ability to contain them, and damage proteins and nucleic acids (Dröge, 2002). These free radicals lead to DSBs that cue to action of the MRN complex proteins.

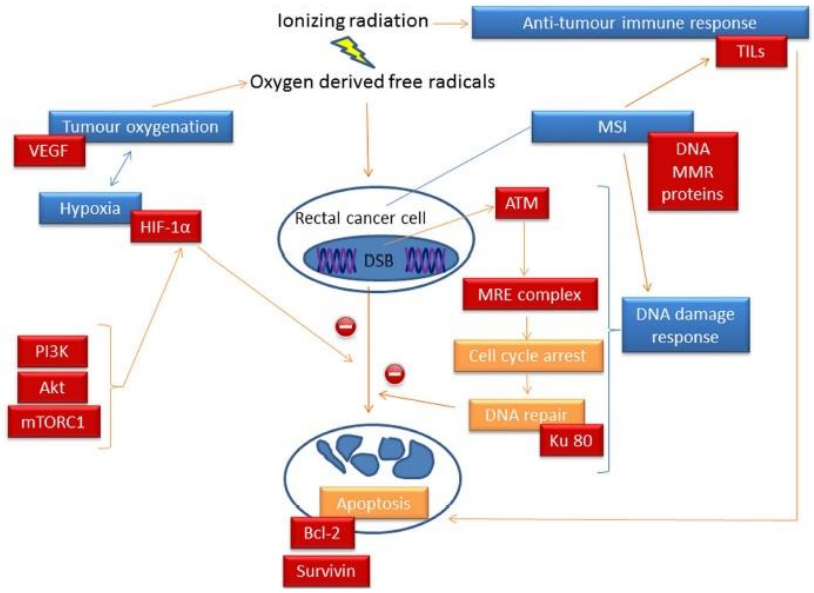
Cancer is often described as a double-edged sword (Hagemann et al., 2007; Tian et al., 2015). As we have seen previously, many tumours carry lesions in DDR proteins – be they germ-line or somatically-acquired defects. These proteins can also be post-translationally impaired in tumours (Oberle and Blattner, 2010). These mechanisms and genomic lesions in DDR proteins allows for cell division and growth, and the expression of diverse phenotypes, to go un-regulated, benefitting tumour growth and functional development. However, when radiation causes DSBs in DNA (affecting a critical gene, say) and it cannot be repaired, the cancer cell will die, while a healthy cell (with intact DSB repair) may fix the problem.

The entire process of radiation-induced cell death is dependent on there being oxygen available in the target tissue (Rockwell et al., 2009), with low levels oxygen and high levels of the hypoxia-sensing HIF-1-alpha transcription factor being associated with a non-response to radiotherapy. In the presence of molecular oxygen and related oxygen species, HIF-1-alpha remains in an inactive state, although when the concentration of its activating species' decline, it remains in an intact, nuclear-localised state, where it results in anti-apoptotic and proliferative gene expression - including AKT, mTOR- and PI3K-regulated pathways (all of which aid the growth of tumour cells) (Courtnay et al., 2015; Hudson et al., 2002; Semenza, 2003).

The centres of tumours, and other regions of the tumour microenvironment (depending on the anatomy of the pathology), have been repeatedly found to be hypoxic (Brown, 2007). Thus, HIF-1-alpha in these environments hampers the success of radiotherapy, and the gene remains an intense focus of investigation in cancer and other fields. To overcome this, radiation is typically delivered in intermittent doses (Kye and Cho, 2014), allowing surface-exposed, non-hypoxic cells to die during the first dose, and healthy tissue time to regenerate before the second. Subsequently, the previously-hidden, previously-hypoxic cells from the tumour centre become surface residents (due to dose one) and increase their oxygen consumption and metabolism (Barker et al., 2015). Given their roles in responding to radiation and oxygen, perhaps HIF-1-alpha, mTOR components and PI3K-pathway components could additionally serve as markers of radiotherapy success.

Following the identification of a double strand break, as occurs following radiotherapy, the DSB repairs proteins of the MRN complex - RAD50, NBS1 and MRE11 – create single-strand

DNA regions with their inherent nuclease (and other) ability, as we will discuss later (Friedberg, 2003). These proteins are essential for detecting and repairing damaged DNA. Subsequent to lesion recognition by the MRN complex, ATM is recruited and cues cell cycle arrest and apoptosis through a variety of mechanisms (Thompson et al., 2005). Thus, we believe these are good candidates to explore with regards to predictive value in rectal cancer. Figure 1.6 comes from our review paper looking at predictive markers of radiotherapy-induced rectal cancer regression and pictorially highlights the key role that these proteins play (Shin et al, 2014).



Potential markers of radiotherapy response in rectal cancer. Within the depicted framework of main processes involved in cellular response to ionising radiation, examples of factors with possible relationship to treatment sensitivity/resistance are highlighted in red. DSB, double strand break; MSI, microsatellite instability; TILs, tumour infiltrating lymphocytes; MMR, mismatch repair; VEGF, vascular endothelial growth factor; HIF, hypoxia-inducible factor; PI3K, phosphatidylinositol-3-kinase; mTOR, mammalian target of rapamycin; ATM, ataxia telangiectasia mutated; MRE, meiotic recombination.

Figure 1-6 Potential markers of radiotherapy response in rectal cancer (Shin et al, 2014)

The biochemistry and mechanisms of action of each of these four proteins (RAD50, NBS1, MRE11, ATM) has started to be introduced, and additional details of how they function as part of the MRN complex is described in the corresponding Results Chapters. Here, it suffices to say that these proteins are integral to maintaining genome stability, and as we will discuss, and many are lethal at an embryonic level if absent in mammals. Variations within these genes is also associated with severe clinical phenotypes; *please see Results*.

Despite their well-understood roles, proving the clinical utility of these proteins in the context of rectal cancer radiotherapy is not straightforward. For example, MRE11 is thought to be responsible for initiating the MRN complex response via its sensing of DNA damage (Kondo et al., 2013), therefore, its absence is expected to correlate with increased radiation sensitivity; something that has been shown in cell lines (Xu et al., 2004). However, its clinical utility in rectal cancer has not been established. As we will discuss, however, MRE11 has been (and

NBS1, ATM, and RAD50) associated with survival and radiotherapy responses in other types of tumours, demonstrating their utility to interrogate the disease pathogenesis.

Demonstrating a predictive role of any of these proteins in the rectal cancer response to radiotherapy would represent an important first step in monitoring and stratifying these patients for therapy.

Beyond proteins that are involved in the DDR, proteins such as VEGF (which mediates tumour angiogenesis and is an important component of carcinogenesis) could also be monitored in response to radiotherapy in rectal cancer. Indeed, a recent study of 62 rectal cancer patients, a complete response to radiotherapy was seen in patient tumours where VEGF expression was low or absent (Zlobec et al., 2005). Furthermore, a meta-analysis by Kuremsky recently suggested 36 potential biomarkers for rectal cancer, based on surveying more than 1,200 peer-reviewed papers on the molecular and clinical nature of the disease. In this report, the genes highlighted included, p53, Ki-67 (involved in cell proliferation), epidermal growth factor receptor (EGFR), BCL-2 (an anti-apoptotic protein), and thymidylate synthase – all key proteins involved in cell growth (Kuremsky et al., 2009). Still, these candidates have not been validated in independent patient cohorts, and many studies included in the meta-analysis were candidate gene studies (and not hypothesis-free genome-wide screens), biasing the output of the meta-analysis.

Thus, although it might seem like a frustrated venture in terms of genetic prognostics and rectal cancer, several important DNA repair proteins are currently used to inform on disease pathology in patients (Peltomaki, 2001). In this case, the markers are used for prognostication and the response to adjuvant therapy. The best-established markers in this context are the DNA mismatch repair proteins, MLH-1, MSH-2, MSH-6, PMS-2, the lack of expression of which defines colorectal cancers with microsatellite instability (MSI) (Arends, 2013), as discussed – variable numbers of tandem repeats are present due to a lack of ability to detect lesions by MMR proteins. This presumably leads to diminished chromosomal integrity and poor mutation repair on the genome-wide scale.

Importantly, colorectal cancers with MSI have their own unique features – such as being relatively poorly differentiated, subjected to an increased immune response, and with a greater tendency to be mucinous – and are not rectal cancers, highlighting the requirement to establish markers for other subtypes of CRC and rectal cancer specifically.

Thus, as radiotherapy remains a mainstay of rectal cancer treatment (Häfner and Debus, 2016), we here attempt to identify markers that are predictive of a successful response to radiotherapy in these patients. As we enter the genomic age (Hua and Bromham, 2017), a focus will be on the stratification of patients into more refined clinical cohorts, where treatment and aetiology are more specific to the individual, so that novel molecular associations can be defined. Importantly, histological assessment, of the nature we here employ, is widely applicable globally to diverse patient cohorts, and benefits from being able to analyse heritage and contemporary tissue samples to learn about disease (De Souza and Greenspan, 2013).

Here, we investigated whether expression of the MRN complex proteins or ATM (in tumour sections), was associated with rectal cancer survival and the response to radiotherapy in a local rectal cancer patient biobank; as well describing correlations found with other clinicohistopathological features.

1.5 Study Aims

Rectal cancers represent approximately one third of all colorectal cancers worldwide, are associated with considerable morbidity and mortality, and their incidence is increasing worldwide.

To date, a large number of molecular mechanisms have been associated with the development and progression of rectal and other colorectal cancers, including aberrant functioning of the DNA damage response, as discussed.

Despite detailed descriptions of many molecular players in cancer, understanding the precise sequence of events causing disease, as well as how and when to intervene in a particular patient, remains a major challenge, as does the ability to prognosticate.

A primary means by which disease mechanism in an individual patient can be understood, and patients stratified, is through the use of biomarkers, which highlight common mechanisms associated with a particular process, i.e. cancer development.

Given the large heritage literature concerning DNA damage response proteins and cancer pathophysiology, we here sought to establish whether any such proteins (and a subset of DNA mismatch repair proteins) were expressed in rectal tumours, and whether their expression was predictive of various clinicohistopathological features, such as: overall survival, disease-free survival, tumour regression, and metastasis. The proteins of interest in our study were: ATM, MRE11, NBS1 and RAD50.

Importantly, although radiotherapy represents a first-line treatment for rectal cancer, highly variable treatment responses have been documented between patients. Therefore, we also set out to investigate whether expression of our candidate proteins – central to repairing damaged DNA generated by radiotherapy – in tumours was associated with patient responses to radiotherapy.

Finally, we also sought to determine whether the expression of these DNA damage response proteins could be considered in combination, within a patient, to increase the predictive power of any putative single biomarker alone.

In order to achieve our aims, we stained (using immunohistochemistry) representative rectal tumour sections from a cohort of 263 rectal cancer specimens collected from the Sydney South West Local Health District (SWSLHD) and analysed protein expression via quantitative means.

1.6 Hypotheses

Given the complex role of DNA damage repair proteins in maintaining genome integrity in the face of double strand breaks (as well as the complex, genetic, oncogenic picture), we hypothesise that several scenarios (and their opposites) are possible:

- Low expression levels of DNA damage response protein expression will be associated with improved radiotherapy responses, and consequently, patient survival; *as a lack of DDR to radiation-induced DSBs will facilitate cancer cell apoptosis.*
- High levels of DNA damage response protein expression in the absence of radiotherapy will be associated with increased disease-free survival
- High/low levels of DDR protein expression will be associated with high/low grade disease subgroups; *high DDR protein levels suggestive of increased genomic lesion burden, as occurs in high-grade rectal cancer*.
- A combination of DDR protein biomarkers (combinatorial panels), will improve predictive power in this setting over single markers alone.

1.7 References

Australian Institute of Health and Welfare & Australasian Association of Cancer Registries. (2017a). Cancer in Australia: in brief 2017. Cancer series no. 102. Cat. no. CAN 101. Canberra: AIHW

Australian Institute of Health and Welfare. (2017b). Australian Cancer Incidence and Mortality (ACIM) books: Colorectal cancer. Canberra. AIHW

Allen, J.I. (1995). Molecular biology of colon polyps and colon cancer. Semin. Surg. Oncol. *11*, 399–405.

Alwis, K.U., de Castro, B.R., Morrow, J.C., and Blount, B.C. (2015). Acrolein exposure in U.S. tobacco smokers and non-tobacco users: NHANES 2005–2006. Environ. Health Perspect. *123*, 1302–1308.

Anand, P., Kunnumakara, A.B., Sundaram, C., Harikumar, K.B., Tharakan, S.T., Lai, O.S., Sung, B., and Aggarwal, B.B. (2008). Cancer is a preventable disease that requires major lifestyle changes. Pharm. Res. *25*, 2097–2116.

Andriole, G.L., Crawford, E.D., Grubb, R.L., Buys, S.S., Chia, D., Church, T.R., Fouad, M.N., Gelmann, E.P., Kvale, P.A., Reding, D.J., et al. (2009). Mortality results from a randomized prostate-cancer screening trial. N. Engl. J. Med. *360*, 1310–1319.

Archibald, J.M. (2015). Endosymbiosis and eukaryotic cell evolution. Curr. Biol. *25*, R911–R921.

Arends, M.J. (2013). Pathways of colorectal carcinogenesis. Appl. Immunohistochem. Mol. Morphol. *21*, 97–102.

Armitage, E.G., and Southam, A.D. (2016). Monitoring cancer prognosis, diagnosis and treatment efficacy using metabolomics and lipidomics. Metabolomics *12*, 146.

Artinyan, A., Soriano, P.A., Prendergast, C., Low, T., Ellenhorn, J.D., and Kim, J. (2008). The anatomic location of pancreatic cancer is a prognostic factor for survival. HPB *10*, 371–376.

Atkin, W.S., Edwards, R., Kralj-Hans, I., Wooldrage, K., Hart, A.R., Northover, J.M. Parkin, D.M., Wardle, J., Duffy, S.W., and Cuzick, J. (2010). Once-only flexible sigmoidoscopy screening in prevention of colorectal cancer: a multicentre randomised controlled trial. Lancet *375*, 1624–1633.

Atta, H.M. (1999). Edwin smith surgical papyrus: The oldest known surgical treatise. Am. Surg. *65*, 1190–1192.

Bae, J.M., Kim, J.H., and Kang, G.H. (2016). Molecular subtypes of colorectal cancer and their clinicopathologic features, with an emphasis on the serrated neoplasia pathway. Arch Pathol Lab Med. *140*, 406–412.

Balkwill, F., and Mantovani, A. (2001). Inflammation and cancer: Back to Virchow? Lancet *357*, 539–545.

Barker, H.E., Paget, J.T.E., Khan, A.A., and Harrington, K.J. (2015). The tumour microenvironment after radiotherapy: Mechanisms of resistance and recurrence. Nat. Rev. Cancer *15*, 409–425.

Barr, T.A., Shen, P., Brown, S., Lampropoulou, V., Roch, T., Lawrie, S., Fan, B., O'Connor, R.A., Anderton, S.M., Bar-Or, A., et al. (2012). B cell depletion therapy ameliorates autoimmune disease through ablation of IL-6–producing B cells. J. Exp. Med. *209*, 1001–1010.

Barros, A., Moreira, L., Santos, H., Ribeiro, N., Carvalho, L., and Santos-Silva, F. (2014).

"Cancer--Educate to Prevent"--high-school teachers, the new promoters of cancer prevention education campaigns. PLoS One *9*, e96672.
Bartek, J., Bartkova, J., and Lukas, J. (2007). DNA damage signalling guards against activated oncogenes and tumour progression. Oncogene *26*, 7773–7779.
Bartkova, J., Hořejší, Z., Koed, K., Krämer, A., Tort, F., Zieger, K., Guldberg, P., Sehested, M., Nesland, J.M., Lukas, C., et al. (2005). DNA damage response as a candidate anti-cancer barrier in early human tumorigenesis. Nature *434*, 864–870.
Baskar, R., Lee, K.A., Yeo, R., and Yeoh, K.W. (2012). Cancer and radiation therapy: current advances and future directions. Int J Med Sci *9*, 193–199.
Baskar, R., Dai, J., Wenlong, N., Yeo, R., and Yeoh, K.-W. (2014). Biological response of cancer cells to radiation treatment. Front. Mol. Biosci. *1*, 24.
Baxevanis, A.D., and Bateman, A. (2015). The importance of biological databases in biological discovery. Curr. Protoc. Bioinforma. 27, 1.1.1-1.1.6.
Baylin, S.B., and Jones, P.A. (2011). A decade of exploring the cancer epigenome — biological and translational implications. Nat. Rev. Cancer *11*, 726–734.
Becher, B., Schlitzer, A., Chen, J., Mair, F., Sumatoh, H.R., Teng, K.W.W., Low, D., Ruedl, C., Riccardi-Castagnoli, P., Poidinger, M., et al. (2014). High-dimensional analysis of the murine myeloid cell system. Nat. Immunol. *15*, 1181–1189.
Bell, E.A., Boehnke, P., Harrison, T.M., and Mao, W.L. (2015). Potentially biogenic carbon preserved in a 4.1 billion-year-old zircon. Proc. Natl. Acad. Sci. *112*, 14518–14521.
Bendall, S.C., Simonds, E.F., Qiu, P., Amir, E.D., Krutzik, P.O., Finck, R., Bruggner, R. V, Melamed, R., Trejo, A., Ornatsky, O.I., et al. (2011). Single-cell mass cytometry of differential immune and drug responses across a human hematopoietic continuum. Science *332*, 687–696.
Berg, J.M., Tymoczko, J.L., and Stryer, L. (2013). Stryer Biochemie.
Berwick, M., Armstrong, B.K., Ben-Porat, L., Fine, J., Kricker, A., Eberle, C., and Barnhill, R. (2005). Sun exposure and mortality from melanoma. J. Natl. Cancer Inst. *97*, 195–199.
Blumenthal, E.Z. (1996). On the origin of cancer. Med. Hypotheses *46*, 581–583.
Boutayeb, A., and Boutayeb, S. (2005). The burden of non communicable diseases in developing countries. Int. J. Equity Health *4*, 2.
Branzei, D., and Foiani, M. (2008). Regulation of DNA repair throughout the cell cycle. Nat Rev Mol Cell Biol *9*, 297–308.
Braunstein, S., and Nakamura, J.L. (2013). Radiotherapy-Induced Malignancies: Review of Clinical Features, Pathobiology, and Evolving Approaches for Mitigating Risk. Front. Oncol. *3*, 73.
Brenner, H., Hoffmeister, M., Stegmaier, C., Brenner, G., Altenhofen, L., and Haug, U. (2007). Risk of progression of advanced adenomas to colorectal cancer by age and sex: estimates based on 840 149 screening colonoscopies. Gut *56*, 1585–1589.
Brodin, P., Jojic, V., Gao, T., Bhattacharya, S., Angel, C.J.L., Furman, D., Shen-Orr, S., Dekker, C.L., Swan, G.E., Butte, A.J., et al. (2015). Variation in the human immune system is largely driven by non-heritable influences. Cell *160*, 37–47.
Brookes, P.,and Lawley, P.D. (1964). Evidence for the binding of polynuclear aromatic hydrocarbons to the nucleic acids of mouse skin: relation between carcinogenic power of hydrocarbons and their binding to deoxynucleic acid. Nature. *202*, 781–784.
Brown, J.M. (2007). Tumor Hypoxia in Cancer Therapy. Methods Enzymol. *435*., 297-321.

Browner, W.S., Kahn, A.J., Ziv, E., Reiner, A.P., Oshima, J., Cawthon, R.M., Hsueh, W.C., and Cummings, S.R. (2004). The genetics of human longevity. Am J Med *117*, 851–860.
Buffery, D. (2015). The 2015 oncology drug pipeline: Innovation drives the race to cure cancer. Am. Heal. Drug Benefits *8*, 216–222.
Burnett-Hartman, A.N., and Newcomb, P. (2008). Infectious agents and colorectal cancer: a review of Helicobacter pylori, Streptococcus bovis, JC virus, and human papillomavirus. Cancer Epidemiol. *216*, 378–388.
Butow, P.N., Maclean, M., Dunn, S.M., Tattersall, M.H.N., and Boyer, M.J. (1997). The dynamics of change: Cancer patients' preferences for information, involvement and support. Ann. Oncol. *8*, 857–863.
Bydder, S., Phillips, M., Joseph, D.J., Cameron, F., Spry, N.A., DeMelker, Y., and Musk, A.W. (2004). A randomised trial of single-dose radiotherapy to prevent procedure tract metastasis by malignant mesothelioma. Br. J. Cancer *91*, 9–10.
Cadet, J., and Wagner, J.R. (2013). DNA base damage by reactive oxygen species, oxidizing agents, and UV radiation. Cold Spring Harb. Perspect. Biol. *5*, 2.
Caldecott, K.W. (2014). DNA single-strand break repair. Exp. Cell Res. *329*, 2–8.
Camma, C., Giunta, M., Fiorica, F., Pagliaro, L., Craxi, A., and Cottone, M. (2000). Preoperative radiotherapy for resectable rectal cancer: A meta-analysis. JAMA *284*, 1008–1015.
Cantor, D. (2007). Uncertain enthusiasm: the American Cancer Society, public education, and the problems of the movie, 1921-1960. Bull. Hist. Med. *81*, 39–69.
Carbone, A., Gloghini, A., and Dotti, G. (2008). EBV-associated lymphoproliferative disorders: classification and treatment. Oncologist *13*, 577–585.
Carmi, Y., Spitzer, M.H., Linde, I.L., Burt, B.M., Prestwood, T.R., Perlman, N., Davidson, M.G., Kenkel, J.A., Segal, E., Pusapati, G. V., et al. (2015). Allogeneic IgG combined with dendritic cell stimuli induce antitumour T-cell immunity. Nature *521*, 99–104.
Carrera-Bastos, P., Fontes, M., O'Keefe, J.H., Lindeberg, S., and Cordain. L. (2011). The western diet and lifestyle and diseases of civilization. Res. Reports Clin. Cardiol. *2*, 15-35.
Casás-Selves, M., and DeGregori, J. (2011). How Cancer Shapes Evolution and How Evolution Shapes Cancer. Evol. Educ. Outreach *4*, 624–634.
Castillo-Fernandez, J.E., Spector, T.D., and Bell, J.T. (2014). Epigenetics of discordant monozygotic twins: implications for disease. Genome Med. *6*, 60.
CCA & ACN (Cancer Council Australia and Australian Cancer Network). (2005). Clinical practice guidelines for the prevention, early detection and management of colorectal cancer. Sydney: CCA & ACN.
Ceballos, G., Ehrlich, P.R., Barnosky, A.D., García, A., Pringle, R.M., and Palmer, T.M. (2015). Accelerated modern human – induced species losses: entering the sixth mass extinction. Sci. Adv. *1*, 1–5.
Chatterjee, A., Mambo, E., and Sidransky, D. (2006). Mitochondrial DNA mutations in human cancer. Oncogene *25*, 4663–4674.
Chen, A.C., Martin, A.J., Choy, B., Fernández-Peñas, P., Dalziell, R.A., McKenzie, C.A., Scolyer, R.A., Dhillon, H.M., Vardy, J.L., Kricker, A., et al. (2015). A Phase 3 Randomized Trial of Nicotinamide for Skin-Cancer Chemoprevention. N. Engl. J. Med. *373*, 1618–1626.
Chen, Y., Argentinis, E., and Weber, G. (2016). IBM Watson: How Cognitive Computing Can

Be Applied to Big Data Challenges in Life Sciences Research. Clin. Ther. *38*, 688–701.
Cho, I., and Blaser, M.J. (2012). The human microbiome: at the interface of health and disease. Nat. Rev. Genet. *13*, 260–270.
Clavarino, A.M., Janda, M., Hughes, K.L., Del Mar, C., Tong, S., Stanton, W.R., Aitken, J.F., Leggett, B.A., and Newman, B. (2004). The view from two sides: A qualitative study of community and medical perspectives on screening for colorectal cancer using FOBT. Prev. Med. (Baltim). *39*, 482–490.
Clayden, J., Greeves, N., and Warren, S.G. (2012). *Organic chemistry*. Oxford, Oxford University Press.Colorectal Cancer Collaborative, G. (2001). Adjuvant radiotherapy for rectal cancer: a systematic overview of 8,507 patients from 22 randomised trials. Lancet *358*, 1291–1304.
Cooper, G.M. (2000). *The Cell: A Molecular Approach. 2nd edition*. Sunderland (MA): Sinauer Associates. The Development and Causes of Cancer. Available from: https://www.ncbi.nlm.nih.gov/books/NBK9963/
Costi, R., Leonardi, F., Zanoni, D., Violi, V., and Roncoroni, L. (2014). Palliative care and end-stage colorectal cancer management: the surgeon meets the oncologist. World J Gastroenterol. *20*, 7602–7621.
Courtnay, R., Ngo, D.C., Malik, N., Ververis, K., Tortorella, S.M., and Karagiannis, T.C. (2015). Cancer metabolism and the Warburg effect: the role of HIF-1 and PI3K. Mol. Biol. Rep. *42*, 841–851.
Coussens, L.M., and Werb, Z. (2002). Inflammation and cancer. Nature *420*, 860–867.
Crosbie, E.J., Einstein, M.H., Franceschi, S., and Kitchener, H.C. (2013). Human papillomavirus and cervical cancer. Lancet *382*, 889–899.
Crown, J., and O'Leary, M. (2000). The taxanes: an update. Lancet *355*, 1176–1178.
Cucina, A., Dinicola, S., Coluccia, P., Proietti, S., D'Anselmi, F., Pasqualato, A., and Bizzarri, M. (2012). Nicotine stimulates proliferation and inhibits apoptosis in colon cancer cell lines through activation of survival pathways. J. Surg. Res. *178*, 233–241.
Curtis, C., Shah, S.P., Chin, S.F., Turashvili, G., Rueda, O.M., Dunning, M.J., Speed, D., Lynch, A.G., Samarajiwa, S., Yuan, Y., et al. (2012). The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. Nature *486*, 346–352.
Czabotar, P.E., Lessene, G., Strasser, A., and Adams, J.M. (2013). Control of apoptosis by the BCL-2 protein family: implications for physiology and therapy. Nat. Rev. Mol. Cell Biol. *15*, 49–63.
Dang, C. V. (2012). MYC on the path to cancer. Cell *149*, 22–35.
Darby, S., McGale, P., Correa, C., Taylor, C., Arriagada, R., Clarke, M., Cutter, D., Davies, C., Ewertz, M., Godwin, J., et al. (2011). Effect of radiotherapy after breast-conserving surgery on 10-year recurrence and 15-year breast cancer death: Meta-analysis of individual patient data for 10 801 women in 17 randomised trials. Lancet *378*, 1707–1716.
De Flora, S., and La Maestra, S. (2014). Epidemiology of cancers of infectious origin and prevention strategies. J. Prev. Med. Hyg. *56*, E15–E20.
DeGregori, J. (2013). Challenging the axiom: does the occurrence of oncogenic mutations truly limit cancer development with age? Oncogene *32*, 1869–1875.
Delellis, R.A., and Xia, L. (2003). Paraneoplastic Endocrine Syndromes: A Review*. Endocr. Pathol. *14*, 303–317.

Denard, B., Lee, C., and Ye, J. (2012). Doxorubicin blocks proliferation of cancer cells through proteolytic activation of CREB3L1. Elife. *1*, e00090.
De Souza, Y.G., and Greenspan, J.S. (2013). Biobanking past, present and future: responsibilities and benefits. AIDS *27*, 303–312.
de Wilt, J.H.W., Vermaas, M., Ferenschild, F.T.J., and Verhoef, C. (2007). Management of locally advanced primary and recurrent rectal cancer. Clin. Colon Rectal Surg. *20*, 255–263.
Diaz, L.A., and Bardelli, A. (2014). Liquid biopsies: Genotyping circulating tumor DNA. J. Clin. Oncol. *32*, 579–586.
Diehl, D., Lahm, H., Wolf, E., and Bauersachs, S. (2005). Transcriptome analysis of a human colorectal cancer cell line shows molecular targets of insulin-like growth factor-binding protein-4 overexpression. Int J Cancer *113*, 588–599.
Donlan, R.M. (2002). Biofilms: Microbial life on surfaces. Emerg. Infect. Dis. *8*, 881–890.
Dopico, X.C., Evangelou, M., Ferreira, R.C., Guo, H., Pekalski, M.L., Smyth, D.J., Cooper, N., Burren, O.S., Fulford, A.J., Hennig, B.J., et al. (2015). Widespread seasonal gene expression reveals annual differences in human immunity and physiology. Nat. Commun. *6*, 7000.
Dorudi, S., Steele, R.J., and McArdle, C.S. (2002). Surgery for colorectal cancer. Br Med Bull. *64*, 101–118.
Douillard, J.Y., Cunningham, D., Roth, A., Navarro, M., James, R., Karasek, P., Jandik, P., Iveson, T., Carmichael, J., Alakl, M., et al. (2000). Irinotecan combined with fluorouracil compared with fluorouracil alone as first-line treatment for metastatic colorectal cancer: a multicentre randomised trial. Lancet *355*, 1041–1047.
Dröge, W. (2002). Free Radicals in the Physiological Control of Cell Function. Physiol. Rev. *82*, 47–95.
Duek SD, Issa N, Hershko DD, et al. (2008). Outcome of transanal endoscopic microsurgery and adjuvant radiotherapy in patients with T2 rectal cancer. Dis Colon Rectum. *51*, 379-84.
Duncan, E., Brown, M., and Shore, E.M. (2014). The revolution in human monogenic disease mapping. Genes (Basel). *5*, 792–803.
Eaden, J.A., Abrams, K.R., and Mayberry, J.F. (2001). The risk of colorectal cancer in ulcerative colitis: a meta-analysis. Gut *48*, 526–535.
Easton, D.F., and Eeles, R.A. (2008). Genome-wide association studies in cancer. Hum. Mol. Genet. *17*, 109-115.
Ellis, P.M., and Vandermeer, R. (2011). Delays in the diagnosis of lung cancer. J. Thorac. Dis. *3*, 183–188.
Emanuel, E.J., Schnipper, L.E., Kamin, D.Y., Levinson, J., and Lichter, A.S. (2003). The costs of conducting clinical research. J. Clin. Oncol. *21*, 4145–4150.
Endreseth, B.H,, Myrvold, H.E,, Romundstad, P., Hestvik, U.E., Bjerkeset, T., Wibe, A., et al. (2005). Transanal excision vs. major surgery for T1 rectal cancer. *Dis Colon Rectum*. *48*, 1380–1388.
Esposito, K., Marfella, R., Ciotola, M., Di Palo, C., Giugliano, F., Giugliano, G., D'Armiento, M., D'Andrea, F., and Giugliano, D. (2004). Effect of a Mediterranean-Style Diet on Endothelial Dysfunction and Markers of Vascular Inflammation in the Metabolic Syndrome. JAMA *292*, 1440-1446.
EUROGAST Study Group (1993). An international association between Helicobacter pylori

infection and gastric cancer. Lancet *341*, 1359–1363.
Eyler, C.E., and Rich, J.N. (2008). Survival of the fittest: cancer stem cells in therapeutic resistance and angiogenesis. J Clin Oncol *26*, 2839–2845.
Fernandez-Medarde, A., and Santos, E. (2011). Ras in Cancer and Developmental Diseases. Genes Cancer *2*, 344–358.
Festino, L., Botti, G., Lorigan, P., Masucci, G. V., Hipp, J.D., Horak, C.E., Melero, I., and Ascierto, P.A. (2016). Cancer Treatment with Anti-PD-1/PD-L1 Agents: Is PD-L1 Expression a Biomarker for Patient Selection? Drugs *76*, 925–945.
Fleshman, J., Branda, M., Sargent, D.J., Boller, A.M., George, V., Abbas, M., Peters, W.R., Maun, D., Chang, G., Herline, A., et al. (2015). Effect of Laparoscopic-Assisted Resection vs Open Resection of Stage II or III Rectal Cancer on Pathologic Outcomes. JAMA *314*, 1346.
Ford, D., Easton, D.F., Stratton, M., Narod, S., Goldgar, D., Devilee, P., Bishop, D.T., Weber, B., Lenoir, G., Chang-Claude, J., et al. (1998). Genetic Heterogeneity and Penetrance Analysis of the BRCA1 and BRCA2 Genes in Breast Cancer Families. Am. J. Hum. Genet. *62*, 676–689.
Frangioni, J. V. (2008). New technologies for human cancer imaging. J. Clin. Oncol. *26*, 4012–4021.
Frayling TM, Timpson NJ, Weedon MN, et al. (2007). A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity. Science, *316*, 889–894.
Freeman, H.-J. (2008). Colorectal cancer risk in Crohn's disease. World J. Gastroenterol. *14*, 1810-1811.
Friedberg, E.C. (2003). DNA damage and repair. Nature *421*, 436–440.
Fu, D., Calvo, J.A., and Samson, L.D. (2012). Balancing repair and tolerance of DNA damage caused by alkylating agents. Nat. Rev. Cancer. *12*, 104-120.
Galiatsatos, P., and Foulkes, W.D. (2006). Familial adenomatous polyposis. Am. J. Gastroenterol. *101*, 385–398.
Garber, J.E., and Offit, K. (2005). Hereditary cancer predisposition syndromes. J. Clin. Oncol. *23*, 276–292.
Garzon, R., Calin, G.A., and Croce, C.M. (2009). MicroRNAs in Cancer. Annu. Rev. Med. *60*, 167–179.
Gattinoni, L., Powell, D.J., Rosenberg, S.A., and Restifo, N.P. (2006). Adoptive immunotherapy for cancer: building on success. Nat. Rev. Immunol. *6*, 383–393.
Gérard, J.P., Romestaing, P., Baulieux, J., Benchimol, and Gerard, J.P. (2003). Local curative treatment of rectal cancer by radiotherapy alone. Color. Dis *5*, 442–444.
Gille, J.J.P., Hogervorst, F.B.L., Pals, G., Wijnen, Jt., van Schooten, R.J., Dommering, C.J., Meijer, G.A., Craanen, M.E., Nederlof, P.M., de Jong, D., et al. (2002). Genomic deletions of MSH2 and MLH1 in colorectal cancer families detected by a novel mutation detection approach. Br. J. Cancer *87*, 892–897.
Ginsburg, G.S., and Willard, H.F. (2009). Genomic and personalized medicine: foundations and applications. Transl. Res. *154*, 277–287.
Glimelius, B., and Isacsson, U. (2001). Preoperative radiotherapy for rectal cancer -- is 5x5 Gy a good or bad schedule?. Acta Oncol *40*, 958-967.
Glimelius, B.L.G. (2002). The role of preoperative and postoperative radiotherapy in rectal

cancer. Clin. Colorectal Cancer *2*, 82–92.
Glimelius, B., Grönberg, H., Järhult, J., Wallgren, A., and Cavallin-Ståhl, E. (2003). A systematic overview of radiation therapy effects in rectal cancer. Acta Oncol. (Madr). *42*, 476–492.
Goldstein, D.A., Stemmer, S.M., and Gordon, N. (2016). The cost and value of cancer drugs - are new innovations outpacing our ability to pay? Isr. J. Health Policy Res. *5*, 40.
Grávalos, C., Cassinello, J., Fernández-Rañada, I., and Holgado, E. (2007). Role of tyrosine kinase inhibitors in the treatment of advanced colorectal cancer. Clin. Colorectal Cancer *6*, 691–699.
Gray, R. (2013). Children could have DNA tested at birth. Telegr. http://www.telegraph.co.uk/news/health/news/105017.
Greaves, M., and Maley, C.C. (2012). Clonal evolution in cancer. Nature *481*, 306–313.
Gribbin, J. (1999). Panspermia revisited. The Observatory. *119*, 284-285.
Grivennikov, S., and Karin, M. (2008). Autocrine IL-6 Signaling: A Key Event in Tumorigenesis? Cancer Cell *13*, 7–9.
Growney, J.D., Clark, J.J., Adelsperger, J., Stone, R., Fabbro, D., Griffin, J.D., and Gilliland, D.G. (2005). Activation mutations of human c-KIT resistant to imatinib mesylate are sensitive to the tyrosine kinase inhibitor PKC412. Blood *106*, 721–724.
Guedez, L., and Stetler-Stevenson, W.G. (2010). The prognostic value of TIMP-1 in multiple myeloma. Leuk. Res. *34*, 576–577.
Guo, G., Li, X., Wang, R., Yu, J., Ye, M., Mao, L., Zhang, S., and Zheng, S. (2016). Association between Oxidative DNA Damage and Risk of Colorectal Cancer: Sensitive Determination of Urinary 8-Hydroxy-2′-deoxyguanosine by UPLC-MS/MS Analysis. Sci Rep *6*, 32581.
Gyllensten, U., Gustavsson, I., Lindell, M., and Wilander, E. (2012). Primary high-risk HPV screening for cervical cancer in post-menopausal women. Gynecol. Oncol. *125*, 343–345.
Häfner, M.F., and Debus, J. (2016). Radiotherapy for colorectal cancer: Current standards and future perspectives. Visc. Med. *32*, 172–177.
Hagemann, T., Balkwill, F., and Lawrence, T. (2007). Inflammation and Cancer: A Double-Edged Sword. Cancer Cell *12*, 300–301.
Han, A., Newell, E.W., Glanville, J., Fernandez-Becker, N., Khosla, C., Chien, Y.-H., and Davis, M.M. (2013). Dietary gluten triggers concomitant activation of CD4+ and CD8+ αβ T cells and γδ T cells in celiac disease. Proc. Natl. Acad. Sci. U. S. A. *110*, 13073–13078.
Han, A., Glanville, J., Hansmann, L., and Davis, M.M. (2014). Linking T-cell receptor sequence to functional phenotype at the single-cell level. Nat. Biotechnol. *32*, 684–692.
Hande, K.R. (2008). Topoisomerase II inhibitors. Update Cancer Ther. *3*, 13–26.
Hellevik, T., and Martinez-Zubiaurre, I. (2014). Radiotherapy and the Tumor Stroma: The Importance of Dose and Fractionation. Front. Oncol. *4*, 1.
Herbst, R.S., Heymach, J. V., and Lippman, S.M. (2008). Lung cancer. N. Engl. J. Med. *359*, 1367–1380.
Hoelder, S., Clarke, P.A., and Workman, P. (2012). Discovery of small molecule cancer drugs: Successes, challenges and opportunities. Mol. Oncol. *6*, 155–176.
Hoffmann, C., Sabranski, M., and Esser, S. (2017). HIV-Associated Kaposi's Sarcoma. Oncol. Res. Treat. *40*, 94–98.

Hogerzeil, H. V (2004). The concept of essential medicines: lessons for rich countries. BMJ Br. Med. J. *329*, 1169–1172.
Hörst, S.M., Yelle, R.V., Buch, A., Carrasco, N., Cernogora, G., Dutuit, O., Quirico, E., Sciamma-O'Brien, E., Smith, M.A., Somogyi, Á., et al. (2012). Formation of Amino Acids and Nucleotide Bases in a Titan Atmosphere Simulation Experiment. Astrobiology *12*, 809–817.
Hua, X., and Bromham, L. (2017). Darwinism for the genomic age: Connecting mutation to diversification. Front. Genet. *8*, 12.
Huang, P., Westmoreland, S. V., Jain, R.K., and Fukumura, D. (2011). Spontaneous nonthymic tumors in SCID mice. Comp. Med. *61*, 227–234.
Hudson, C.C., Liu, M., Chiang, G.G., Otterness, D.M., Loomis, D.C., Kaper, F., Giaccia, A.J., and Abraham, R.T. (2002). Regulation of hypoxia-inducible factor 1alpha expression and function by the mammalian target of rapamycin. Mol. Cell. Biol. *22*, 7004–7014.
Huffman, M. A. (1997). Current evidence for self-medication in primates: A multidisciplinary perspective. Am. J. Phys. Anthropol. *104*, 171–200.
ICRP (2007). ICRP 103: The 2007 Recommendations of the International Commission on Radiological Protection. Ann. ICRP *37*, 330.
Imperiale, T.F., Ransohoff, D.F., Itzkowitz, S.H., Turnbull, B.A., and Ross, M.E. (2004). Fecal DNA versus Fecal Occult Blood for Colorectal-Cancer Screening in an Average-Risk Population. N. Engl. J. Med. *351*, 2704–2714.
Ingle, R.A., Carstens, M., and Denby, K.J. (2006). PAMP recognition and the plant-pathogen arms race. BioEssays *28*, 880–889.
Isella, C., Terrasi, A., Bellomo, S.E., Petti, C., Galatola, G., Muratore, A., Mellano, A., Senetta, R., Cassenti, A., Sonetto, C., et al. (2015). Stromal contribution to the colorectal cancer transcriptome. Nat. Genet. *47*, 312–319.
James, J., Thomas, P., and Kerr, D. (2007). Preventing childhood obesity: two year follow-up results from the Christchurch obesity prevention programme in schools (CHOPPS). BMJ *335*, 762.
Jan, S., Kimman, M., Kingston, D., and Woodward, M. (2012). The socioeconomic burden of cancer in member countries of the Association of Southeast Asian Nations (ASEAN)--stakeholder meeting report. Asian Pac. J. Cancer Prev. *13*, 407–409.
Jansen, J.G., Tsaalbi-Shtylik, A., and de Wind, N. (2015). Roles of mutagenic translesion synthesis in mammalian genome stability, health and disease. DNA Repair (Amst). *29*, 56–64.
Jemal, A., Bray, F., and Ferlay, J. (2011). Global Cancer Statistics: 2011. CA Cancer J Clin *49*, 1,33-64.
Jiricny, J. (2006). The multifaceted mismatch-repair system. Nat. Rev. Mol. Cell Biol. *7*, 335–346.
Jocham, H.R., Dassen, T., Widdershoven, G., and Halfens, R. (2006). Quality of life in palliative care cancer patients: A literature review. J. Clin. Nurs. *15*, 1188–1195.
Jögi, A., Vaapil, M., Johansson, M., and Påhlman, S. (2012). Cancer cell differentiation heterogeneity and aggressive behavior in solid tumors. Ups. J. Med. Sci. *117*, 217–224.
Johnson, D.A., Barclay, R.L., Mergener, K., Weiss, G., König, T., Beck, J., and Potter, N.T. (2014). Plasma Septin9 versus fecal immunochemical testing for colorectal cancer screening: A prospective multicenter study. PLoS One *9*, e98238.

Joyce, J.A., and Fearon, D. (2015). T cell exclusion, immune privilege, and the tumor microenvironment. Science.*348*, 74–80.
Kalos, M., and June, C.H. (2013). Adoptive T Cell Transfer for Cancer Immunotherapy in the Era of Synthetic Biology. Immunity *39*, 49–60.
Kanavos, P. (2006). The rising burden of cancer in the developing world. Ann Oncol. *17(Suppl 8)*: viii15–viii23.
Kang, G.H. (2011). Four molecular subtypes of colorectal cancer and their precursor lesions. Arch. Pathol. Lab. Med. *135*, 698–703.
Karadag, F., Kirdar, S., Karul, A.B., and Ceylan, E. (2008). The value of C-reactive protein as a marker of systemic inflammation in stable chronic obstructive pulmonary disease. Eur. J. Intern. Med. *19*, 104–108.
Karamanou, M., Androutsos, G., and Diamantis, A. (2009). Cancer and Roman medicine. J. B.U.ON. *14*, 537–540.
Karanicolas, P.J., Metrakos, P., Chan, K., Asmis, T., Chen, E., Kingham, T.P., Kemeny, N., Porter, G., Fields, R.C., Pingpank, J., et al. (2014). Hepatic arterial infusion pump chemotherapy in the management of colorectal liver metastases: expert consensus statement. Curr. Oncol. *21*, e129-36.
Kaye, J., Boddington, P., de Vries, J., Hawkins, N., and Melham, K. (2010). Ethical implications of the use of whole genome methods in medical research. Eur J Hum Genet *18*, 398–403.
Kelly, C., and Cassidy, J. (2007). Chemotherapy in metastatic colorectal cancer. Surg. Oncol. *16*, 65–70.
Kim, E.R., and Chang, D.K. (2014). Colorectal cancer in inflammatory bowel disease: The risk, pathogenesis, prevention and diagnosis. World J. Gastroenterol. *20*, 9872–9881.
Kim, S.H., Chang, H.J., Kim, D.Y., Park, J.W., Baek, J.Y., Kim, S.Y., Park, S.C., Oh, J.H., Yu, A., and Nam, B.H. (2016). What is the ideal tumor regression grading system in rectal cancer patients after preoperative chemoradiotherapy? Cancer Res. Treat. *48*, 998–1009.
Kondo, T., Kobayashi, J., Saitoh, T., Maruyama, K., Ishii, K.J., Barber, G.N., Komatsu, K., Akira, S., and Kawai, T. (2013). DNA damage sensor MRE11 recognizes cytosolic double-stranded DNA and induces type I interferon by regulating STING trafficking. Proc. Natl. Acad. Sci. *110*, 2969–2974.
Koriech, O. (1994). Diet and Cancer. J Fam. Community Med *1*, 2–11.
Kosaka, P.M., Pini, V., Ruz, J.J., da Silva, R.A., González, M.U., Ramos, D., Calleja, M., and Tamayo, J. (2014). Detection of cancer biomarkers in serum using a hybrid mechanical and optoplasmonic nanosensor. Nat. Nanotechnol. *9*, 1047–1053.
Kraman, M., Bambrough, P.J., Arnold, J.N., Roberts, E.W., Magiera, L., Jones, J.O., Gopinathan, A., Tuveson, D.A., and Fearon, D.T. (2010). Suppression of antitumor immunity by stromal cells expressing fibroblast activation protein-alpha. Science *330*, 827–830.
Kranz, L.M., Diken, M., Haas, H., Kreiter, S., Loquai, C., Reuter, K.C., Meng, M., Fritz, D., Vascotto, F., Hefesha, H., et al. (2016). Systemic RNA delivery to dendritic cells exploits antiviral defence for cancer immunotherapy. Nature *534*, 396–401.
Kulis, M., Queirós, A.C., Beekman, R., and Martín-Subero, J.I. (2013). Intragenic DNA methylation in transcriptional regulation, normal differentiation and cancer. Biochim. Biophys. Acta - Gene Regul. Mech. *1829*, 1161–1174.

Kuremsky, J.G., Tepper, J.E., and McLeod, H.L. (2009). Biomarkers for response to neoadjuvant chemoradiation for rectal cancer. Int J Radiat Oncol Biol Phys *74*, 673–688.
Kye, B., and Cho, H. (2014). Overview of radiation therapy for treating rectal cancer. Ann. Coloproctol. *30*, 165–166.
Lachenmeier, D.W., Kanteres, F., and Rehm, J. (2009). Carcinogenicity of acetaldehyde in alcoholic beverages: Risk assessment outside ethanol metabolism. Addiction *104*, 533–550.
Laghi, L., Randolph, A.E., Chauhan, D.P., et al. (1999). JC virus DNA is present in the mucosa of the human colon and in colorectal cancers. Proc Natl Acad Sci U S A. 96, 7484–7489.
Lee, Y.C., Lee, Y.L., Chuang, J.P., and Lee, J.C. (2013). Differences in survival between colon and rectal cancer from SEER data. PLoS One. *8*, e78709.
Li, F., and Lai, M. (2009). Colorectal cancer, one entity or three. J. Zhejiang Univ. Sci. B *10*, 219–229.
Li, X., Sundquist, J., and Sundquist, K, (2009), Risks of rheumatic diseases in first- and second-generation immigrants in Sweden A Nationwide Followup Study. Arthritis Rheum *60*, 1588–1596.
Li, B., Li, T., Pignon, J.C., Wang, B., Wang, J., Shukla, S.A., Dou, R., Chen, Q., Hodi, F.S., Choueiri, T.K., et al. (2016). Landscape of tumor-infiltrating T cell repertoire of human cancers. Nat. Genet. *48*, 725–732.
Li, Y., Innocentin, S., Withers, D.R., Roberts, N.A., Gallagher, A.R., Grigorieva, E.F., Wilhelm, C., and Veldhoen, M. (2011). Exogenous stimuli maintain intraepithelial lymphocytes via aryl hydrocarbon receptor activation. Cell *147*, 629–640.
Lindahl, T., Satoh, M.S., Poirier, G.G., and Klungland, A. (1995). Post-translational modification of poly(ADP-ribose) polymerase induced by DNA strand breaks. Trends Biochem. Sci. *20*, 405–411.
Lopez, A.D., Mathers, C.D., Ezzati, M., Jamison, D.T., and Murray, C.J. (2006). Global and regional burden of disease and risk factors, 2001: systematic analysis of population health data. Lancet *367*, 1747–1757.
Lord, C.J., and Ashworth, A. (2012). The DNA damage response and cancer therapy. Nature *481*, 287–294.
Lung cancer guide book (2013). Chemotherapy side effects profile. Lung Cancer Guid. B. 384–395.
Lutz, W., and Nowakowska-Swirta, E. (2002). Gene p53 mutations, protein p53, and anti-p53 antibodies as biomarkers of cancer process. Int. J. Occup. Med. Environ. Health *15*, 209–218.
Lutz, S.T., Jones, J., and Chow, E. (2014). Role of radiation therapy in palliative care of the patient with cancer. J. Clin. Oncol. *32*, 2913–2919.
Lynch D., and Murphy, A. (2016). The emerging role of immunotherapy in colorectal cancer. Ann. Transl. Med. *4*, 305.
MacEwen, E.G. (1990). Spontaneous tumors in dogs and cats: Models for the study of cancer biology and treatment. Cancer Metastasis Rev. *9*, 125–136.
MacPherson, D., Sage, J., Kim, T., Ho, D., McLaughlin, M.E., and Jacks, T. (2004). Cell type-specific effects of Rb deletion in the murine retina. Genes Dev. *18*, 1681–1694.
Maecker, H.T., McCoy, J.P., and Nussenblatt, R. (2012). Standardizing immunophenotyping for the Human Immunology Project. Nat Rev Immunol *12*, 191–200.
Maher, K.A., and Stevenson, D.J. (1988). Impact frustration of the origin of life. Nature *331*,

612–614.
Mali, P., Yang, L., Esvelt, K.M., Aach, J., Guell, M., DiCarlo, J.E., Norville, J.E., Church, G.M., Wiedenheft, B., Sternberg, S.H., et al. (2013). RNA-guided human genome engineering via Cas9. Science *339*, 823–826.
Mariano, C. (2007). Holistic nursing as a specialty: holistic nursing - scope and standards of practice. Nurs. Clin. North Am. *42*, 165–88.
Markman, B., Rodríguez-Freixinos, V., and Tabernero, J. (2010). Biomarkers in colorectal cancer. Clin. Transl. Oncol. *12*, 261–270.
Marley, A.R., and Nan, H. (2016). Epidemiology of colorectal cancer. Int. J. Mol. Epidemiol. Genet. *7*, 105–114.
Martin, W., Sousa, F., and Lane, N. (2014). Energy at life's origin. Science. *344*, 1092–1093.
Mikkelsen, T., and Thorn, C. (2011). Methotrexate pathway. Pharmacogenet Genomics *21*, 679–686.
Miller, E.C., and Miller, J.A. (1981). Mechanisms of chemical carcinogenesis. Cancer *47*, 1055–1064.
Minsky, B.D., and Rodel, C. (2014). Identifying the most predictive post-chemoradiation TRG system for rectal cancer. J. Natl. Cancer Inst. *106*, 10.
Mockenhaupt, F.P., Ehrhardt, S., Gellert, S., Otchwemah, R.N., Dietz, E., Anemana, S.D., and Bienzle, U. (2004). Alpha(+)-thalassemia protects African children from severe malaria. Blood *104*, 2003–2006.
Moss, S.F., Neugut, A.I., Garbowski, G.C., Wang, S., Treat, M.R., and Forde, K.A. (1995). Helicobacter pylori seroprevalence and colorectal neoplasia: Evidence against an association. J. Natl. Cancer Inst. *87*, 762–763.
Moulder, J.W. (1985). Comparative biology of intracellular parasitism. Microbiol Rev. *49*, 298-337.
Munger, K., Baldwin, A., Edwards, K.M., Hayakawa, H., Nguyen, C.L., Owens, M., Grace, M., and Huh, K. (2004). Mechanisms of Human Papillomavirus-Induced Oncogenesis. J. Virol. *78*, 11451–11460.
Muñoz, N., Bosch, F.X., de Sanjosé, S., Herrero, R., Castellsagué, X., Shah, K. V., Snijders, P.J.F., and Meijer, C.J.L.M. (2003). Epidemiologic Classification of Human Papillomavirus Types Associated with Cervical Cancer. N. Engl. J. Med. *348*, 518–527.
Murchison, E.P. (2008). Clonally transmissible cancers in dogs and Tasmanian devils. Oncogene *27*, S19–S30.
Nagao, M., and Sugimura, T. (1993). Carcinogenic factors in food with relevance to colon cancer development. Mutat. Res. - Fundam. Mol. Mech. Mutagen. *290*, 43–51.
Navari, R.M., Qin, R., Ruddy, K.J., Liu, H., Powell, S.F., Bajaj, M., Dietrich, L., Biggs, D., Lafky, J.M., and Loprinzi, C.L. (2016). Olanzapine for the Prevention of Chemotherapy-Induced Nausea and Vomiting. N. Engl. J. Med. *375*, 134–142.
Nelles, J.L., Hu, W.-Y., and Prins, G.S. (2011). Estrogen action and prostate cancer. Expert Rev. Endocrinol. Metab. *6*, 437–451.
Neu, J., and Rushing, J. (2011). Cesarean Versus Vaginal Delivery: Long-term Infant Outcomes and the Hygiene Hypothesis. Clin. Perinatol. *38*, 321–331.
O'Connor, M.J. (2015). Targeting the DNA Damage Response in Cancer. Mol. Cell *60*, 547–560.

Oberle, C., and Blattner, C. (2010). Regulation of the DNA Damage Response to DSBs by Post-Translational Modifications. Curr. Genomics *11*, 184–198.
Oddone, E., Modonesi, C., and Gatta, G. (2014). Occupational exposures and colorectal cancers: A quantitative overview of epidemiological evidence. World J. Gastroenterol. *20*, 12431–12444.
Oka, K., Tanaka, T., Enoki, T., Yoshimura, K., Ohshima, M., Kubo, M., Murakami, T., Gondou, T., Minami, Y., Takemoto, Y., et al. (2010). DNA damage signaling is activated during cancer progression in human colorectal carcinoma. Cancer Biol. Ther. *9*, 246–252.
Okada, H., Kuhn, C., Feillet, H., and Bach, J.F. (2010). The "hygiene hypothesis" for autoimmune and allergic diseases: An update. Clin. Exp. Immunol. *160*, 1–9.
Oppenheim, J.J. (2007). IL-2: More Than a T Cell Growth Factor. J. Immunol. *179*, 1413–1414.
Österlund, P., Ruotsalainen, T., Korpela, R., Saxelin, M., Ollus, A., Valta, P., Kouri, M., Elomaa, I., and Joensuu, H. (2007). Lactobacillus supplementation for diarrhoea related to chemotherapy of colorectal cancer: a randomised study. Br. J. Cancer *97*, 1028–1034.
Ouyang, D.L., Chen, J.J., Getzenberg, R.H., and Schoen, R.E. (2005). Noninvasive testing for colorectal cancer: A review. Am. J. Gastroenterol. *100*, 1393–1403.
Park, B., Yee, C., and Lee, K.M. (2014). The effect of radiation on the immune response to cancers. Int. J. Mol. Sci. *15*, 927–943.
Parker, W.B. (2009). Enzymology of purine and pyrimidine antimetabolites used in the treatment of cancer. Chem. Rev. *109*, 2880–2893.
Paschos, K.A., Canovas, D., and Bird, N.C. (2009). The role of cell adhesion molecules in the progression of colorectal cancer and the development of liver metastasis. Cell. Signal. *21*, 665–674.
Pearl, L.H., Schierz, A.C., Ward, S.E., Al-Lazikani, B., and Pearl, F.M.G. (2015). Therapeutic opportunities within the DNA damage response. Nat. Rev. Cancer *15*, 166–180.
Peltomaki, P. (2001). Deficient DNA mismatch repair: a common etiologic factor for colon cancer. Hum. Mol. Genet. *10*, 735–740.
Phang, P.T., Macfarlane, J.K., Taylor, R.H., Cheifetz, R.E., Davis, N., Hay, J.H., et al. (2002). Effects of positive resection margin and tumor distance from anus on rectal cancer treatment outcomes. Am J Surg. *183*, 504–508.
Pool-Zobel, B.. L., and Leucht, U. (1997). Induction of DNA damage by risk factors of colon cancer in human colon cells derived from biopsies. Mutat. Res. *375*, 105–115.
Post, M.J. (2012). Cultured meat from stem cells: Challenges and prospects. Meat Sci. *92*, 297–301.
Price, J. V, Tangsombatvisit, S., Xu, G., Yu, J., Levy, D., Baechler, E.C., Gozani, O., Varma, M., Utz, P.J., and Liu, C.L. (2012). On silico peptide microarrays for high-resolution mapping of antibody epitopes and diverse protein-protein interactions. Nat. Med. *18*, 1434–1440.
Proctor, L.M. (2011). The human microbiome project in 2011 and beyond. Cell Host Microbe *10*, 287–291.
Pui, C.-H., Robison, L.L., and Look, A.T. (2008). Acute lymphoblastic leukaemia. Lancet *371*, 1030–1043.
Puntoni, R., Russo, L., Zannini, D., Vercelli, M., Gambaro, R.P., Valerio, F., and Santi, L. (1977). Mortality among dock yard workers in Genoa, Italy. Tumori *63*, 91–96.

Rahman, M., Deleyrolle, L., Vedam-Mai, V., Azari, H., Abd-El-Barr, M., and Reynolds, B.A. (2011). The cancer stem cell hypothesis: Failures and pitfalls. Neurosurgery *68*, 531–545.
Reagan, J.L., and Castillo, J.J. (2014). Ofatumumab as front-line therapy in untreated chronic lymphocytic leukemia. Futur. Oncol. *10*, 1147–1155.
Ren, R. (2005). Mechanisms of BCR–ABL in the pathogenesis of chronic myelogenous leukaemia. Nat. Rev. Cancer *5*, 172–183.
Resta, N., Simone, C., Mareni, C., Montera, M., Gentile, M., Susca, F., Gristina, R., Pozzi, S., Bertario, L., Bufo, P., et al. (1998). STK11 mutations in Peutz-Jeghers syndrome and sporadic colon cancer. Cancer Res. *58*, 4799–4801.
Ricciardiello, L., Baglioni, M., Giovannini, C., Pariali, M., Cenacchi, G., Ripalti, A., Landini, M.P., Sawa, H., Nagashima, K., Frisque, R.J., et al. (2003). Induction of Chromosomal Instability in Colonic Cells by the Human Polyomavirus JC Virus. Cancer Res. *63*, 7256–7262.
Rini, B.I. (2005). VEGF-targeted therapy in metastatic renal cell carcinoma. Oncologist *10*, 191–197.
Roberts, M. (2016). UK air pollution "linked to 40,000 early deaths a year." BBC News http://www.bbc.co.uk/news/health-35629034.
Robertson, M.P., and Joyce, G.F. (2012). The origins of the RNA World. Cold Spring Harb. Perspect. Biol. *4*, 5.
Rockwell, S., Dobrucki, I.T., Kim, E.Y., Marrison, S.T., and Vu, V.T. (2009). Hypoxia and radiation therapy: past history, ongoing research, and future promise. Curr. Mol. Med. *9*, 442–458.
Rodu, B., Delzell, E., Beall, C., and Sathiakumar, N. (2001). Mortality among employees at a petrochemical research facility. Am. J. Ind. Med. *39*, 29–41.
Rose, A.M., and Bhattacharya, S.S. (2016). Variant haploinsufficiency and phenotypic non-penetrance in PRPF31-associated retinitis pigmentosa. Clin. Genet. *90*, 118–126.
Rostagno, M.H. (2009). Can Stress in Farm Animals Increase Food Safety Risk? Foodborne Pathog. Dis. *6*, 767–776.
Ryan, K.M., Phillips, A.C., and Vousden, K.H. (2001). Regulation and function of the p53 tumor suppressor protein. Curr. Opin. Cell Biol. *13*, 332–337.
Saif, M.W., and Reardon, J. (2005). Management of oxaliplatin-induced peripheral neuropathy. Ther. Clin. Risk Manag. *1*, 249–258.
Salmond, R.J., Brownlie, R.J., Morrison, V.L., and Zamoyska, R. (2014). The tyrosine phosphatase PTPN22 discriminates weak self peptides from strong agonist TCR signals. Nat. Immunol. *15*, 1–10.
Sampson, J.R., Jones, S., Dolwani, S., and Cheadle, J.P. (2005). MutYH (MYH) and colorectal cancer. Biochem. Soc. Trans. *33*, 679–683.
Samuels, Y., Wang, Z., Bardelli, A., Silliman, N., Ptak, J., Szabo, S., et al. (2004). High frequency of mutations of the *PIK3CA* gene in human cancers. Science. *304*, 554.
Schaffner, C., Idler, I., Stilgenbauer, S., Dohner, H., and Lichter, P. (2000). Mantle cell lymphoma is characterized by inactivation of the ATM gene. Proc. Natl. Acad. Sci. *97*, 2773–2778.
Schlom, J. (2012). Therapeutic cancer vaccines: Current status and moving forward. J. Natl. Cancer Inst. *104*, 599–613.
Schuchard, M., Landers, J.P., Sandhu, N.P., and Spelsberg, T.C. (1993). Steroid hormone

regulation of nuclear proto-oncogenes. Endocr Rev *14*, 659–669.
Schumacher, T.N., and Schreiber, R.D. (2015). Neoantigens in cancer immunotherapy. Science. *348*, 69–74.
Semenza, G.L. (2003). Targeting HIF-1 for cancer therapy. Nat. Rev. Cancer *3*, 721–732.
Seok, J., Warren, H.S., Cuenca, A.G., Mindrinos, M.N., Baker, H. V, Xu, W., Richards, D.R., McDonald-Smith, G.P., Gao, H., Hennessy, L., et al. (2013). Genomic responses in mouse models poorly mimic human inflammatory diseases. Proc. Natl. Acad. Sci. U. S. A. *110*, 3507–3512.
Shain, A.H., Yeh, I., Kovalyshyn, I., Sriharan, A., Talevich, E., Gagnon, A., Dummer, R., North, J., Pincus, L., Ruben, B., et al. (2015). The Genetic Evolution of Melanoma from Precursor Lesions. N. Engl. J. Med. *373*, 1926–1936.
Shalapour, S., Font-Burgada, J., Di Caro, G., Zhong, Z., Sanchez-Lopez, E., Dhar, D., Willimsky, G., Ammirante, M., Strasner, A., Hansel, D.E., et al. (2015). Immunosuppressive plasma cells impede T-cell-dependent immunogenic chemotherapy. Nature *521*, 94–98.
Shanks, N., and Pyles, R.A. (2007). Evolution and medicine: the long reach of "Dr. Darwin." Philos. Ethics, Humanit. Med. *2*, 4.
Shay, J.W., and Wright, W.E. (2000). Hayflick, His Limit, and Cellular Ageing. Nat. Rev. Mol. Cell Biol. *1*, 72–76.
Shen, X.J., Rawls, J.F., Randall, T., Burcal, L., Mpande, C.N., Jenkins, N., Jovov, B., Abdo, Z., Sandler, R.S., and Keku, T.O. (2010). Molecular characterization of mucosal adherent bacteria and associations with colorectal adenomas. Gut Microbes *1*, 138–147.
Shendure, J., and Ji, H. (2008). Next-generation DNA sequencing. Nat. Biotechnol. *26*, 1135–1145.
Shin, J.S., Tut, T.G., Ho, V**.,** and Lee, C.S. (2014). Predictive markers of radiotherapy-induced rectal cancer regression. J Clin Pathol. *67*, 859–864.
Shipitsin, M., Campbell, L.L., Argani, P., Weremowicz, S., Bloushtain-Qimron, N., Yao, J., Nikolskaya, T., Serebryiskaya, T., Beroukhim, R., Hu, M., et al. (2007). Molecular Definition of Breast Tumor Heterogeneity. Cancer Cell *11*, 259–273.
Siegel, R.L., Miller, K.D., and Jemal, A. (2016). Cancer statistics. CA Cancer J Clin. *66*, 7–30.
Siva, N. (2008). 1000 Genomes project. Nat. Biotechnol. *26*, 256.
Solal-Céligny, P., Roy, P., Colombat, P., White, J., Armitage, J.O., Arranz-Saez, R., Au, W.Y., Bellei, M., Brice, P., Caballero, D., et al. (2004). FLIPI: Follicular Lymphoma International Prognostic Index. Blood *104*, 1258–1265.
Sommer, A. (2008). Vitamin A Deficiency and Clinical Disease: An Historical Overview. J. Nutr *138*, 1835–1839.
Soussi, T., and Wiman, K.G. (2007). Shaping Genetic Alterations in Human Cancer: The p53 Mutation Paradigm. Cancer Cell *12*, 303–312.
Staib, C., Pesch, J., Gerwig, R., Gerber, J.K., Brehm, U., Stangl, a, and Grummt, F. (1996). p53 inhibits JC virus DNA replication in vivo and interacts with JC virus large T-antigen. Virology *219*, 237–246.
Standring, S., Borley, N. R., and Gray, H. (2008). *Gray's anatomy: the anatomical basis of clinical practice.* 40th ed., anniversary ed. [Edinburgh]: Churchill Livingstone/Elsevier.
Steele GD, Herndon JE, Bleday R, et al. (1999). Sphincter-sparing treatment for distal rectal adenocarcinoma. Ann Surg Oncol. *6*, 433-41.

Stirdivant, S. M., H. E. Huber, D. R. Patrick, D. Defeo-Jones, E. M. McAvoy, V. M. Garsky, A. Oliff, and D. C. Heimbrook. (1992). Human papillomavirus type 16 E7 protein inhibits DNA binding by the retinoblastoma gene product. Mol. Cell. Biol. *12***,** 1905-1914.

Stockert, E., Jäger, E., Chen, Y.T., Scanlan, M.J., Gout, I., Karbach, J., Arand, M., Knuth, A., and Old, L.J. (1998). A survey of the humoral immune response of cancer patients to a panel of human tumor antigens. J. Exp. Med. *187*, 1349–1354.

Stratton, M.R., Campbell, P.J., and Andrew F, P. (2009). The cancer genome. Nature *458*, 719–724.

Stubbe, C.E., and Valero, M. (2013). Complementary strategies for the management of radiation therapy side effects. J. Adv. Pract. Oncol. *4*, 219–231.

Subramaniam, R., Mizoguchi A., and Mizoguchi, E. (2016). Mechanistic roles of epithelial and immune cell signaling during the development of colitis-associated cancer. Cancer Research Frontiers. *2*, 1-21.

Sudhakar, A. (2009). History of Cancer, Ancient and Modern Treatment Methods. J. Cancer Sci. Ther. *1*, 1-4.

Sung, J.J.Y., Lau, J.Y.W., Goh, K.L., Leung, W.K., Chen, M.H., Li, C.J., Tandon, R., Makharia, G., Abdullah, M., Fujita, R., et al. (2005). Increasing incidence of colorectal cancer in Asia: Implications for screening. Lancet Oncol. *6*, 871–876.

Supriya, M., Denholm, S., and Palmer, T. (2008). Seeding of tumor cells after fine needle aspiration cytology in benign parotid tumor: a case report and literature review. Laryngoscope *118*, 263–265.

Suzuki, T., Sadahiro, S., Tanaka, A., Okada, K., Saito, G., Kamijo, A., Akiba, T., and Kawada, S. (2014). Relationship between histologic response and the degree of tumor shrinkage after chemoradiotherapy in patients with locally advanced rectal cancer. J. Surg. Oncol. *109*, 659–664.

Talpaz, M., Shah, N.P., Kantarjian, H., Donato, N., Nicoll, J., Paquette, R., Cortes, J., O'Brien, S., Nicaise, C., Bleickardt, E., et al. (2006). Dasatinib in Imatinib-Resistant Philadelphia Chromosome–Positive Leukemias. N. Engl. J. Med. *354*, 2531–2541.

Teldo-Pereyra, L.H. (2009). X-Rays Surgical Revolution. Investig. Surg. *22*, 327–332.

Thomas, G., Jacobs, K.B., Yeager, M., Kraft, P., Wacholder, S., Orr, N., Yu, K., Chatterjee, N., Welch, R., Hutchinson, A., et al. (2008). Multiple loci identified in a genome-wide association study of prostate cancer. Nat. Genet. *40*, 310–315.

Thompson, D., Duedal, S., Kirner, J., McGuffog, L., Last, J., Reiman, A., Byrd, P., Taylor, M., and Easton, D.F. (2005). Cancer risks and mortality in heterozygous ATM mutation carriers. J. Natl. Cancer Inst. *97*, 813–822.

Tian, H., Gao, Z., Li, H., Zhang, B., Wang, G., Zhang, Q., Pei, D., and Zheng, J. (2015). DNA damage response - A double-edged sword in cancer prevention and cancer therapy. Cancer Lett. *358*, 8–16.

Tie, J., Wang, Y., Tomasetti, C., Li, L., Springer, S., Kinde, I., Silliman, N., Tacey, M., Wong, H.-L., Christie, M., et al. (2016). Circulating tumor DNA analysis detects minimal residual disease and predicts recurrence in patients with stage II colon cancer. Sci. Transl. Med. *8*, 346ra92-346ra92.

Tjalsma, H., Boleij, A., and Kato, I. (2012). Streptococcus bovis and colorectal cancer. In Bacteria and Cancer, pp. 61–78.

Tol, J., Koopman, M., Cats, A., Rodenburg, C.J., Creemers, G.J.M., Schrama, J.G., Erdkamp, F.L.G., Vos, A.H., van Groeningen, C.J., Sinnige, H.A.M., et al. (2009). Chemotherapy, Bevacizumab, and Cetuximab in Metastatic Colorectal Cancer. N. Engl. J. Med. *360*, 563–572.
Travis, R.C., and Key, T.J. (2003). Oestrogen exposure and breast cancer risk. Breast Cancer Res. *5*, 239–247.
Truninger, K., Menigatti, M., Luz, J., Russell, A., Haider, R., Gebbers, J.O., Bannwart, F., Yurtsever, H., Neuweiler, J., Riehle, H.M., et al. (2005). Immunohistochemical analysis reveals high frequency of PMS2 defects in colorectal cancer. Gastroenterology *128*, 1160–1171.
Tubiana, M. (1987). [Radiotherapy-chemotherapy combinations]. Bull. Cancer *74*, 541–554.
Unicomb, L. (2009). Food Safety: Pathogen Transmission Routes, Hygiene Practices and Prevention. J Heal Popul Nutr. *27*, 599–601.
US Preventive Services Task Force. (2016). Screening for Colorectal Cancer: US Preventive Services Task Force Recommendation Statement. *JAMA. 315*, 2564–2575.
Van Meter, M., Simon, M., Tombline, G., May, A., Morello, T.D., Hubbard, B.P., Bredbenner, K., Park, R., Sinclair, D.A., Bohr, V.A., et al. (2016). JNK Phosphorylates SIRT6 to Stimulate DNA Double-Strand Break Repair in Response to Oxidative Stress by Recruiting PARP1 to DNA Breaks. Cell Rep. *16*, 2641–2650.
Vasen, H.F., Watson, P., Mecklin, J.P., Lynch, H.T., Lundell, L., Osteen, R., Lewis, F.R., Huntsman, D.G., Pharoah, P.D.P., Jankowski, J.A., et al. (1999). New clinical criteria for hereditary nonpolyposis colorectal cancer (HNPCC, Lynch syndrome) proposed by the International Collaborative group on HNPCC. Gastroenterology *116*, 1453–1456.
Venitt, S., and Phillips, D.H. (2012). Philip D. Lawley (1927-2011). Mutagenesis *27*, 255–256.
Voena, C., and Chiarle, R. (2016). Advances in cancer immunology and cancer immunotherapy. Discov Med *21*, 125–133.
Wang, J., Joshi, A.D., Corral, R., Siegmund, K.D., Marchand, L. Le, Martinez, M.E., Haile, R.W., Ahnen, D.J., Sandler, R.S., Lance, P., et al. (2012). Carcinogen metabolism genes, red meat and poultry intake, and colorectal cancer risk. Int. J. Cancer *130*, 1898–1907.
Wang, X., Fu, A.Q., McNerney, M.E., and White, K.P. (2014). Widespread genetic epistasis among cancer genes. Nat. Commun. *5*, 4828.
Wang, X.-S., Armstrong, M.E.G., Cairns, B.J., Key, T.J., Travis, R.C., and Nicholson, P.J. (2011). Shift work and chronic disease: the epidemiological evidence. Occup. Med. (Chic. Ill). *61*, 443–4; author reply 444.
Weinstein, J.N., Collisson, E.A., Mills, G.B., Shaw, K.R.M., Ozenberger, B.A., Ellrott, K., Shmulevich, I., Sander, C., and Stuart, J.M. (2013). The Cancer Genome Atlas Pan-Cancer analysis project. Nat. Genet. *45*, 1113–1120.
Weis, S.M., and Cheresh, D.A. (2011). Tumor angiogenesis: molecular pathways and therapeutic targets. Nat. Med. *17*, 1359–1370.
Weiss, R.A., and Vogt, P.K. (2011). 100 years of Rous sarcoma virus. J. Exp. Med. *208*, 2351–2355.
Werner, H.M.J., Mills, G.B., and Ram, P.T. (2014). Cancer Systems Biology: a peek into the future of patient care? Nat. Rev. Clin. Oncol. *11*, 167–176.
Westra, H.-J., Peters, M.J., Esko, T., Yaghootkar, H., Schurmann, C., Kettunen, J., Christiansen, M.W., Fairfax, B.P., Schramm, K., Powell, J.E., et al. (2013). Systematic

identification of trans eQTLs as putative drivers of known disease associations. Nat. Genet. *45*, 1238–1243.
Wheeler, H.E., Maitland, M.L., Dolan, M.E., Cox, N.J., and Ratain, M.J. (2012). Cancer pharmacogenomics: strategies and challenges. Nat. Rev. Genet. *14*, 23–34.
Wheeler, J.M.D., Warren, B.F., Mortensen, N.J.M.C., Ekanyaka, N., Kulacoglu, H., Jones, A.C., George, B.D., and Kettlewell, M.G.W. (2002). Quantification of histologic regression of rectal cancer after irradiation: A proposal for a modified staging system. Dis. Colon Rectum *45*, 1051–1056.
White, M.C., Holman, D.M., Boehm, J.E., Peipins, L.A., Grossman, M., and Jane Henley, S. (2014). Age and Cancer Risk. Am. J. Prev. Med. *46*, S7–S15.
WHO (2017). WHO Cancer Fact Sheet. WHO Fact Sheets http://www.who.int/mediacentre/factsheets/fs297/en.
Winawer, S,J,, Krabshuis, J,, Lambert, R,, O'Brien, M,, and Fried, M. (2011). Cascade colorectal cancer screening guidelines: a global conceptual model. J Clin Gastroenterol. *45*, 297–300.
Wolpin, B.M., and Mayer, R.J. (2008). Systemic Treatment of Colorectal Cancer. Gastroenterology *134*, 1296-1310.
Wood, L.D., Parsons, D.W., Jones, S., Lin, J., Sjöblom, T., Leary, R.J., Shen, D., Boca, S.M., Barber, T., Ptak, J., et al. (2007). The genomic landscapes of human breast and colorectal cancers. Science *318*, 1108–1113.
Xia, T., Konno, H., Ahn, J., and Barber, G.N. (2016). Deregulation of STING Signaling in Colorectal Carcinoma Constrains DNA Damage Responses and Correlates With Tumorigenesis. Cell Rep. *14*, 282–297.
Xu, F., and Li, X. (2012). The role of histone methyltransferase EZH2 in myelodysplastic syndromes. Expert Rev. Hematol. *5*, 177–185.
Xu, M., Myerson, R.J., Hunt, C., Kumar, S., Moros, E.G., Straube, W.L., and Roti Roti, J.L. (2004). Transfection of human tumour cells with Mre11 siRNA and the increase in radiation sensitivity and the reduction in heat-induced radiosensitization. Int. J. Hyperth. *20*, 157–162.
Yin, L. lan, Chung, C.M., Chen, J., Fok, K.L., Ng, C.P., Jia, R.R., Ren, X., Zhou, J., Zhang, T., Zhao, X.H., et al. (2009). A suppressor of multiple extracellular matrix-degrading proteases and cancer metastasis. J. Cell. Mol. Med. *13*, 4034–4041.
Yu, L.G., Milton, J.D., Fernig, D.G., and Rhodes, J.M. (2001). Opposite effects on human colon cancer cell proliferation of two dietary Thomsen-Friedenreich antigen-binding lectins. J. Cell. Physiol. *186*, 282–287.
Zhang, L., Smit-McBride, Z., Pan, X., Rheinhardt, J., and Hershey, J.W.B. (2008). An oncogenic role for the phosphorylated h-subunit of human translation initiation factor eIF3. J. Biol. Chem. *283*, 24047–24060.
Zlobec, I., Steele, R., Nigam, N., and Compton, C.C. (2005). A predictive model of rectal tumor response to preoperative radiotherapy using classification and regression tree methods. Clin. Cancer Res. *11*, 5440–5443.
Zumkeller, N., Brenner, H., Zwahlen, M., and Rothenbacher, D. (2006). Helicobacter pylori infection and colorectal cancer risk: A meta-analysis. Helicobacter *11*, 75–80.

CHAPTER 2
Materials and Methods

2.1 Introduction

The results presented in subsequent sections concern the value of DDR proteins in informing rectal cancer prognosis. In total, the expression levels of four key DDR proteins (ATM, MRE11, NSB1 and RAD50) were analysed in rectal tumour sections using histological staining with well validated antibodies. All samples were previously collected and stored from a local rectal cancer cohort, and each protein of interest was analysed independently of the others – aside from experiments in which MRE11 and ATM together served as an informative marker panel, as we will discuss. Accordingly, the results concerning each protein will be presented as individual results chapters, with the combinatorial panel presented in Chapter Four.

2.2 Ethical considerations and approvals

Robust ethical standards and protocols - with regards to patient welfare and safety, and scientific conduct - are essential to safely and accurately progress biological and clinical research (Emanuel et al., 2000). Indeed, since Hippocrates, medical practitioners have appreciated the importance of patient confidentiality, respect, communication and privacy. The past two decades have witnessed a welcome expansion in the controversial medical ethics public and private sectors, which now employs diverse committees of experts, philosophies and specialist administrators to better safeguard participants and their data whilst allowing the scientific community to benefit from medical advancements (Edwards et al., 2007; Jamrozik, 2004).

However, ethics does not only concern patient and data security, but also scientific conduct; although this is more dependent and susceptible to personal intervention. Along these lines, academic and private science sectors have been besieged by erroneous, irreproducible publications in recent history (Baker, 2016); which may be a consequence of the highly competitive environment in academic research, where one must "publish or perish" (Neill, 2008). As a responsible author, I here guarantee the integrity of the data presented. Every effort has been made to evaluate data in a statistically-stringent, blinded and objective manner. All stages of our study have been frequently and critically evaluated by collaborators and professional colleagues, allowing feedback to be incorporated in real-time.

In order to secure a strong ethical framework for our studies, we sought approval for our work from the local ethics committee. Following review and discussion, ethical approval for the ATM, MRE11 and NSB1 studies was granted on the 22nd June 2012 by the South-Western Sydney Local Health District (SWLHD) Human Research Ethics Committee (HREX); reference number HREC/12/LPOOL/102. Specifically, this allowed for the assessment of paraffin-embedded rectal cancer tissues and retrospective survival analyses, and was conducted under issued protocol numbers X01-0138 and X03-0291. Ethical approval for the histological studies concerning RAD50 was also granted by the SWLHD HREC; reference number HREC/14/LPOOL/186, project number 14/103.

Given the retrospective and exploratory nature of our study, all projects were deemed to be of low or negligible risk to participants. Accordingly, the aforementioned ethical committee waived the need for written and informed consent from participants, facilitating samples access and our studies. Despite this waiver, and in all cases, to avoid the introduction of experimental biases and safeguard patient identity, all identifiable information was anonymised and coded prior to use by researchers. Furthermore, all data analyses were carried out in a (researcher) blinded-manner, and only after the experimental protocol had been completed for all samples within the experiment. Furthermore, all histological scoring was carried out by independent pathologists and results averaged.

All data arising from this work will be deposited in the public domain via this thesis and peer-reviewed publications, thus ensuring open public access to our findings.

2.3 Patient Samples

All patient samples (paraffin wax embedded tumour cores) used in our studies were obtained from the local SWLHD Pathology Database (Australia), which maintained a high-quality collection of primary tumours, involved regional lymph nodes, and healthy tissue from rectal cancer patients.

All samples were collected from patients undergoing tumour resection therapy in the district between 2000 and 2011, and stored for distribution to different research groups upon application. All slides from each case were available for review. In our cohort, surgery (for rectal or rectosigmoid tumours) consisted of total mesorectal excision with anterior or abdominoperineal resection, and samples were collected post-operatively for comparative analyses, as described (Tut et al., 2015).

Throughout the study period, patient follow-up consisted of specialist clinic visits, routine blood biochemistry analyses, colonoscopy, and gastrointestinal imaging - based on the recommendations of the administering physician and requirements of the patient at the time.

2.4 Radiotherapy and chemotherapy treatments

As radiosensitivity is influenced by proteins involved in the DDR, this subgroup of patients is highly relevant to our studies. The patients enrolled in our study that underwent radiotherapy treatment, as in the ATM or RAD50 experiments (*please see* Results Chapters), received either a 25 Gy dose over five treatment fractions, or a 50.4 Gy dose over 28 fractions. Furthermore, all of these patients also received 5-fluorouracil-based chemotherapy and were under the care of professionals within the SWLHD.

2.5 Outcome measures

In order to evaluate whether the DDR proteins of interest had a predictive purpose in rectal cancer, patient disease course was considered. Outcome measures across our studies included:

disease-free survival (DFS) time; overall survival (OS) time; and histological tumour regression grade (TRG) in the resected bowel. DFS was defined as the time between diagnosis and the first recurrence of disease, and OS was defined as the time between diagnosis and the last follow-up date or death.

In patients receiving adjuvant radiotherapy, the short-term response to treatment was measured by TRG, according to the guidelines set out in the 7th edition of the American Joint Committee on Cancer manual (Edge and Compton, 2010; Hari et al., 2013), and modified by Ryan *et al*(Ryan et al., 2005). Accordingly, TRG is described on a scale of 0 – 3: 0 represents a complete response, without any viable malignant cells being detected; 1 represents a moderate response, with small groups of malignant cells still detectable; 2 represents a minimal response to treatment, with fibrosis outgrowing residual malignancy; while 3 typifies a poor response, with abundant residual malignancy. TRG scoring was carried out by two (or more) independent pathologist researchers in a blinded manner. Patients categorised as having a TRG score of 0, 1, or 2 were considered *responders*, in our studies, whilst those with a TRG score of 3 were considered *non-responders*.

To gain additional insight into disease pathophysiology, additional variables of interest were interrogated after staining the tissue sections. Clinicohistopathological variables of interest across our studies included: age, sex, pathological TNM stage, tumour grade, the presence of vascular or perineural invasion, the level of tumour-infiltrating lymphocytes, and treatment.

2.6 Sample preparation and tissue microarrays

Archival formalin-fixed, paraffin wax embedded tissue blocks from postoperative rectal cancers were retrieved for each patient in the study. From each tissue block, two cores (0.6 mm in diameter) were obtained from each of five different sampling sites, namely: the tumour centre (TC); the tumour periphery at the invasive edge (TP); normal mucosa close/adjacent to the tumour (NCT); normal mucosa more distal (away) to the tumour (NAT); and any involved regional lymph nodes (LN). For all samples, corresponding (previously-prepared) hematoxylin and eosin-stained sections were reviewed to determine the most representative areas of oncogenesis and associated normal colorectal mucosa for our sampling and analyses.

Tissue microarrays (TMAs) are now well recognised to provide an efficient means to achieve high-throughput screening of biomarkers from tissue sections. Thus, we exploited array profiling technology for our investigations. Originally, such microarrays were designed to interrogate a single biomarker in samples from a large number of patients (Kononen et al, 1998); as TMAs work by placing several samples on a single glass slide. Each sample is approximately 0.6 – 2.0 mm diameter (Horvath and Henshall, 2001).

The process of constructing a TMA is here described. A 37 x 24 x 5 mm sized mould was created from the recipient paraffin wax block. From H& E stained histological sections, a manual tissue arrrayer (Beecher Manual Tissue Arrayer, Model-MTA-1, Beecher Instruments, Inc.) was used to take 1 mm diameter tissue cores, creating a hole in the wax block (Cardano

et al., 2013). Tissue cores were subsequently inserted into pre-punched wells. Samples can be arranged in up to 8 rows and 10 columns, with each sample being 2 mm apart from the adjacent specimen. Human error is reduced by the fact that each position on the TMA can be identified by a coded recorded pertaining to the sample.

A schematic representation showing the steps involved in the manual construction of a tissue microarray (TMA) is shown in Figure 2.1 (Horvath and Henshall, 2001). After construction of TMA blocks, they were heated for 7 minutes at 60°C; to seal any gaps between tissue cores and surrounding wax. TMA blocks were left at room temperature for 5 minutes before being heated again to 60°C for 5 minutes. After removal from the oven, TMA blocks left to cool at room temperature. The heating and cooling cycle can be repeated to facilitate gap closing. TMA blocks were store at 4°C overnight. Tissue paraffin cores were shaved into thin sections (4µm) in the manual microtome (Microm HM325 Germany GMBH, Thermo Scientific, Cat No: 902100) and the shaved thin sections were then added to a circular paraffin section flotation bath (Thermo Scientific Cat No: 3120059) before being added to specially-coated slides (Bancroft and Gamble 2002, Bancroft and Gamble 2008). It is important to remove all paraffin wax present on the slide, as this will facilitate aqueous antibody solution penetration into the tissue. Then the glass slide was fixed and stained by H&E.

Detailed TMA maps were constructed from blocks and are available for access upon request in an Excel password protected file containing patient identifiers. A TMA map example is shown in Figure 2.2 as an example, with patient surnames abbreviated to an initial to preserve privacy. Positive controls were included in the TMA blocks. Antibodies selected for the studies were commercially tested by companies in different methods including immunocytochemistry staining, western blot, immunofluorescence staining and immunoprecipitations with strong specificity in human tissues and cells.

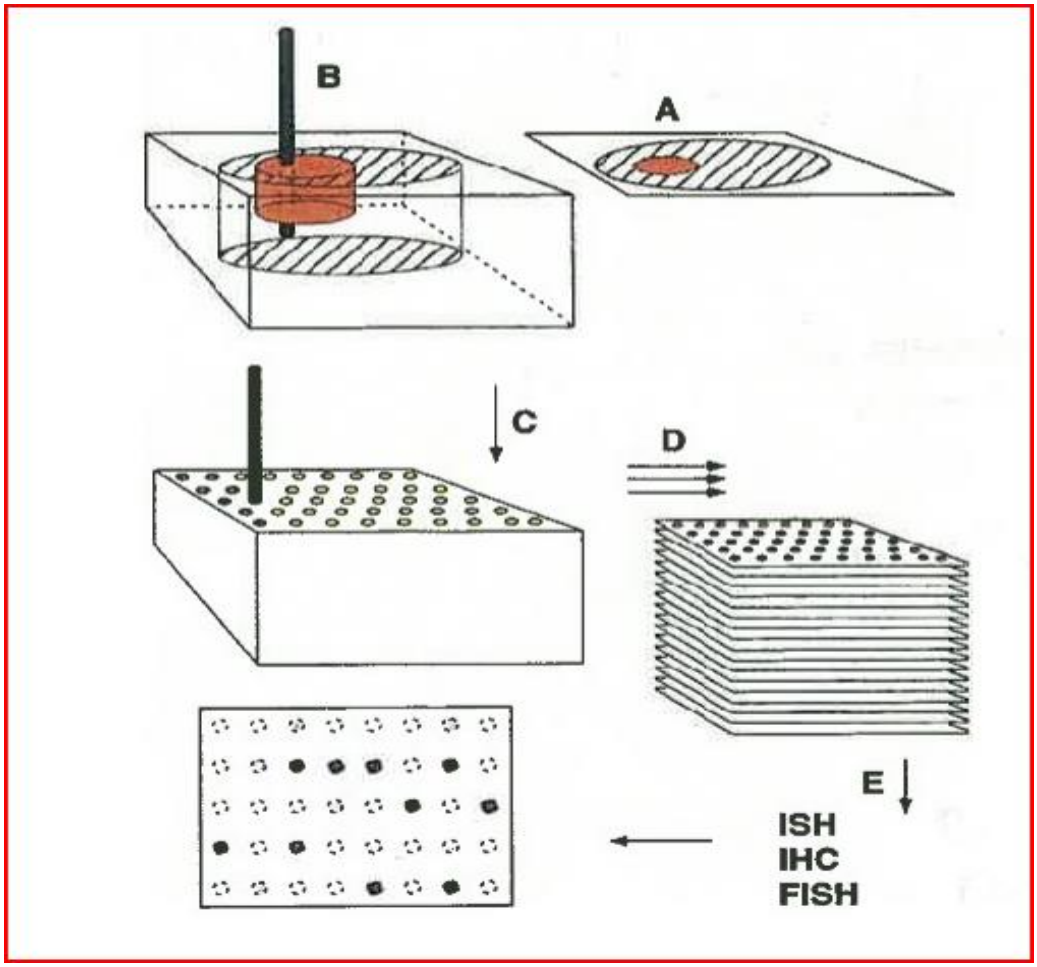


Figure 2-1Immunohistochemistry tissue microarray

(A) The histopathological tumour lesion is first identified on the donor paraffin blocks, which has already been stained with H & E, as described.
(B) A cylindrical core tissue biopsy (between 0.6 - 2.0 mm in diameter) was then cut from donor paraffin block and inserted into a pre-prepared recipient block. Samples are arranged in a grid for easier high-throughput analysis.
(C) In our study, samples from one patient were placed adjacent to one another, repeating the punching and depositing procedure.
(D) The newly-filled recipient paraffin block can then be sectioned to produce various tissue sections from the same tumour core.
(E) TMAs are then sent for light microscopy analysis.

Patient A		Patient B		Patient C	
L	H000012281	C1	H00002424	C2	H00003086
K	H00012814	Z	H00014995	B1	H00017273
P1	H00017494	B2	H00019872	B3	H01000931
N	H01002814	S1	H01005139	C3	H01007300
H	H01012045	J	H01017580	S2	H01018708
X	H01019327	P2	H01019755	T	H01020975
A	H010220401	M	H02000226	P3	H02000259

	Patient A				Patient B				Patient C			
	Cancer	Cancer	Normal	Normal	Cancer	Cancer	Normal	Normal	Cancer	Cancer	Normal	Normal
Control	**12**	**12**	**50%strong**	X	**12**	**12**	X	X	X	X	95%strong	95%strong
	12	X	X	X	**12**	**12**	**9**	**12**	X	X	X	X
	X	X	X	X	**12**	**12**	**12**	**12**	**12**	**12**	**9**	**9**
	X	X	X	X	X	X	X	X	**X**	**X**	**6**	X
	12	**12**	**80%strong**	**80%stong**	**12**	**12**	**9**	**9**	X	X	X	X
	X	X	X	X	X	X	x	x	x	9	9	X
	X	X	X	X	**1**	**x**	X	X	X	X	X	X

Figure 2-2 Original TMA blocks presented with patient identifiers removed and corresponding TMA maps constructed

2.7 Immunohistochemistry

The different DNA damage response proteins of interested were analysed in overlapping sets of biopsy/surgical samples from the named Pathology Database, as described: ATM expression was quantified in 263 rectal cancer patients, 54 of which received preoperative radiotherapy; MRE11 expression was quantified in tumours and involved regional lymph nodes from 262 patients, 54 of whom also underwent neoadjuvant radiotherapy; NSB1 expression was analysed in tumours and regional lymph nodes from an overlapping set of 260 patients; and RAD50 expression was quantified histologically in a collection of 266 rectal cancer patients receiving chemotherapy and radiotherapy. Detailed cohort descriptions are presented in their respective results chapters.

Slides were deparaffinized in xylene and rehydrated in graded ethanol. Slides were pre-heated in the oven at 60^{o}C for 1 hour and 30 minutes, before the slides were fully immersed in decreasing concentration ethanol solutions (100%, 95%, 70%, 50% EtOH) for 6 minutes each. Antigen retrieval was performed with Envision™ FLEX Target Retrieval Solution, pH 9.0, in a 98°C water bath for 10 minutes for ATM and 45 minutes for MRE11, NBS1 and RAD50. This method has been previously described for colorectal tumour staining (Holck et al, 2015). This was followed by incubation for 5 minutes at room temperature with Envision™ FLEX Peroxidase-Blocking Reagent to block endogenous peroxidases.

Blocked and washed slides were then incubated with the chosen primary (unconjugated) antibody raised against the DNA damage response protein epitope of interest, as detailed:

- ATM expression was analysed with the murine monoclonal 2C1(1A1) obtained from Abcam, UK or Sapphire Bioscience, Australia. 2C1(1A1) was used at a 1:800 dilution for a minimum of 30 minutes.
- MRE11 expression was analysed with the murine monoclonal 12D7 obtained from Abcam, UK. 12D7 was used at a 1:600 dilution for a minimum of 30 minutes.
- NSB1 expression was analysed using the antibody clone NBP1-06609 obtained from Novus Bioscience, USA. This reagent was used at a 1:800 dilution for a minimum of 30 minutes.
- RAD50 expression was analysed with the murine monoclonal 13B3/2C6 obtained from Abcam, UK. 13B3/2C6 was used at a 1:400 dilution for a minimum of 30 minutes.
- Following antigen retrieval in the ATM experiments (Chapter Three), primary antibodies MLH1 (clone ES05, dilution 1:50; Dako), MSH2 (clone FE11, dilution 1:50; Dako), MSH6 (clone EP49, dilution 1:50; Dako), and PMS2 (clone EP51, dilution 1:40; Dako) were applied to the slides for 15 minutes using the DAKO Autostainer in 1 mM EDTA buffer, pH 8.0.

The antibody's specificity and reproducibility are two key elements in antibody validation and verification. Particularly, when it comes to IHC, standardization can be quite challenging due to the number of pre-analytical, analytical and post-analytical factors known to influence staining in the TMA assay. In our laboratory with a standardised analytical procedure setup we

included both positive (using colorectal normal tissue samples) and negative (IgG only) controls to ensure the antibody's reproducibility in the IHC assays.

All primary antibodies were incubated with the sample slides for 30-60 minutes (depending on the antigen and antibody) at room temperature. Subsequently, slides were washed twice in TBST (as before) and incubated for 15 minutes with DAKO mouse/rabbit linker, as per the manufacturer's instructions. Slides were again washed with TBST solution prior to being incubated for 60 minutes with an anti-mouse/rabbit secondary antibody for primary antibody detection. Following incubation with secondary antibodies, slides were thoroughly washed three times in TBST, as before, to remove unbound antibody and background staining levels.

The peroxidase substrate used for development and antigen detection was a mixture of Envision™ FLEX DAB + Chromogen DM827 and Envision™ FLEX Substrate Buffer DM823 (DAKO, Denmark), which was incubated with the slides for 2-5 minutes, or until a brown colour developed on the stained section. As soon as any brown colour had developed and was noticeable by eye, the stained sections were immediately washed twice in TBST.

Sections were then counterstained in an automated stainer with haematoxylin solution for 10 seconds before being washed with cold water, and dipped 10 times in Scott Bluing solution. Haematoxylin, the oxidized forms of haematin, needs a metallic salt to act as a mordant in linking with nuclear chromatin (anionic cellular components). Eosin binds to collagen and cytoplasm (cationic components) of the cell reciprocally (Bancroft and Gamble 2002, Bancroft and Gamble 2008). Slides were then immediately rinsed with cold water prior to dehydration and mounting with Water-Based Immuno Mount (ThermoFisher, USA).

2.8 Histological scoring

For all samples, scoring was carried out in a blinded manner by two independent, experienced pathologists, who evaluated the intensity and proportion of positive immunohistochemical staining in each sample objectively and blinded to identifiable information. In all cases, scoring was based upon the methods previously described, with differences in the scoring method representing the antibodies and antigens in question.

In all cases, immuno-stained sections were examined by manual counting of cells in each tissue microarray dot, with the observers blinded to clinical outcomes. Briefly:

- ATM expression was scored as the product of percentage and intensity of staining, based on the work of Angèle and colleagues (2004). The percentage of positive staining was categorised as: < 25% (1), 25-50% (2), 50-75% (3), or > 75% (4); whilst the intensity of staining was graded as null (0), low (1), moderate (3), or high (5). Both scores were multiplied to create a composite score between 0 and 20, which allowed for accurate comparison of ATM expression between tissues. ATM expression was then categorised into negative (0) or positive (1-20) groups and analysed against available clinicohistopathological and clinical outcome data. This protocol has

previously been published and reliable results have been achieved (Ho et al, 2016; Ho et al, 2017; Ho et al, 2018).

- MRE11 expression was quantified in a similar manner, as described by Rodel *et al* (2010). The percentage of positively-staining cells and the staining intensity were scored. Intensity was graded as: negative (0), weak (1), moderate (2), or strong (3). The percentage of positive cells was graded as: < 5% (0); 5%–25% (1); 26%–50% (2); 51%–75% (3); and > 75% (4). These two measures were multiplied to produce a weighted score between 0 and 12, and dichotomized into low (0–5) or high (6–12) expression.
- NSB1 expression was scored into either a negative expression group (score: 0) or positive expression group (score: 1-12).
- RAD50 expression was scored in a similar manner to MRE11, and as described in our publication arising from this work, Ho et al (2017). Intensity was graded as: negative (0); weak (1); moderate (2); or strong (3). The percentage of positive cells was graded as: < 5% (0); 5%–25% (1); 26%–50% (2); 51%–75% (3); > 75% (4). These two measures were multiplied to obtain weighted scores ranging from 0–12. Again, samples were categorised into either a low expression group (0–5) or a high expression group (6–12).
- ATM and MRE11 combinatorial panel scoring was achieved as follows. Firstly, both ATM and MRE11 staining intensity were analysed as described for each individual antigen. Subsequently, all cases were categorised into either a low expression group (score range: 0-5) or a high expression group (score range: 6-32), in which both proteins are considered. All immunohistochemical scoring system for both MRE11 and ATM was based on the methods published previously (Ho et al, 2016; Ho et al, 2018).
- MRE11, NBS1 and RAD50 combinatorial panel staining was scored as follows. Two independent pathologists scored the intensity and percentage of positive immunohistochemical staining in each sample. Protein expression was calculated as the product of the staining percentage and intensity (Ho et al, 2018). Intensity was graded as: 0, negative; 1, weak; 2, moderate; or 3, strong. The percentage of positive cells was graded as: 0, <5%; 1, 5%–25%; 2, 26%–50%; 3, 51%–75%; or 4, >75%. These two measures were multiplied to obtain weighted expression scores ranging from 0–12. All tumour samples were categorized into either a low (score range: 0–<6) or high (score range: 6–12) expression group. The two scores for each biological sampling site were averaged, yielding final average weighted scores.

2.9 Statistical Analyses

All statistical analyses were carried out in SPSS Statistics for Windows 20.0 (Chicago, IL, USA).

2.9.1 ATM

Differences in ATM expression between tissue types were analysed using paired-sample t-tests. ATM expression was compared with MMR protein expression and clinicohistopathological data using Pearson χ^2 and Fisher's exact test. Survival analyses were performed separately in patients who received preoperative radiotherapy and those who did not. Disease-free and overall survival data were analysed using the Kaplan-Meier method and a log-rank test. Univariate and multivariate analyses were performed using Cox proportional hazards survival modelling for ATM expression in the TC and TP. Covariates included sex, age, TNM stage, tumour grade, vascular invasion, perineural invasion, treatment with chemotherapy and radiotherapy, and TRG. $P<0.05$ was considered statistically significant.

2.9.2 MRE11

Paired t-tests were used to compare weighted scores between the low and high expressing MRE11 groups in the TC and TP. The scores between the tumour and healthy tissue were also analysed in this way. The Fisher's exact test was used to test for associations between MRE11 status and the available clinicopathologic variables for our cohort. For these tests, the significance threshold was set to $P< 0.05$.

Furthermore, Cox regression models were used to test for univariate associations between the clinicopathological variables according to MRE11 status in the TC and TP. The assumption of proportional hazards was tested using log minus log plots; where variables met the assumption of proportional hazards if the plot showed a parallel curve between the groups over time. An interaction term with time was created for variables that did not meet the proportional hazards criteria and that interaction term was added to the model (grade did not meet the criteria for proportional hazards). If a variable was significant in one or both MRE11 groups for TC and TP in the univariate analysis, it was included in the multivariate analysis. Variables were excluded until only significant variables remained in one or both of the MRE11 groups. Two-way interactions between the remaining variables were explored and included in the models if they were significant. Subgroup analysis was performed on lymph node-positive subjects and adenoma-positive subjects to determine whether there was a difference in survival between these groups according to MRE11 high or low status.

2.9.3 ATM/MRE11 combinatorial panel

ATM expression was compared with clinicohistopathological data using Pearson's χ^2 test, and the association of MRE11 expression with clinicohistopathological variables was assessed by Fisher's exact test. ATM and MRE11 expression levels were compared and combined by binary logistic regression, as described previously. Survival analyses were performed in the overall cohort and separately in patients who received preoperative radiotherapy. Univariate and multivariate analyses were performed using Kaplan Meier curves and Cox's proportional hazards survival modeling for the combined two-marker expression levels from cancer core and periphery samples. Covariates were sex, age, TNM

stage, grade, vascular invasion, perineural invasion, treatment with chemotherapy and radiotherapy, and TRG. Univariate analysis by the Mann–Whitney U test was also used to assess associations between the single and combined two-marker expression levels in rectal tumour tissue with TRG, which was further characterized with receiver operating characteristic—area under curve (ROC-AUC) analysis. $P<0.05$ was considered statistically significant.

2.9.4 NSB1

Survival analyses were conducted both for the entire cohort, and separately, in samples from patients who received pre-operative radiotherapy, as has been described previously. In addition, further subgroup analysis was conducted with early tumour stage and low-grade tumours as covariates. Univariate and multivariate analyses were performed using Kaplan–Meier curves and Cox's proportional hazards survival modelling for NBS1 protein expression in the TC and TP. The covariates included were sex, age, TNM stage, tumour grade, vascular invasion, perineural invasion, chemotherapy and radiotherapy, and TRG. Univariate analysis was performed using the Mann–Whitney U test. $P < 0.05$ was considered statistically significant.

2.9.5 RAD50

Survival analyses were conducted both for the entire cohort, and separately, in samples from patients who received pre-operative radiotherapy. In addition, further subgroup analysis was conducted with early tumour stage and low-grade tumours as covariates. Univariate and multivariate analyses were performed using Kaplan–Meier curves and Cox's proportional hazards survival modelling for RAD50 protein expression in the TC and TP. The covariates included were sex, age, TNM stage, tumour grade, vascular invasion, perineural invasion, chemotherapy and radiotherapy, and TRG. Univariate analysis was performed using the Mann–Whitney U test. $P < 0.05$ was considered statistically significant.

2.9.6 MRE11, NBS1, RAD50 combinatorial panel

Statistical analysis was performed with SPSS Statistics for Windows 20.0 (Chicago, IL, USA). Survival analysis was conducted both for the entire cohort and, separately, in patients who received preoperative radiotherapy. MRE11, RAD50, and NBS1 expression were compared and combined by binary logistic regression (detailed raw data available upon request). Univariate and multivariate analyses of the combined expression of the three proteins at the TC and TP were performed using Kaplan–Meier curves and Cox's proportional hazards survival modelling. Covariates were sex, age, TNM stage, histological grade, vascular invasion, perineural invasion, chemotherapy, and radiotherapy. Univariate analysis was performed using the Mann–Whitney U test. $P<0.05$ was considered statistically significant.

2.10 References

Angèle, S., Falconer, A., Foster, C.S., Taniere, P., Eeles, R.A., and Hall, J. (2004). ATM Protein Overexpression in Prostate Tumors: Possible Role in Telomere Maintenance. Am. J. Clin. Pathol. *121*, 231–236.

Baker, M. (2016). 1,500 scientists lift the lid on reproducibility. Nature *533*, 452–454.

Cardano, M., Diaferia, G.R., Falavigna, M., Spinelli, C.C., Sessa, F., DeBlasio, P., and Biunno, I. (2013). Cell and tissue microarray technologies for protein and nucleic acid expression profiling. J. Histochem. Cytochem. *61*, 116–124.

Edge, S.B., and Compton, C.C. (2010). The American Joint Committee on Cancer: the 7th Edition of the AJCC Cancer Staging Manual and the Future of TNM. Ann. Surg. Oncol. *17*, 1471–1474.

Edwards, S.J.L., Stone, T., and Swift, T. (2007). Differences between research ethics committees. Int. J. Technol. Assess. Health Care *23*, 17–23.

Emanuel, E.J., Wendler, D., and Grady, C. (2000). What makes clinical research ethical? JAMA *283*, 2701–2711.

Hari, D.M., Leung, A.M., Lee, J.H., Sim, M.S., Vuong, B., Chiu, C.G., and Bilchik, A.J. (2013). AJCC cancer staging manual 7th edition criteria for colon cancer: Do the complex modifications improve prognostic assessment? J. Am. Coll. Surg. *217*, 181–190.

Ho, V., Chung, L., Revoltar, M., Lim, S.H., Tut, T.G., Abubakar, A., Henderson, C.J., Chua, W., Ng, W., Lee, M., et al. (2016). MRE11 and ATM expression levels predict rectal cancer survival and their association with radiotherapy response. PLoS One *11*, e0167675.

Ho, V., Chung, L., Singh, A., Lea, V., Revoltar, M., Lim, SH., Tut, TG., Ng, W., Lee, M., de Souza, P., Shin, J and Lee, CS. (2017). Early postoperative low expression of RAD50 in rectal cancer patients associates with disease free survival. Cancers. *9*, 163.

Ho, V., Chung, L., Singh, A., Lea, V., Abubakar, A., Lim, SH., Ng, W., Lee, M., de Souza, P., Shin, J and Lee, CS. (2018). Overexpression of the MRE11-RAD50-NBS1 (MRN) complex in rectal cancer correlates with poor response to neoadjuvant radiotherapy and prognosis. BMC Cancer. *18*, 869.

Holck, S., Nielsen, H.J., Pedersen, N., and Larsson, L.-I. (2015). Phospho-ERK1/2 levels in cancer cell nuclei predict responsiveness to radiochemotherapy of rectal adenocarcinoma. Oncotarget *6*, 34321–34328.

Horvath, L., and Henshall, S. (2001). The application of tissue microarrays to cancer research. Pathology. *33*, 125–129.

Jamrozik, K. (2004). Research ethics paperwork: what is the plot we seem to have lost? BMJ *329*, 286–287.

Kononen, J., Bubendorf, L., Kallioniemi, A., et al. (1998). Tissue microarrays for high-throughput molecular profiling of tumor specimens. Nat Med. *4*, 844–847.

Neill, U.S. (2008). Publish of perish, but at what cost? J. Clin. Invest. *118*, 2368.

Rodel, F., Keppner, S., Capalbo, G., Bashary, R., Kaufmann, M., Rodel, C., Strebhardt, K., and Spankuch, B. (2010). Polo-like kinase 1 as predictive marker and therapeutic target for radiotherapy in rectal cancer. Am J Pathol *177*, 918–929.

Ryan, R., Gibbons, D., Hyland, J.M.P., Treanor, D., White, A., Mulcahy, H.E., O'Donoghue, D.P., Moriarty, M., Fennelly, D., and Sheahan, K. (2005). Pathological response following

long-course neoadjuvant chemoradiotherapy for locally advanced rectal cancer. Histopathology *47*, 141–146.
Tut, T.G., Lim, S.H.S., Dissanayake, I.U., Descallar, J., Chua, W., Ng, W., de Souza, P., Shin, J.-S., and Lee, C.S. (2015). Upregulated Polo-Like Kinase 1 Expression Correlates with Inferior Survival Outcomes in Rectal Cancer. PLoS One *10*, e0129313.

CHAPTER 3
The Association Between ATM Expression and Survival Outcomes in Rectal Cancer

3.1 Introduction

Surgical resection remains the definitive treatment option for colorectal cancer, and advances in surgical techniques, will continue to serve patients well. However, as we will discuss in this and subsequent sections, in order to improve patient diagnoses and prognoses, we need to precisely determine the molecular mechanisms underpinning cancer development, progression and treatment in the individual; with markers of these processes serving not only to improve patient welfare and lifespan (i.e. by using biomarkers of radiotherapy responses, or cancer phenotypes, tailoring treatment), but also resource usage in under-pressure public healthcare systems.

Such approaches are particularly pertinent to rectal cancer, as anatomy limits surgical access to the pelvic space, and the close proximity of the rectum to other potential sites for metastases, make surgical approaches more complicated than for colon cancer. That said, laparoscopic techniques continue to advance and improve patient outcomes in rectal cancer, as it reduces the time spent in surgery, results in a more rapid return of normal bowel function and reduces other comorbidities, such as intestinal adhesions and postoperative abdominal bleeding (Trastulli et al., 2012). In all cases, advancements will come from the laboratory-clinic interface, which serves to unite fundamental biology with clinical manifestations, increasingly, in the individual (Chan and Ginsburg, 2011; Di-Paolo et al., 2017).

As introduced previously, the DNA damage response is central to cancer development and progression, and has been repeatedly associated with clinical utility, with various participating members of the response being associated with patient survival times and the response to radiotherapy (Bartkova et al., 2005; Choudhury et al., 2010). Here, we will describe the utility of a single DDR protein biomarker, namely *ATM* protein expression in tumour tissue, to inform upon prognosis and disease-free survival in a cohort of rectal cancer patients.

Importantly, a subset of the cohort received pre-operative (neoadjuvant) radiotherapy, allowing us to investigate any associations between treatment and biomarker expression. Preoperative radiotherapy attempts to downstage the tumour, improving outcomes after surgical resection. However, the response to radiotherapy is highly heterogeneous, with some patients responding to a much greater degree than others, as discussed. Across all cancers, approximately 60% of patients show a response to pre-operative radiotherapy (Jaffray and Gospodarowicz, 2015), whilst a third of these will display what is judged as a complete response. This variation in radiotherapy sensitivity is likely due to a combination of genetic, stochastic and environmental factors between individuals, and ultimately, the oncogenic mechanisms at work. That said, Kalady and colleagues suggest that the greatest biological determinant of pathological complete response in rectal cancer patients is an extended interval between the completion of neo-adjuvant radiotherapy and surgery (Kalady et al., 2009).

Crucially, predictive markers of radiotherapy sensitivity will help target treatment protocols and reduce exposure to potentially harmful radiation in patients unlikely to benefit; patients that already carry a genetic burden predisposing towards cancer. As radiation therapy remains

central to the clinical management of diverse cancers, with more than 50% of cancer patients estimated to receive some radiotherapy protocol during treatment (Baskar et al., 2012), markers helping guide these treatments would ultimately help save more lives.

3.1.1 ATM: biochemistry and mechanisms of action

The focus of our investigations in this Chapter, is the serine/threonine protein kinase, *ATM*; also known as *ataxia-telangiectasia mutated*, due to its pathological association with disease, as we will discuss. In this section, we will review what is known about the protein's function, and consider some specific examples in which the gene/protein has been implicated in cancer pathology, listing clinical applications where appropriate.

ATM is a very highly-conserved gene, being found in *Arabidopsis*, *Drosophila*, and *Xenopus*, to name some non-mammals or birds (Zdobnov et al., 2017). Across all species containing the gene, nucleotide and amino acid sequence homology is high, and across taxa the gene has been described to carry out many similar functions concerning DNA repair (Zdobnov et al., 2017). A good example comes from work in flies, where the authors report that the *ATM* orthologue, dATM, mediates the DNA damage response to ionising radiation, or that arising due to normal tissue development. An absence of dATM led to pupal or larval death (Song et al., 2004).

In humans, the *ATM* gene is located on chromosome 11 and encodes a protein of 3,056 amino acids in length, with the major isoform weighing approximately 350 kDa(Gately et al., 1998). It is member of the phosphatidylinositol 3-kinase-related kinases (PIKKs) superfamily, which has six-members (*ATM*, *ATR*, *PRKDC*, *MTOR*, *SMG1*, *TRRAP*), all of which share sequence homology with phosphatidylinositol 3-kinases (PI3Ks). PI3Ks themselves are potent intracellular signalling enzymes that regulate cellular growth, division, differentiation, apoptosis, mobility, and are themselves associated with cancer (Fruman and Rommel, 2014).

ATM has been extensively characterised, given its prominent role in human disease. The full-length protein is described to have five known functional domains, from the N- to C-terminus being: a HEAT repeat domain (composed of two alpha helices connected by a linker region that mediates binding to the C-terminus of NBS1, a protein we will discuss in its own right in later Chapters); a FRAP-ATM-TRRAP, or FAT, domain; a kinase domain; the PIKK-regulatory domain; and another short FAT domain of 30 amino acids, called FATC (Wang et al., 2016). The consensus model at present states that the FAT domain stabilises the activatory kinase domain, whose activity is regulated by both the PIKK-regulatory domain and FATC (Jiang et al., 2006). As PI3K and PIKK families are widespread and signalling molecules that participate in a range of processes, the functions of ATM are not limited to DSB, as we will discuss later.

For our interests, however, we will focus on the role of *ATM* in DSB repair. In the absence of DSBs in DNA (which may be induced by various mechanisms), ATM is held in an inactive state in cells. This inactivation is mediated by ATM's tendency to form dimers or higher-order multimers, which allows the kinase domain to associate with serine 1,981 of the FAT domain

to enter an activated state (Bakkenist and Kastan, 2003). Dimer and multimer dissociation can presumably be triggered by a range of events in the cell, and ionising radiation has been shown to cue the auto-phosphorylation of serine 1,981 required for activated ATM monomers to be produced. However, the extent of ATM triggers remains to be fully determined, and could include small molecule metabolites and a number of auxiliary factors guiding ATM to where it is required. Subsequent to de-dimersation, auto-phosphorylation of serines 367 and 1,893 drives the full activation of ATM monomers. Doses of ionising radiation as low as 0.5 Gy have been found to trigger monomer formation, as can a limited handful of DSBs, demonstrative of the fine sensitivity of the ATM sensor (Kozlov et al., 2011). In the context of evolution, and given the remarkable conservation of the gene in plants and other species, it is likely that ATM represents one of the major ways in which organic life responds to the threat damaging radiation poses to the genetic code; making the gene a great candidate when it comes to evaluating the effects of neoadjuvant radiation on cancer progression; as are MRE11, NBS1 and RAD50.

To facilitate DNA repair, ATM collaborates with members of the MRN complex – MRE11, RAD50 and NBS1, the other foci of our investigations. In eukaryotic organisms, the MRN complex is responsible for detecting and processing double-strand breaks immediately prior to homologous recombination, or NHEJ (Lamarche et al., 2010). The remarkable affinity of the complex for DSBs is shown by its association with such lesions *in vitro*, implying strongly conserved structural motifs at the damaged DNA ends underlie lesion recognition. Such innate structural recognition is presumably essential for the rapid tethering of broken strands *in vivo*.

The rest of the MRN complex is responsible for recruiting ATM to the site of the DSB, where the complex may serve to tether the free ends whilst awaiting repair. A fully assembled and functional MRN complex is required to recruit ATM to DSBs (Uziel et al., 2003). This association between the MRN complex and ATM is mediated by binding of ATM to the protein MDC1 (Eliezer et al., 2014), which in turns binds to MRE11, a central MRN complex member, discussed in Chapter Four. ATM kinase activity as part of the MRN complex is dependent upon the interaction of the kinase with NBS1 via the HEAT domain (Lee and Paull, 2007). Once associated with NBS1, ATM is able to phosphorylate the serine at position 139 of H2A histone family member X (H2AX). ATR and PRKDC (other PIKK family members) are also able to carry out this phosphorylation, ensuring redundancy in this essential pathway (Stiff et al., 2004).

Phosphorylation of H2AX by ATM is likely to have a number of consequences for DNA DSB repair. Firstly, it may allow for DNA de-condensation and fewer spatial restrictions for lesion repair, and secondly, it has been shown to allow for the recruitment of additional DDR proteins to the site. This recruitment is thought to be dependent on adaptor proteins possessing a BRCT domain (often found in proteins involved in cell cycle checkpoint regulation, such as BRCA1), which are then able to recruit some familiar regulators of the cell cycle, such as the TSG p53 and checkpoint kinase 2 (CHK2) (Baldock et al., 2015).

However, histone modifications are only one aspects of ATMs functions as part of the MRN complex, and the DDR can take different paths; with many others, likely to be described in the future. However, in all cases, it is thought that conformational changes, induced first by the DNA DSB on the MRN complex, and then by the MRN complex on ATM, allow for the repair to proceed (Lee et al., 2013); again, presumably allowing a rapid response to lesions. Such changes are thought to underpin an increased affinity of ATM for some of its substrates, although the precise contacts and interactions are unknown. Recent advances in cryo-EM methodology could be particularly useful for further refining the mechanism of action of ATM, for which no crystal structure has been solved to date. However, given sequence and functional homology, it is thought to function in a similar manner to another member of the same superfamily, PRKDC, which uses a head and neck region to wrap around DNA after conformational changes in the molecule (Hill and Lee, 2010).

Regardless of the precise motifs and cues dictating ATM function, the protein is known to halt cell cycle progression in the face of DSBs via various biochemical means. In one pathway, ATM (when associated with the MRN complex) has been shown to phosphorylate CHK2, which leads to its activation (Matsuoka et al., 2000). CHK2 encodes a TSG with various functions in DNA repair, cell cycle regulation and apoptosis. After being phosphorylated by ATM, CHK2 then goes on to phosphorylate one of its key targets, the phosphatase CDC25A (Donzelli and Draetta, 2003). The phosphorylation of CDC25A by CHK2 leads to its degradation by the proteasome. In the absence of CDC25A, CDK2-cyclin remains in a phosphorylated state of its own, cueing cell cycle arrest at the G1/S transition (Donzelli and Draetta, 2003).

However, given the rapid kinetics and dynamic of the intracellular environment, where lesions can just as easily progress as be fixed, it is perhaps not surprising to learn that ATM can further delay cell cycle progression. This is achieved via the effects ATM has on two additional substrates, p53 and MDM2 (Khosravi et al., 1999). Phosphorylation of *p53* by ATM (and the ATM target CHK2) leads to its stabilisation and the engagement of gene expression programmes designed to keep cells in a quiescent state or undergo apoptosis (Morgan and Kastan, 1997). Indeed, if DSB repair is unsuccessful, ATM is able to cue apoptosis by activating p53 and BID, a BCL2 family member, and inhibiting MDM2 and MDMX (Kamer et al., 2005). ATM also participates in the recruitment of DNA ligase IV to the DSB lesion site, allowing closure of the break (Jackson, 2002).

The *ATM* gene is also known to have a number of other important functions in the cell worth considering before evaluating the results with regards to the DDR and rectal cancer. For example, *ATM* has recently been found to be an important mediator of mitochondrial autophagy, where old organelles are recycled and new ones produced (Valentin-Vega and Kastan, 2012). *ATM* is also known to be important for meiotic prophase, and is highly expressed in testes and oocytes, compared to somatic cells (Plug et al., 1997). In germ cells, ATM is known to repair DSBs (along with other DDR proteins), although its expression in these cells declines with age, allowing for mutation accumulation over time. An absence of the gene causes premature ageing of reproductive germ cells and infertility in man and other

mammals (Elson et al., 1996). Furthermore, as ATM has a wide range of substrates, defects in its activity have far-reaching, debilitating consequences. To date, ATM has been reported to be involved in apoptosis (Maclean et al., 2007), translation initiation (So and Ouchi, 2014), insulin signalling (Yang and Kastan, 2000), gene expression regulation (Heinloth et al., 2003), telomere maintenance (Tong et al., 2015), and G_1/S and G_2/M check-point control (Goodarzi et al., 2003), to name but a few known functions of the gene. Some key target proteins of ATM kinase activity are shown in Figure 3.1.

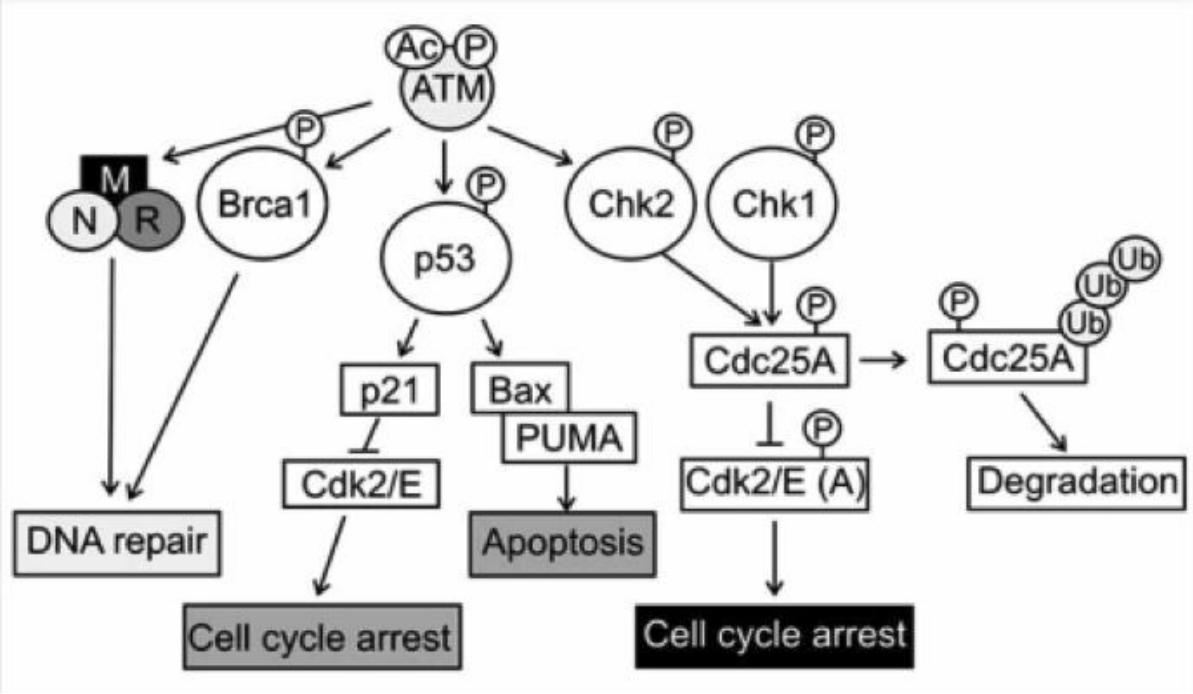


Figure 3-1 A schematic representing ATM and some of its best-known target proteins that halt cancer development.
Image taken from Oberle and Blattner, 2010.

3.1.2 Associations between ATM and cancer

As alluded to in the previous section, ATM is known to be central to many cancers. Indeed, ATM was so named for its association with Ataxia telangiectasia, a rare, autosomal recessive human disease in which patients exhibit marked radiation sensitivity, cerebellar degeneration, and a predisposition to cancer (Chun and Gatti, 2004). Indeed, the discovery of extreme radiation sensitivity in A-T patient cells first stimulated the intense research that has accompanied this gene for the past three decades. In all cases, Ataxia telangiectasia is caused by mutations in ATM, with clinical symptoms typically manifesting first during early childhood, where affected patients first appear to lack co-ordination, and develop oculomotor apraxia, dysphagia and infections of the respiratory tract, before succumbing to systemic disease burden (Rothblum-Oviatt et al., 2016). Importantly, A-T shares many

pathophysiological features with cancer, not least because patients with the disease struggle to repair DSBs.

To date several different types of cancer have been associated with mutations in, or aberrant activity of, ATM, including: Mantle cell lymphoma, T cell lymphoma, and B cell chronic lymphocytic leukaemia (Stankovic and Skowronska, 2014). If spontaneous cancers are considered – i.e. those in patients not known to have a family history of the disease, and as surveyed by the Catalogue of Somatic Mutations in Cancer – heterozygous ATM mutations have been found in approximately 11% of haematopoietic cancers, 7% of lung cancers, 5% of colon cancers, and 2% of both kidney and lung cancers (Forbes et al., 2017). ATM allelic imbalance and p53 gene mutations have also been found to occur during the progression from diploid to aneuploid cell populations in multiploid colorectal carcinomas (Sugai et al., 2001).

The precise means by which ATM lesions predispose to cancer are likely to be different depending on the tissue of origin and the other variants driving oncogenesis. For example, recent work concerning pancreatic ductal adenocarcinoma, demonstrates ATM to be essential for avoiding the increased epithelial-mesenchymal transition that predisposes to such cancers in mice and man (Russell et al., 2015), highlighting the tissue-specific phenotypes and dysregulation arising from ATM variants. Subsequent work has shown pancreatic lesions deficient for ATM to have a higher mutation burden and that disruption of the gene accelerates *Kras*-mediated oncogenesis, without altering the tumour phenotype (Drosos et al., 2017).

Given such associations and what is known about the gene, many studies have attempted to explore whether ATM variants lead to phenotypic differences in cancer phenotypes, similar to the approach we have taken. For example, Balleine and colleagues undertook histopathological analyses of breast cancer carried the *ATM* variants, IVS10-6T → G, 2424V →G or 1420L→F (and lacked known disease variants in *BRCA1* and *BRCA2*), to determine whether clinically-relevant differences could be determined (Balleine et al., 2006). However, in contrast to the results we will present in the following sections, the authors did not find any significant differences in histopathological features between ATM variant breast cancers, and those lack such lesions. This is in agreement with other studies that conclude IVS10-6T → G does not confer an increased breast cancer risk, with germ-line variants in ATM only occurring rarely familial breast cancer cases (Szabo et al., 2004). Balleine*et al* did, however, find the ATM tumour phenotype to be different to that of tumours carrying BRCA1/2 mutations, demonstrating that different oncogenic mechanisms are at play depending on the causal lesion.

Of direct relevance to our work is that of Heike Grabsch and colleagues, who show that tumour histology with panel of two anti-ATM and -BRCA1 antibodies can predict survival in colorectal cancer patients (Grabsch et al., 2006); remembering that ATM regulates *BRCA1* and *BRCA2* expression after DSB recognition. In this study, tumours from 330 patients with colorectal adenocarcinoma were analysed, with the authors reporting increased survival in patients in whom ATM expression was highest – equating to an approximate 20% difference in survival 12 years post-surgery. Furthermore, patients receiving adjuvant radiotherapy were also found to have a higher disease-free survival interval if ATM expression was increased.

In a recent study from Feng and colleagues, low ATM expression in the tumour and tumour-associated stroma both served as independent prognostic factors in breast cancer (Feng et al., 2015). More specifically, in hormone-negative and -positive breast cancer, patients with low ATM expression levels showed reduced survival times, compared to patients with high levels of expression. These associations were found to be independent of tumour size and draining lymph node status (Feng et al., 2015). The authors suggest that low levels of ATM allow the cell division to progress in the face of accumulating DSBs, and the study demonstrates the potential utility of ATM as a prognostic cancer marker. In this case, it is possible that oncogenic variants in the gene lead to low protein expression, of the protein's function is suppressed by other mechanisms at work in the tumour – both of which would facilitate cancer development.

Aside from ATM lesions characterised directly from patient tumours, studies in cell lines have also shed light on the role of ATM in the aetiology of colorectal cancer. In the work of Ejima and colleagues, 50 sequence alterations in the *ATM* gene region were identified in 16 colon tumour cell lines (Ejima et al., 2000). In the five lines tested that displayed microsatellite instability, the most common lesions identified (accounting for 62% of identified lesions in these lines) were deletions within the intronic mononucleotide tracts of ATM. These deletions were found to disrupt ATM splicing before exons 8 and 12, leading to reduced levels of the protein – highlighting that ATM is a target of microsatellite instability in colorectal cancer.

Finally, apart from what we know about changes in the base pair sequence altering ATM function, epigenetic mechanisms affecting the *ATM* gene region have also been associated with oncogenesis. For example, 73% of brain tumours have been found to contain the gene promoter in a hyper-methylated state (Mehdipour et al., 2015), leading to reduced levels of ATM mRNA, and a blunted DDR that facilitates unregulated division. Similar observations were made in early-stage breast cancer samples, where between 53-78% of all tumours from a heterogeneous cohort were found to have a hyper-methylated *ATM* promoter (Delmonico et al., 2015), and in squamous cell carcinoma of the head and neck where 42% of samples showed the phenotype (Bolt et al., 2005).

Together, these associations demonstrate the importance of the gene with regards to maintaining genomic integrity and cellular fidelity, and that it can serve an important predictive value in different cancers – suggestive of common genetic mechanisms. Although somatic cancer-causing mutations in ATM are much less common than in other important TSG genes, such as p53 (Vogelstein et al., 2013) (probably due to the redundancy in the DDR), they remain an important contributor to pathology. As advancement of cancer cataloguing projects continue – at ever higher genetic and epigenetic resolution in large numbers of clinically well-characterised patients (Hudson et al., 2010; Stratton et al., 2009; Wheeler and Wang, 2013) – we will further our understanding of the genes regulating different types of cancers, and how they interact to cause disease.

3.2 Hypothesis and study overview

Given the prominent role of ATM in mediating DNA DSB repair, we hypothesise that variants within the gene that reduce protein expression levels in rectal tumours, will hamper DSB repair, increase genomic instability, and result in more tumour cell death following radiotherapy.

The expression of ATM in tumours (as measured by immunohistochemistry) was compared to the mismatch repair proteins, MLH1, MSH2, MSH6 and PMS2; deficiency of these proteins is responsible for microsatellite instability in colorectal cancer.

A primary interest is determining whether ATM expression in tumours can be used as a corollary marker for radiation sensitivity in rectal cancer patients. Furthermore, we investigate any associations between ATM expression and: available clinicohistopathological data; mismatch repair protein expression; and survival outcomes in the cohort.

*Part of the results here described has been presented previously at a scientific meeting prior to the completion of this thesis. I contributed to the design and analysis of these experiments. Citation: Revoltar, M., Shin, J., Lim, S., Tut, T., Dissanayake, I., Descallar, J., **Ho, V**., Chua, W., Ng, W., Lee, M., et al. (2016). Early marker of DNA damage response, atm, as a predictor of clinical outcome following radiotherapy in rectal cancer patients. Pathology *48*, s153.

3.3 Results

3.3.1 Patient Populations

Cohort characteristics for this study are laid out in Table 3.1. All samples represented tumours isolated from rectal cancer patients during surgery, as previously described.

The median age of participants in this study was 71 years (and ranged between 35 to 100), making this a predominantly elderly cohort. Notwithstanding the age associations with cancer, it is important to remember that younger cells have a lower mutation burden, and any associations here reported for ATM expression in older adults require validation in cohorts with lower mean ages.

In total, one hundred and seventy-five out of 263 (66.5%) patients analysed were male, whilst 88 (33.5%) were female, a consideration discussed in detail in the Discussion Chapter. Seventy-six patients (from 245 for which data is available; although missing data can be a problem for analyses in limited sample sizes, this dataset on the whole benefits from detailed patient annotations) (31.0%) received a radiotherapy treatment protocol during the course of their disease, and 54 of these (71.1%) also received preoperative radiotherapy (i.e. prior to sample collection). A chemotherapy regimen was used in 98/219 (44.7%) patients involved in the study, and disease-free and overall survival data were available for 215 and 248 patients, respectively.

Local recurrence of disease was documented to have occurred in 82/215 (38.1%) patients in the study, with the median time to recurrence being 2.12 years. At the time of our study, 141/248 (56.9%) patients were alive, and the median time to death in the cohort was 2 and a half years following surgery (ranging between 0 − 11.1 years). Amongst those patients who received preoperative radiotherapy, local recurrence occurred in 23/54 (42.6%) patients, and the median time to recurrence was found to be 2.08 years (ranging between 37 days − 10.5 years). Of these, 21/23 had died by the completion of the study. Median time to death following recurrence was 3.81 years (range 0.6−10.9 years).

Table 3-1 Patient cohort characteristics

	All Patients	**Preoperative Radiotherapy Group**
Total, *n*	263	54
Sex, *n* (%)		
Male	175 (66.5)	38 (70.4)
Female	88 (33.5)	16 (29.6)
Mean age, yrs.	71.0	66.6
pT category, *n* (%)		
T1-2	86/257 (33.5)	17/54 (31.5)
T3-4	171/257 (66.5)	37/54 (68.5)
pN category, *n* (%)		
N0	137/256 (53.5)	28/54 (51.9)
N1-2	119/256 (46.5)	26/54 (48.1)
pM category, *n* (%)		
M0	220/237 (92.8)	52/53 (98.1)
M1	17/237 (7.2)	1/53 (1.9)
Grade, *n* (%)		
1-2	243/263 (92.4)	50/54 (92.6)
3	20/263 (7.6)	4/54 (7.4)

Vascular invasion, *n* (%)		
Absent	198/260 (76.2)	46/54 (85.2)
Present	62/260 (23.8)	8/54 (14.8)
Perineural invasion, *n* (%)		
Absent	218/260 (83.8)	41/54 (75.9)
Present	42/260 (16.2)	13/54 (24.1)
Radiotherapy, *n* (%)		
Total	76/245 (31.0)	
Neoadjuvant	54/76 (71.1)	54/54 (100)
Adjuvant	22/76 (28.9)	0/54 (0)
Tumour regression grade, *n* (%)		
0	-	0/51 (0)
1-2	-	9/51 (17.6)
3	-	42/51 (82.4)
Chemotherapy, *n* (%)		
Total	98/219 (44.7)	38/54 (70.4)
Neoadjuvant	38/98 (38.8)	31/38 (81.6)
Adjuvant	67/98 (68.4)	27/38 (71.1)
Surgery alone, *n* (%)	130/218 (59.6)	-
Median follow up, years	3.2	3.2

3.3.2 ATM Expression

ATM protein in tumour tissue sections was immuno-stained and analysed in 259 central, and 260 peripheral cores.

Two hundred and eighteen samples (from a total of 259; 84%) taken from the central core of the tumour displayed positive ATM expression; quantitated as previously described in Chapter 2. Negative expression was scored in 41/259 (16%) tumour central cores. Samples from peripheral cores were positive in 205/260 (79%) samples, and negative in 55/260 (21%).

Representative ATM staining examples are shown in Figure 3.2.

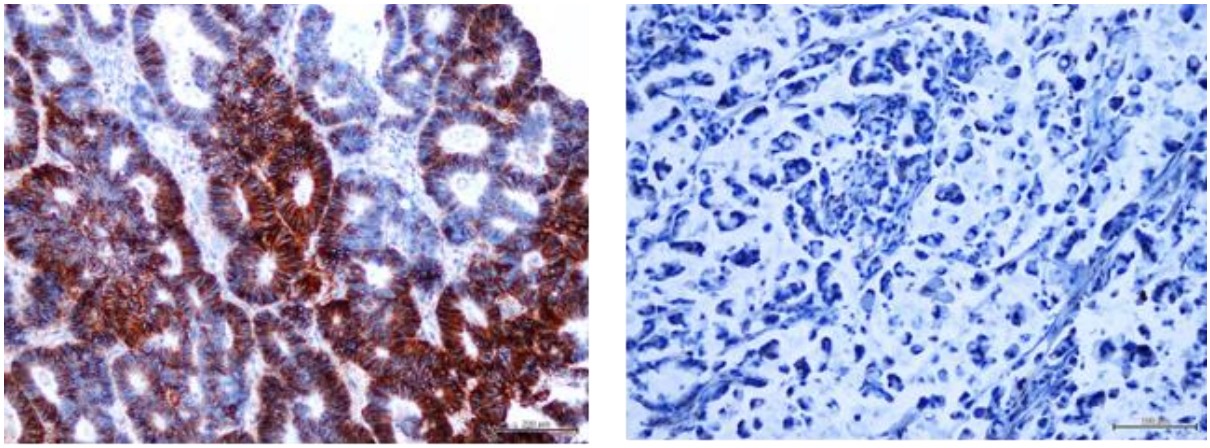


Figure 3-2 Representative positive and negative ATM staining in rectal cancer samples from our patient cohort.

Positive ATM staining is shown by the brown colour present in tumour cells, whereas an absence of staining is shown by the right-hand panel.

Based on continuous variable scoring, the observed expression differences between the tumour central and peripheral tissues was found to be significant (M_{A-B}=1.56, t (256) = 4.40, P< 0.001). ATM expression was found to be higher in the tumour centre than periphery. Given these differences, available clinicohistopathological and clinical outcome data were analysed separately for associations between central and peripheral tumour ATM expression.

In order to accurately quantitate ATM staining, and to have a biological comparator in the experiment, we stained ATM using the same protocol in 250 tumour-adjacent tissue samples, and 228 distal samples from the same patients. ATM expression was observed in 199/250 (80%) tumour-adjacent cores, and in 204/228 (89%) distal cores. ATM expression was found to be higher in samples in closer proximity to the tumour (M_{A-B} = 2.19, t (218) = 4.59, P< 0.001), suggestive of a spreading mutation burden associated with the malignancy.

In support of this, ATM expression was also found to be significantly higher in both central and peripheral tumour samples than in distal healthy mucosa (M_{A-B} = 2.80, t (246) = 5.79, P < 0.001 and M_{A-B} = 1.16, t (247) = 2.75, P = 0.006, respectively).

3.3.3 Association between ATM and mismatch repair protein expression

To evaluate the association between ATM and other mis-match repair proteins, MLH1, MSH2, MSH6 and PMS2 were stained for in the same samples. Unfortunately for our analyses, in our cohort there were no MSI-high (MMR-negative) cases of rectal cancer. All cases were positive for MLH1 and MSH2 expression. Therefore, these two proteins were excluded from all further analyses. Expression of MSH6 and PMS2 was found to be negative in only 2/257 (0.8%) and 9/253 (3.6%) cases, respectively. No associations between MSH6 or PMS2 expression and ATM expression, in either central or peripheral tumour samples, were found in this dataset (P> 0.1).

3.3.4 Associations between ATM expression and clinicohistopathological variables

To further explore the implications of ATM expression levels and rectal cancer, we next sought to determine whether the immunohistochemical score (product of proportion and intensity) correlated with known clinicohistopathological variables.

Firstly, we found negative ATM expression in the tumour periphery to be associated with older age (χ^2 = 6.21 (1, n = 260), P = 0.013) and higher grade of disease (P = 0.044) (Table 3.2), again suggesting that older age is associated with increased DNA damage.

However, expression of ATM in tumour centres was not found to be associated with any clinicohistopathological variables associated with the patient samples. Additionally, we did not observe there to be any correlations with sex, TNM category, vascular invasion, or perineural invasion.

Table 3-2 Associations between ATM expression and clinicohistopathological data

	Tumour centre			**Tumour periphery**		
	Negative ***n*** **(%)**	**Positive** ***n*** **(%)**	***P***	**Negative** ***n*** **(%)**	**Positive** ***n*** **(%)**	***P***
Sex						
Male	26 (15)	146 (85)	0.658	36 (21)	136 (79)	0.902
Female	15 (17)	72 (83)		19 (22)	69 (78)	

Age						
≤70 yrs.	17 (14)	103 (86)	0.496	17 (14)	102 (86)	**0.013**
>70 yrs.	24 (17)	115 (83)		38 (27)	103 (73)	
pT category						
T1-2	13 (16)	71 (84)	0.824	20 (24)	64 (76)	0.417
T3-4	28 (17)	141 (83)		33 (19)	137 (81)	
pN category						
N0	19 (14)	116 (86)	0.401	30 (22)	104 (78)	0.667
N1-3	21 (18)	96 (82)		24 (20)	95 (80)	
pM category						
M0	38 (18)	178 (82)	0.083	45 (21)	173 (79)	1.000
M1	0 (0)	17 (100)		3 (18)	14 (82)	
Grade						
1-2	36 (15)	203 (85)	0.333	47 (20)	193 (80)	**0.044**
3	5 (25)	15 (75)		8 (40)	12 (60)	
Vascular invasion						
Absent	32 (17)	162 (83)	0.712	45 (23)	150 (77)	0.150
Present	9 (15)	53 (85)		9 (15)	53 (85)	
Perineural invasion						
Absent	37 (17)	177 (83)	0.210	48 (22)	167 (77)	0.242
Present	4 (10)	38 (90)		6 (14)	37 (86)	

Tumour regression grade 1-2 3	2 (25) 7 (17)	6 (75) 34 (83)	0.628	5 (56) 8 (19)	4 (44) 34 (81)	**0.036**

3.3.5 Associations between ATM expression and Tumour Regression Grade

Next, we sought to evaluate the relationship between ATM expression and TRG following radiotherapy, helping us determine whether expression of the protein can be an informative marker of radiotherapy outcomes.

TRG scores were available for 51/54 (94.4%) patients in the study cohort. Of these, 42 (82.4%) showed a poor clinical response to radiotherapy, and nine (17.6%) showed a minimal-to-moderate response (Table 3.1). No tumours displayed a complete response to radiotherapy in these patients.

In these patients, we found positive ATM expression in the tumour periphery to be associated with higher TRG (i.e. a poorer response to radiotherapy) ($P = 0.036$). This relationship was not observed between ATM expression in central tumour cores (Table 3.2).

3.3.6 Associations between ATM and MMR protein expression, clinicohistopathological variables, and disease-free and overall survival: in patients receiving preoperative radiotherapy

As discussed previously, predictive biomarkers hold much promise for cancer diagnosis and treatment. Accordingly, we next sought to determine whether ATM expression in rectal tumours was associated with clinical outcome measures, namely survival.

To this end, disease-free and overall survival outcomes were analysed in the 55 patients in the cohort that had received preoperative radiotherapy (Figure 3.3). A longer disease-free survival interval was found in patients with negative ATM expression in the tumour periphery ($P = 0.025$). However, negative ATM expression in the tumour centre was not associated with survival measures ($P = 0.140$). Nonetheless, the association between disease-free survival and negative ATM expression at both sites was found to be significant after adjusting for known confounders in our multivariate analysis (Tables 3.3 and 3.4) [ATM, tumour centre (HR = 6.948 (1.192−40.504), $P = 0.031$); and ATM, tumour periphery (HR = 34.636 (2.160−555.293), $P = 0.012$)]. Overall survival was not found to be affected by ATM expression in either central ($P = 0.203$) or peripheral (P=0.570) tumour cells.

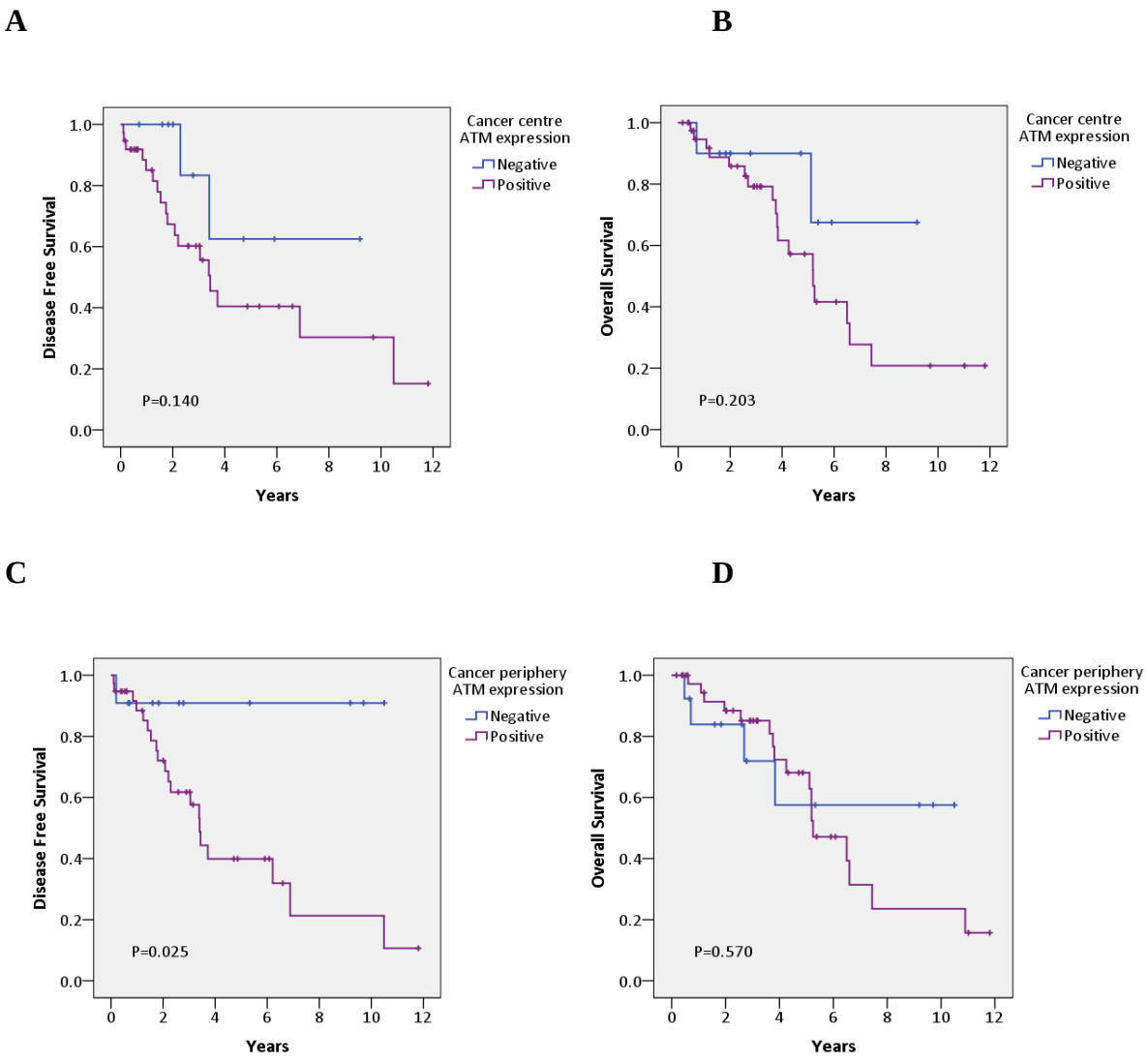


Figure 3-3 ATM expression associations with DFS and OS. Association between ATM expression in the tumour centre and (A) disease-free survival and (B) overall survival of patients. Association between ATM expression in the tumour periphery and (C) disease-free survival and (D) overall survival of patients.

Table 3-3 Multivariate analysis of ATM expression with disease-free survival in patients who received preoperative radiotherapy

	Disease-free survival	
	HR (95% CI)	***P***
ATM, tumour centre Positive vs. negative	6.948 (1.192−40.504)	**0.031**
Grade 3 vs. 1-2	22.167 (3.086−159.215)	**0.002**
Vascular invasion Presence vs. absence	9.216 (2.506−33.884)	**0.001**
	Disease-free survival	
	HR (95% CI)	***P***
ATM, tumour periphery Positive vs. negative	36.717 (2.103− 641.065)	**0.014**
Grade 3 vs. 1-2	92.465 (7.039− 1214.575)	**0.001**
Vascular invasion Presence vs. absence	5.735 (1.656− 19.863)	**0.006**

Table 3-4 Multivariate analysis of ATM expression with overall survival in patients who received preoperative radiotherapy

	Overall survival	
	HR (95% CI)	***P***
ATM, tumour centre Positive vs. negative	2.748 (0.612−12.337)	0.187
Sex Male vs. female	0.338 (0.118−0.970)	**0.044**
pM category M1 vs. M0	11.705 (1.240−110.498)	**0.032**
Grade 3 vs. 1-2	9.388 (2.180−40.435)	**0.003**
	Overall survival	
	HR (95% CI)	**P**
ATM, tumour periphery Positive vs. negative	4.256 (0.994−18.222)	0.051
Sex Male vs. female	0.267 (0.091−0.784)	**0.016**
pM category M1 vs. M0	63.865 (4.007−1017.971)	**0.003**
Grade 3 vs. 1-2	26.337 (4.567−151.891)	**<0.001**

Multivariate analysis in patients receiving preoperative radiotherapy, revealed that a higher histological grade negatively impacted disease-free survival [ATM, tumour centre [(HR =

22.167 (3.086−159.215), *P* = 0.002) vs. ATM, tumour periphery (HR = 58.640 (4.867−706.503), *P* =0.001)]. A higher histological grade also negatively impacted overall survival [ATM, tumour centre (HR = 9.388 (2.180−40.435), *P* = 0.003) and ATM, tumour periphery (HR = 26.337 (4.567−151.891), *P*<0.001)].

Furthermore, the presence of vascular invasion was found to be associated with a poorer disease-free survival interval [ATM, tumour centre (HR = 5.735 (1.656, 19.863), *P* =0.006) and ATM, tumour periphery (HR = 9.216 (2.506−33.884), *P* = 0.001)], and males were also found to exhibit longer overall survival than females [ATM, tumour centre (HR = 0.338 (0.118−0.970), *P* = 0.044) and ATM, tumour periphery (HR = 0.267 (0.091−0.784), *P* = 0.016)].

Finally, the presence of metastatic disease significantly reduced overall survival (ATM, tumour centre (HR=11.705 (1.240−110.498), P=0.032) and ATM, tumour periphery (HR=63.865 (4.007−1017.971), P=0.003)].

3.4 Concluding remarks

In these studies, we have demonstrated that there is a higher level of ATM expression in rectal tumour cells compared to normal mucosa, suggestive of genomic instability in the lesion. Furthermore, we found ATM expression to be lower in the tumour periphery than in the centre, and tumours with negative ATM expression in the periphery displayed a better response to radiotherapy. Finally, low ATM expression in peripheral tumour cells was also found to be associated with improved disease-free survival, and expression patterns from central tumour cells were also found to be significant after adjusting for confounders, demonstrating the biomarker potential of ATM in this scenario.

In general, in this dataset, the peripheral tumour tissue proved to be more informative than the central cores, which is representative of the distinct microenvironments and mechanisms present in the evolving tumour.

Unfortunately, given the extremely low frequency of MSI in the patient cohort studied, associations between dysregulated MMR and ATM expression, as previously reported, could not be undertaken.

3.5 References

Bakkenist, C.J., and Kastan, M.B. (2003). DNA damage activates ATM through intermolecular autophosphorylation and dimer dissociation. Nature *421*, 499–506.
Baldock, R.A.A., Day, M., Wilkinson, O.J.J., Cloney, R., Jeggo, P.A.A., Oliver, A.W.W., Watts, F.Z.Z., and Pearl, L.H.H. (2015). ATM Localization and Heterochromatin Repair Depend on Direct Interaction of the 53BP1-BRCT2 Domain with γH2AX. Cell Rep. *13*, 2081–2089.
Balleine, R.L., Murali, R., Bilous, A.M., Farshid, G., Waring, P., Provan, P., Byth, K., Thorne, H., and Kirk, J.A. (2006). Histopathological features of breast cancer in carriers of ATM gene variants. Histopathology *49*, 523–532.
Bartkova, J., Hořejší, Z., Koed, K., Krämer, A., Tort, F., Zieger, K., Guldberg, P., Sehested, M., Nesland, J.M., Lukas, C., et al. (2005). DNA damage response as a candidate anti-cancer barrier in early human tumorigenesis. Nature *434*, 864–870.
Baskar, R., Lee, K.A., Yeo, R., and Yeoh, K.W. (2012). Cancer and radiation therapy: current advances and future directions. Int J Med Sci *9*, 193–199.
Bolt, J., Vo, Q.N., Kim, W.J., McWhorter, A.J., Thomson, J., Hagensee, M.E., Friedlander, P., Brown, K.D., and Gilbert, J. (2005). The ATM/p53 pathway is commonly targeted for inactivation in squamous cell carcinoma of the head and neck (SCCHN) by multiple molecular mechanisms. Oral Oncol. *41*, 1013–1020.
Chan, I.S., and Ginsburg, G.S. (2011). Personalized medicine: progress and promise. Annu. Rev. Genomics Hum. Genet. *12*, 217–244.
Choudhury, A., Nelson, L.D., Teo, M.T.W., Chilka, S., Bhattarai, S., Johnston, C.F., Elliott, F., Lowery, J., Taylor, C.F., Churchman, M., et al. (2010). MRE11 expression is predictive of cause-specific survival following radical radiotherapy for muscle-invasive bladder cancer. Cancer Res. *70*, 7017–7026.
Chun, H.H., and Gatti, R.A. (2004). Ataxia-telangiectasia, an evolving phenotype. DNA Repair (Amst). *3*, 1187–1196.
Delmonico, L., Moreira, A.D.S., Franco, M.F., Esteves, E.B., Scherrer, L., Gallo, C.V.D.M., Nascimento, C.M. Do, Ornellas, M.H.F., De Azevedo, C.M., and Alves, G. (2015). CDKN2A ($p14^{ARF}$/$p16^{INK4a}$) and ATM promoter methylation in patients with impalpable breast lesions. Hum. Pathol. *46*, 1540–1547.
Di-Paolo, A., Sarkozy, F., Ryll, B., and Siebert, U. (2017). Personalized medicine in Europe: not yet personal enough? BMC Health Serv. Res. *17*, DOI: 10.1186/s12913-017-2205-4.
Donzelli, M., and Draetta, G.F. (2003). Regulating mammalian checkpoints through Cdc25 inactivation. EMBO Rep. *4*, 671–677.
Drosos, Y., Escobar, D., Chiang, M., Roys, K., Valentine, V., Valentine, M., Rehg, J., Sahai, V., Begley, L., Ye, J., et al. (2017). ATM-deficiency increases genomic instability and metastatic potential in a mouse model of pancreatic cancer. Sci Rep *7*, 11144.
Ejima, Y., Yang, L., and Sasaki, M. (2000). Aberrant splicing of the ATM gene associated with shortening of the intronic mononucleotide tract in human colon tumor cell lines: a novel mutation target of microsatellite instability. Int J Cancer *15*, 262–268.
Eliezer, Y., Argaman, L., Kornowski, M., Roniger, M., and Goldberg, M. (2014). Interplay between the DNA damage proteins MDC1 and ATM in the regulation of the spindle assembly checkpoint. J. Biol. Chem. *289*, 8182–8193.
Elson, A., Wang, Y., Daugherty, C.J., Morton, C.C., Zhou, F., Campos-Torres, J., and Leder, P. (1996). Pleiotropic defects in ataxia-telangiectasia protein-deficient mice. Proc. Natl. Acad. Sci. *93*, 13084–13089.
Feng, X., Li, H., Dean, M., Wilson, H.E., Kornaga, E., Enwere, E.K., Tang, P., Paterson, A., Lees-Miller, S.P., Magliocco, A.M., et al. (2015). Low ATM protein expression in malignant tumor as well as cancer-associated stroma are independent prognostic factors in a retrospective study of early-stage hormone-negative breast cancer. Breast Cancer Res. *17*, 65.

Forbes, S.A., Beare, D., Boutselakis, H., Bamford, S., Bindal, N., Tate, J., Cole, C.G., Ward, S., Dawson, E., Ponting, L., et al. (2017). COSMIC: Somatic cancer genetics at high-resolution. Nucleic Acids Res. *45*, D777–D783.
Fruman, D.A., and Rommel, C. (2014). PI3K and cancer: lessons, challenges and opportunities. Nat. Rev. Drug Discov. *13*, 140–156.
Gately, D.P., Hittle, J.C., Chan, G.K., and Yen, T.J. (1998). Characterization of ATM expression, localization, and associated DNA-dependent protein kinase activity. Mol. Biol. Cell *9*, 2361–2374.
Goodarzi, A., Block, W.D., and Lees-Miller, S.P. (2003). The role of ATM and ATR in DNA damage-induced cell cycle control. Prog. Cell Cycle Res. *5*, 393–411.
Grabsch, H., Dattani, M., Barker, L., Maughan, N., Maude, K., Hansen, O., Gabbert, H.E., Quirke, P., and Mueller, W. (2006). Expression of DNA double-strand break repair proteins ATM and BRCA1 predicts survival in colorectal cancer. Clin. Cancer Res. *12*, 1494–1500.
Heinloth, A., Shackelford, R., Innes, C., Bennett, L., Li, L., Amin, R., Sieber, S., Flores, K., Bushel, P., and Paules, R. (2003). ATM-dependent and -independent gene expression changes in response to oxidative stress, gamma irradiation, and UV irradiation. Radiat Res *160*, 273–290.
Hill, R., and Lee, P.W.K. (2010). The DNA-dependent protein kinase (DNA-PK): More than just a case of making ends meet? Cell Cycle *9*, 3460–3469.
Hudson, T.J., Anderson, W., Aretz, A., and Barker, A.D. (2010). International network of cancer genome projects. Nature *464*, 993–998.
Jackson, S.P. (2002). Sensing and repairing DNA double-strand breaks. Carcinogenesis *23*, 687–696.
Jaffray, D., and Gospodarowicz, M. (2015). Radiation Therapy for Cancer. In Cancer: Disease Control Priorities, Third Edition, (Washington DC: The International Bank for Reconstruction and Development / The World Bank), 1-16. Available from: http://www.dcp-3.org/sites/default/files/chapters/CANCER%20ch14%20Radiation%20Therapy.pdf
Jiang, X., Sun, Y., Chen, S., Roy, K., and Price, B.D. (2006). The FATC domains of PIKK proteins are functionally equivalent and participate in the Tip60-dependent activation of DNA-PKcs and ATM. J. Biol. Chem. *281*, 15741–15746.
Kalady, M.F., de Campos-Lobato, L.F., Stocchi, L., Geisler, D.P., Dietz, D., Lavery, I.C., and Fazio, V.W. (2009). Predictive factors of pathologic complete response after neoadjuvant chemoradiation for rectal cancer. Ann. Surg. *250*, 582–589.
Kamer, I., Sarig, R., Zaltsman, Y., Niv, H., Oberkovitz, G., Regev, L., Haimovich, G., Lerenthal, Y., Marcellus, R.C., and Gross, A. (2005). Proapoptotic BID is an ATM effector in the DNA-damage response. Cell *122*, 593–603.
Khosravi, R., Maya, R., Gottlieb, T., Oren, M., Shiloh, Y., and Shkedy, D. (1999). Rapid ATM-dependent phosphorylation of MDM2 precedes p53 accumulation in response to DNA damage. Proc. Natl. Acad. Sci. U. S. A. *96*, 14973–14977.
Kozlov, S. V., Graham, M.E., Jakob, B., Tobias, F., Kijas, A.W., Tanuji, M., Chen, P., Robinson, P.J., Taucher-Scholz, G., Suzuki, K., et al. (2011). Autophosphorylation and ATM activation: Additional sites add to the complexity. J. Biol. Chem. *286*, 9107–9119.
Lamarche, B.J., Orazio, N.I., and Weitzman, M.D. (2010). The MRN complex in double-strand break repair and telomere maintenance. FEBS Lett. *584*, 3682–3695.
Lee, J.-H., and Paull, T.T. (2007). Activation and regulation of ATM kinase activity in response to DNA double-strand breaks. Oncogene *26*, 7741–7748.
Lee, J.H., Mand, M.R., Deshpande, R.A., Kinoshita, E., Yang, S.H., Wyman, C., and Paull, T.T. (2013). Ataxia Telangiectasia-Mutated (ATM) kinase activity is regulated by ATP-driven conformational changes in the Mre11/Rad50/Nbs1 (MRN) complex. J. Biol. Chem. *288*,

12840–12851.
Maclean, K.H., Kastan, M.B., and Cleveland, J.L. (2007). ATM deficiency affects both apoptosis and proliferation to augment Myc-induced lymphomagenesis. Mol. Cancer Res. *5*, 705–711.
Matsuoka, S., Rotman, G., Ogawa, A., Shiloh, Y., Tamai, K., and Elledge, S.J. (2000). Ataxia telangiectasia-mutated phosphorylates Chk2 in vivo and in vitro. Proc. Natl. Acad. Sci. U. S. A. *97*, 10389–10394.
Mehdipour, P., Karami, F., Javan, F., and Mehrazin, M. (2015). Linking ATM Promoter Methylation to Cell Cycle Protein Expression in Brain Tumor Patients: Cellular Molecular Triangle Correlation in ATM Territory. Mol. Neurobiol. *52*, 293–302.
Morgan, S.E., and Kastan, M.B. (1997). p53 and ATM: cell cycle, cell death, and cancer. Adv Cancer Res *71*, 1–25.
Oberle, C., and Blattner, C. (2010). Regulation of the DNA Damage Response to DSBs by Post-Translational Modifications. Curr. Genomics *11*, 184–198.
Plug, A.W., Peters, A.H., Xu, Y., Keegan, K.S., Hoekstra, M.F., Baltimore, D., de Boer, P., and Ashley, T. (1997). ATM and RPA in meiotic chromosome synapsis and recombination. Nat. Genet. *17*, 457–461.
Revoltar, M., Shin, J., Lim, S., Tut, T., Dissanayake, I., Descallar, J., Ho, V., Chua, W., Ng, W., Lee, M., et al. (2016). Early marker of DNA damage response, atm, as a predictor of clinical outcome following radiotherapy in rectal cancer patients. Pathology *48*, s153.
Rothblum-Oviatt, C., Wright, J., Lefton-Greif, M.A., McGrath-Morrow, S.A., Crawford, T.O., and Lederman, H.M. (2016). Ataxia telangiectasia: a review. Orphanet J. Rare Dis. *11*, 159.
Russell, R., Perkhofer, L., Liebau, S., Lin, Q., Lechel, A., Feld, F.M., Hessmann, E., Gaedcke, J., Güthle, M., Zenke, M., et al. (2015). Loss of ATM accelerates pancreatic cancer formation and epithelial-mesenchymal transition. Nat. Commun. *6*, 7677.
So, E.Y., and Ouchi, T. (2014). Translational initiation regulated by ATM in dendritic cells development. Cell Death Dis. *5*, e1418.
Song, Y.H., Mirey, G., Betson, M., Haber, D.A., and Settleman, J. (2004). The Drosophila ATM ortholog, dATM, mediates the response to ionizing radiation and to spontaneous DNA damage during development. Curr. Biol. *14*, 1354–1359.
Stankovic, T., and Skowronska, A. (2014). The role of ATM mutations and 11q deletions in disease progression in chronic lymphocytic leukemia. Leuk. Lymphoma *55*, 1227–1239.
Stiff, T., O'Driscoll, M., Rief, N., Iwabuchi, K., Löbrich, M., and Jeggo, P.A. (2004). ATM and DNA-PK Function Redundantly to Phosphorylate H2AX after Exposure to Ionizing Radiation. Cancer Res. *64*, 2390–2396.
Stratton, M.R., Campbell, P.J., and Andrew F, P. (2009). The cancer genome. Nature *458*, 719–724.
Sugai, T., Habano, W., Uesugi, N., Jiao, Y.F., Nakamura, S.I., Yoshida, T., and Higuchi, T. (2001). Frequent allelic imbalance at the ATM locus in DNA multiploid colorectal carcinomas. Oncogene *20*, 6095–6101.
Szabo, C.I., Schutte, M., Broeks, A., Houwing-Duistermaat, J.J., Thorstenson, Y.R., Durocher, F., Oldenburg, R.A., Wasielewski, M., Odefrey, F., Thompson, D., et al. (2004). Are ATM Mutations 7271T→G and IVS10-6T→G Really High-Risk Breast Cancer-Susceptibility Alleles? Cancer Res. *64*, 840–843.
Tong, A.S., Stern, J.L., Sfeir, A., Kartawinata, M., de Lange, T., Zhu, X.D., and Bryan, T.M. (2015). ATM and ATR Signaling Regulate the Recruitment of Human Telomerase to Telomeres. Cell Rep. *13*, 1633–1646.
Trastulli, S., Cirocchi, R., Listorti, C., Cavaliere, D., Avenia, N., Gullà, N., Giustozzi, G., Sciannameo, F., Noya, G., and Boselli, C. (2012). Laparoscopic vs open resection for rectal cancer: A meta-analysis of randomized clinical trials. Color. Dis. *14*, e277-296.

Uziel, T., Lerenthal, Y., Moyal, L., Andegeko, Y., Mittelman, L., and Shiloh, Y. (2003). Requirement of the MRN complex for ATM activation by DNA damage. EMBO J. *22*, 5612–5621.
Valentin-Vega, Y.A., and Kastan, M.B. (2012). A new role for ATM: Regulating mitochondrial function and mitophagy. Autophagy *8*, 840–841.
Vogelstein, B., Papadopoulos, N., Velculescu, V.E., Zhou, S., Diaz, L.A., and Kinzler, K.W. (2013). Cancer Genome Landscapes. Science. *339*, 1546–1558.
Wang, X., Chu, H., Lv, M., Zhang, Z., Qiu, S., Liu, H., Shen, X., Wang, W., and Cai, G. (2016). Structure of the intact ATM/Tel1 kinase. Nat. Commun. *7*, 11655.
Wheeler, D.A., and Wang, L. (2013). From human genome to cancer genome: The first decade. Genome Res. *23*, 1054–1062.
Yang, D.Q., and Kastan, M.B. (2000). Participation of ATM in insulin signalling through phosphorylation of eIF-4E-binding protein 1. Nat. Cell Biol. *2*, 893–898.
Zdobnov, E.M., Tegenfeldt, F., Kuznetsov, D., Waterhouse, R.M., Simao, F.A., Ioannidis, P., Seppey, M., Loetscher, A., and Kriventseva, E. V. (2017). OrthoDB v9.1: Cataloging evolutionary and functional annotations for animal, fungal, plant, archaeal, bacterial and viral orthologs. Nucleic Acids Res. *45*, D744–D749.

CHAPTER 4
EXPLORING ASSOCIATIONS BETWEEN MRE11 EXPRESSION AND SURVIVAL IN RECTAL CANCER

4.1 Introduction

This Chapter will further investigate the role of the DDR protein, MRE11, in rectal cancer. Similar to the ATM study, the major focus was determining whether expression of MRE11 in colorectal tumour tissue samples is associated with patient survival and related clinicohistopathological features.

MRE11 has a long history of association with colorectal cancers (Hoeijmakers, 2001). Indeed, it is known that 15-20% of all colorectal cancers carry mutations in the DNA mismatch repair genes resulting in defective mismatch repair. These cancers are typically characterised as displaying microsatellite instability (as a result of destabilisation of simple repeat sequence repair), and there is an increase in genetic mutations in microsatellite instability. These cancers are often found to harbour mutations in the DDR proteins MRE11 and RAD50, which will cover in a subsequent chapter. In a 2007 study by Miquel and colleagues, the vast majority of MSI-positive colorectal tumours had lesions leading to ablation of MRE11 expression at the protein level; allowing error-prone replication to proceed unsupervised (Miquel et al., 2007). From this study and others, it is clear that DNA metabolism, in general, is associated with colorectal cancer development and progression, with other implicated genes including *ATR*, *MSH6* and *POLD3* (Peltomaki, 2001).

4.1.1 MRE11: biochemistry and mechanism of action

Double-strand break repair protein (meiotic recombination protein 11), or MRE11, is encoded on human chromosome 11 (11q21), as is ATM. Similar to *ATM*, *MRE11* is highly-conserved amongst present day species, carrying out analogous DDR functions in yeasts, fish, plants and diverse birds, not to mention mammals (Zdobnov et al., 2017). An orthologue of MRE11 has recently been documented in the archaeon *Sulfolobusacidocaldarius*, which lives by volcanic vents at 82^{o}C and pH 2-3. When this organism is subjected to experimental radiation, Mre11 participates in DSB repair along with RAD50 (Quaiser et al., 2008); and suggests that this gene evolved from prokaryotic lineages.

In humans, MRE11 has a molecular weight between 70-90 kDa, with the predominant isoform consisting of 708 amino acids, making it less than a third the size of ATM, discussed previously. The N-terminus of the protein encodes the di-manganese-dependent phosphoesterase domain, whilst the C-terminus possesses two unique DNA-binding domains (Lamarche et al., 2010). Compared to many proteins, MRE11 has a relatively simple structure, perhaps a key feature allowing it to participate in diverse biochemical processes. As has been discussed for ATM, conformational changes are key to the protein's function; that is considered ancient and pleiotropic proteins malleable and applicable to many tasks. For example, structural data has recently revealed that a RAD50-ATP-driven conformation switch in MRE11 controls the exonuclease function of the latter (Hopfner et al., 2001). MRE11 is also known to be controlled by a number of post-translational mechanisms to influence its localisation, function and longevity. For example, in order for ATM to function in DNA damage repair, it has to be arginine methylated by PRMT1 (Boisvert et al., 2005).

The most notable functions of MRE11 are described in the context of the previously introduced MRN complex, which it forms along with RAD50 and NBS1. As part of this macromolecular assembly, MRE11 participates in the repair of DSBs using a combination of homology directed repair, and classical and alternative non-homologous end-joining. Overall, the predominant function of the MRN complex is to repair DSBs via homology directed repair between sister chromatids (Bressan et al., 1999). Furthermore, in addition to MRE11's DNA-binding roles, described below, the protein also has exo- and endo-nuclease activity against both single- and double-stranded DNA. Importantly, although MRE11 is a potent sequence modifier, it does not possess the 5'-3' exonuclease activity required to generate 3' single-strand overhangs required for homologous recombination (Lamarche et al., 2010). Therefore, auxiliary factors are likely recruited to the MRN complex to assist with homologous recombination, and to date, several 5'-3' exonucleases, such as XRN2, have been identified in man (West et al., 2004).

In-keeping with its best described functions, MRE11 has a mostly nuclear localisation, endowing it with close approximation to the genomic material it aims to repair. However, it is also able to translocate to the cytoplasm, as required, to carry out other functions. A good example is illustrated by the instructive role of MRE11 in driving type I interferon signalling and STING trafficking in response to detecting pathogen dsDNA in the cytoplasm (Kondo et al., 2013); which also highlights an important unrelated role of the protein. Although it is useful to consider the sub-cellular localisation of proteins when considering their potential functions, much remains to be determined by using more modern microscopy approaches that allow spatial and temporal dynamics to be interrogated in response to different stimuli (Lanzano et al., 2015).

The nuclear MRN complex as a whole, has a large central globular domain, where MRE11 and NBS1 associate with RAD50 via the latter's extended coil-coil and Walker A and B domains (Stracker and Petrini, 2011). It is also the globular domain that mediates nucleic acid binding, usually as part of a higher-order assembly of multiple entities (De Jager et al., 2004), and is dependent on MRE11 and RAD50, but not (according to the majority of studies) NBS1 (Schiller et al., 2014). Furthermore, it is known that dimerization of MRE11 is required for DNA binding, with dimerization being mediated by the N-terminal region of the protein; and persists in recombinant MRE11 *in vitro*(Williams et al., 2008). Unfortunately, the crystal structure of the entire assembled globular domain is yet to be determined, which would help address these points relating to function further; although again, advances in cryo-EM may overcome this limitation in the near future. With respect to MRE11, however, crystallographic data has revealed that the protein contacts DNA via 17 conserved residues distributed across DNA recognition loops, with the aforementioned residues forming minor-groove sugar-phosphate contacts (Stracker and Petrini, 2011). This dependency on sugar-phosphate contacts, and not nucleotide bases for DNA binding, allows MRE11 to bind a wide array of sequences, irrespective of the base pair sequence – a feature shared with ATM.

Given the ancient nature of the proteins discussed thus far, it is perhaps unsurprising that the DDR is a major function of MRE11. For example, in our species, and in quite remarkable

experiments, Schwartz and colleagues also found the MRN complex to also be a potent mediator of anti-viral immunity. In cells infected with parvoviruses, MRN complex dissolution led to the accumulation of MRE11, RAD50 and NBS1 on viral inverted terminal repeat regions (Schwartz et al., 2007), suggesting the complex is able to detect foreign nucleic acids as well as those of host origin. Work by Deng and colleagues has also shown multiple roles for MRE11 at the telomere end (Deng et al., 2009). In elegant work, the authors show that in the absence of TRF, MRE11 is able to remove 3' telomeric overhangs, allowing for chromosomal fusions, and protects nascent strands from NHEJ. Essentially, MRE11 is involved in sensing telomere dysfunction and maintenance. This work came on the back of that of Zhong and colleagues, who showed that an absence of the MRN complex resulted in reduced telomere maintenance (Zhong et al., 2007). This study also showed, as have others, that a knock-down of MRE11, leads to a knock-down of the entire MRN complex, highlighting the central role of this DNA binder to one of the most studied biochemical complexes. Furthermore, in mice, expression of a hypomorphic *Mre11* allele caused the premature elimination of oocytes harbouring DNA mutations, although oocyte attrition took much longer than in animals with a fully competent MRN complex (Inagaki et al., 2016). Thus, variants that alter the expression level or isoforms of MRE11 will need to be considered as putative pathological variants, and may lead to an increased oocyte and sperm mutation burden with disastrous consequences for the offspring.

Before moving on to consider some of the specific associations between MRE11 and cancer, it is worth considering the distribution of this gene's expression to learn more about its potential function and importance to cancer. According to the human tissue-wide compendium of mRNA expression, BioGPS, MRE11 mRNA is highly expressed in haematopoietic lineages, especially in cell lines derived from leukaemia and lymphoma patients (Wu et al., 2013). Highly proliferative B cells, or lymphoblasts, also express high amount of MRE11 mRNA, in keeping with its important role in genomic integrity maintenance in cells with a high turnover rate, such as those of the immune system. Adult human tissues at baseline expressed MRE11, but not significantly above mean levels; suggesting that MRE11 has more important functions in some lineages over others.

In agreement with such a distribution in mammals is data from the publicly available Immunological Genome Project (Heng and Painter, 2008) and Symatlas (Su et al., 2004). In these datasets, which analyse gene expression across the finely-dissected mouse, show the bone marrow, testis, and the embryo between days 6 and 10 of gestation, to have the highest expression levels. Interestingly, gastrointestinal inflammation, of diverse forms, is implicated with colorectal and rectal cancer development and progression (Kim and Chang, 2014; Tjalsma et al., 2012). If drivers of inflammation coincide with developmental processes typified by a high cell turnover rate, perhaps the foundations of cancer can be laid. Although much remains to be determined with regards to the aetiology of rectal cancer, immune cell dysregulation in specialised gastrointestinal microenvironments (such as Peyer's patches) can drive oncogenesis arising in other lineages (Chapkin et al., 2007; Nascimbeni et al., 2005; Sipos and Muzes, 2011).

Much remains to be determined about the aetiology of rectal cancer, although it looks

increasingly likely that individual clinical histories, and high-resolution analysis of our different organ systems (such as the microbiome) will be increasingly important in identifying causality – in line, of course, with our expanding knowledge of how genetics underpins and responds to environmental cues.

4.1.2 Associations between MRE11 and cancer

In this section, we will consider some additional and relevant recent associations discovered between MRE11 and cancer.

To date, MRE11 variants have been associated with diverse forms of cancer, including those affecting the rectum/colon, uterus, breast, bladder, or ovary (Damiola et al., 2014; Koppensteiner et al., 2014; Rebbeck et al., 2011). Indeed, missense variants of MRE11, such as 140C→T and 1773_1774delAA have been in the germ-line and have been associated with cancer development; this is not to mention the wide array of mutations reported to lead to A-T and cancer like pathophysiology (Kim et al., 2017). Furthermore, when germ-line mutations were surveyed in 5,552 colorectal cancer cases and 6,792 healthy controls, 16% of familial colorectal cancer cases were found to carry highly-penetrant variants in *POT1*, *POLE2* and *MRE11* (Chubb et al., 2016). Low-frequency alleles with moderate effects were not identified in this study. Amongst cases of colorectal cancer, 17 colorectal cancer cells lines studies typified by microsatellite instability were found to harbour mutations in the polyT(11) tract of MRE11 intro 4 (Vilar et al., 2011). When the frequency of this variant was analysed in primary tumour samples from the MECC study, 82% of samples were found to carry the same variant in intro 4; suggesting that in cases of microsatellite instability, colorectal cancer progression leads to reduced MRE11 function. The ablation of MRE11 expression in colorectal cancers positive for microsatellite instability has been well reported (Giannini et al., 2002), although the mechanisms leading to this phenotype could be diverse, as we will discuss in subsequent sections.

Very relevant to the work presented in this thesis, is that of Choudhury and colleagues, who show that MRE11 expression serves a predictive purpose in muscle invasive bladder cancer (Choudhury et al., 2010); which may represent sequelae of metastatic disease. The study samples were collected prior to radiation therapy. Specifically, the authors show that low MRE11 expression (RAD50, NBS1, ATM and H2AX levels were simultaneously reported but not associated) is associated with worse cancer survival after radiotherapy, in two independent cohorts of bladder cancer patients. The study also exploited immune-histology of archived samples to make the observation, providing strong evidence to support and pursue our investigations. When developing truly applicable biomarkers for patients, an understanding of inter-cohort heterogeneity and known independence/dependence of clinical variables is paramount (Elefsinioti et al., 2016).

Thomas Pavelitz and colleagues also undertook an investigation of the effect of MRE11 expression levels and colorectal cancer development (Pavelitz et al., 2014). Out of the 625 tumour samples analysed, 11% contained mutations in the familiar polyT (11) tract, which

could be bi-allelic or mono-allelic, and reduced expression of MRE11 protein. These patients with mutations in the T11 tract were found to have an increased long-term DFS and OS, with the study demonstrating the utility of MRE11 expression in guiding prognosis.

Finally, in a recent study taking a very similar histological approach to ours in colorectal cancer, Ihara and colleagues found MRE11 expression at the protein level was also found to predict the response to oxaliplatin-based chemotherapy in colorectal cancer patients, further demonstrating its biomarker potential (Ihara et al., 2015). In this study, low expression of MRE11 (or negative expression) was associated with significantly better tumour volume reduction compared to tumours with a high expression – in-keeping with the hypothesis that low DDR proteins make tumours more susceptible to chemo- and radio-therapy.

Together, these results provide compelling evidence to further explore associations between MRE11 expression and clinicohistopathological variables and patient outcomes in rectal cancer.

4.2 Towards the development of a ATM/MRE11 combinatorial marker panel

Whilst the study of MRE11 in tumour samples alone may be able to yield important insights into disease and prognosis, the analysis of groups of proteins together may facilitate the identification of similar and overlapping mechanisms in different patients, which may not share all features of a disease. Equally, combinatorial panels can shed light on the severity of disease, if many proteins involved in the same pathway are ablated, for example.

Ultimately, when outbred, highly heterogeneous patient populations are analysed, having two markers as opposed to one may help increase the power to detect prognostic variables; especially when commonly mutated genes, such as *ATM* and *MRE11*, are considered. Although we have seen that MRE11 and ATM both have predictive value in different cancers, neither biomarker has been tested in large-scale, randomised clinical trials; increasing the speed of biomarker discovery and validation is essential to translate basic research findings to the patient bedside.

Given our previously reported association between ATM expression and rectal cancer survival, and the results of MRE11 here presented, we sought to determine whether a combinatorial panel including both markers could serve a predictive purpose in rectal cancer. As discussed, of interest will be determining whether single- or multi-marker panels are more sensitive and specific for quantifying patient outcome measures, which we will discuss in more detail in the Discussion Chapter.

The results of the combinatorial panel experiments are shown in Chapter Six.

4.3 Hypothesis and study overview

Given the prominent role of MRE11 in mediating DNA DSB repair, we hypothesise that variants within the gene that reduce protein expression levels in rectal tumours, will hamper DSB repair, increase genomic instability, and result in more tumour cell death following radiotherapy.

We hypothesise that reduced ATM/MRE11 expression in a combinatorial panel would result in the similar outcomes to either marker alone.

We will also determine whether the expression levels of MRE11 in tumours is associated with various clinicohistopathological variables available to us as part of the dataset.

4.4 MRE Results

4.4.1 Patient Populations

A total of 262 patients (66% male, 34% female) were used to study MRE11 expression in rectal tumours. The median age of participants in this study was 71 years (ranging between 35–100 years). With regards to tumour characteristics, 33% of lesions were classified as stage T1/2, whilst 67% were classed as T3/4. Ninety-three percent were M0, 54% were lymph node negative, 92% were of tumour grade 1/2, 76% did not show evidence of vascular invasion, and 84% did not show any evidence of perineural invasion. Twenty-two percent of study the patients received neo-adjuvant radiotherapy, and 30% received adjuvant therapy. The median OS in the cohort was found to be 3.2 years (ranging between 0–12.6 years). A detailed breakdown of the cohort characteristics is shown in Table 4.1.

Table 4-1 Cohort characteristics for MRE11 and MRE11/ATM studies

	All Patients (%)	Preoperative Radiotherapy Group
Total, n	262	54
Age median	71	67
Sex		
Male	174 (66.4)	38 (70.4)
Female	88 (33.6)	16 (29.6)
Tumour stage		
T1-2	86/257 (33.5)	17/54 (31.5)
T3-4	171/257 (66.5)	37/54 (68.5)
Node stage		
N0	137/256 (53.5)	28/54 (51.9)
N1-2	118/256 (46.5)	26/54 (48.1)
Metastasis stage		
M0	220/237 (92.8)	52/53 (98.1)
M1	17/237 (7.2)	1/53 (1.9)
Grade		
1–2	242/262 (92.4)	50/54 (92.6)
3	20/262 (7.6)	4/54 (7.4)
Vascular invasion		

Absent	198/260 (76.2)	46/54 (85.2)
Present	62/260 (23.8)	8/54 (14.8)
Perineural invasion		
Absent	218/260 (83.8)	41/54 (75.9)
Present	42/260 (16.2)	13/54 (24.1)
Radiotherapy		
Total	76/245 (31.0)	-
Neoadjuvant	54/76 (71.1)	54/54 (100)
Adjuvant	22/76 (28.9)	0/54 (0)
Recurrence		
Absent	129/211 (61.1)	48/54 (88.9)
Present	82/211 (38.9)	6/54 (11.1)
Tumour regression grade		
0–2 (good response)	-	9/54 (16.7)
3 (poor response)	-	45/54 (83.3)
Chemotherapy		
Total	98/219 (44.7)	38/54 (70.4)
Neoadjuvant	38/98 (38.8)	31/38 (81.6)
Adjuvant	60/98 (61.2)	7/38 (18.4)

Patients were followed for a median period of 3.2 years during the study (ranging between 0−12.6 years). Local recurrence of disease was documented to have occurred in 82/211 (38.9%) patients, with the median time to recurrence being 2.12 years. At the time of the study, 141/248 (56.9%) patients were alive, and the median time to death was 2.5 years following surgery (ranging between 0−11.1 years). Among patients who received preoperative radiotherapy, local recurrence occurred in 6/54 (11.1%) patients, and the median time to recurrence was 2.61 years (ranging between 0.75−4.29 years). All of these six patients (100%) had died by the end of the study. The median time to death following recurrence was 3.81 years (range: 0.6 – 10.9 years).

The pre-operative tissue sampling protocol was limited by sample size. Preoperative endoscopic biopsy tissues were available for analysis in only 66 patients. Nine biopsies were derived from patients who underwent preoperative treatment (neoadjuvant radiotherapy administered either alone [n=5] or as chemoradiotherapy [n=4]), whereas eight were obtained from patients treated postoperatively with either adjuvant chemoradiotherapy (n=3) or chemotherapy alone (n=5).

4.4.2 MRE11 scores

In order to determine whether any associations existed in the dataset, the weighted scores for MRE11 staining were generated by two independent, expert scorers. Scores between 0–5 were considered to represent *low* staining, whilst scores of 6–12 were considered to be *high* staining samples. The mean weighted MRE11 scores were 5.5 for the TC, and 5.8 for the TP. This difference in weighted scores between the TC and TP was found to statistically significantly (paired t-test, $P< 0.05$), although the difference between both sites is relatively small.

The distribution of weighted scores in the TC and TP is shown in Figure 4.1. In the TC, 53% (n = 136) of patients were defined as having a low staining score, whilst 47% (n = 119) were found to have a high staining intensity. In the TP, 44% (n = 115) of participants had a low score, and 56% (n = 144) were found to have a high score.

In order to compare expression between the tumour and adjacent and morphologically healthy tissues, MRE11 was analysed in NAT and NCT. The mean MRE11 score was 4.2 for both adjacent and distal normal tissue, which was significantly different from the score for both the TC and TP ($P<0.001$).

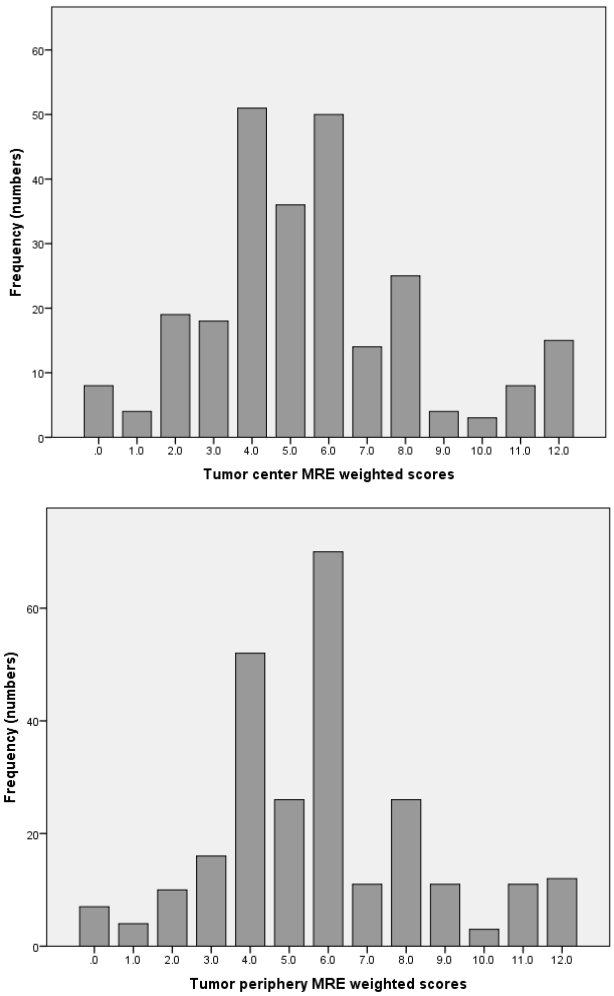


Figure 4-1The distribution of weighted scores for MRE11 in the tumour centre and periphery

4.4.3 Associations between MRE11 scores and clinicohistopathological characteristics

In contrast to what we observed for ATM, in this dataset we did not find any evidence for there being significant associations between the MRE11 score in the TP or TC and clinicohistopathological characteristics (Table 4.2).

Table 4-2 Fisher's test of association between MRE11 score in the tumour centre and tumour periphery (high vs. low) and clinicohistopathological variables

		Tumour centre			Tumour periphery		
		Low (%)	High (%)	p-value	Low (%)	High (%)	p-value
Sex	Male	65.40	68.10	0.69	65.20	66.70	0.90
	Female	34.60	31.90		34.80	33.30	
Age	<71	48.50	42.90	0.38	42.60	48.60	0.38
	71+	51.50	57.10		57.40	51.40	
Tumour stage	T1–2	33.10	33.60	1.00	30.40	34.80	0.50
	T3–4	66.90	66.40		69.60	65.20	
Node stage	Negative	49.60	55.50	0.38	46.40	59.30	0.06
	Positive	50.40	44.50		53.60	40.70	
Metastasis stage	M0	94.30	90.60	0.32	94.10	91.60	0.61
	M1	5.70	9.40		5.90	8.40	
Grade	1–2	90.40	94.10	0.35	92.20	92.40	1.00
	3	9.60	5.90		7.80	7.60	
Vascular invasion	No	74.40	76.50	0.77	72.80	78.20	0.38
	Yes	25.60	23.50		27.20	21.80	
Perineural invasion	No	81.20	87.40	0.23	81.60	85.20	0.50
	Yes	18.80	12.60		18.40	14.80	
Adjuvant therapy	No	67.60	72.00	0.55	67.00	71.10	0.55
	Yes	32.40	28.00		33.00	28.90	
Neoadjuvant therapy	No	76.40	81.50	0.42	77.10	79.90	0.64
	Yes	23.60	18.50		22.90	20.10	

4.4.4 Survival analysis

In order to determine whether MRE11 expression in rectal tumours was associated with patient outcomes, we next determined survival intervals according to MRE11 immunohistochemical score (product of proportion and intensity). By Kaplan–Meier analysis, patients found to exhibit a low level of MRE11 expression in the TC have a slightly higher OS than those with high MRE11 expression. This difference however is not significant (6.90 vs. 6.10 years, P = 0.42; Figure 4.2).

Similarly, the OS interval was found to be slightly shorter in patients with a low MRE11 score in the TP, although this also did not reach statistical significance (6.59 years in the low MRE11 group, compared to 6.68 years in the high MRE11 group, P = 0.77; Figure 4.2).

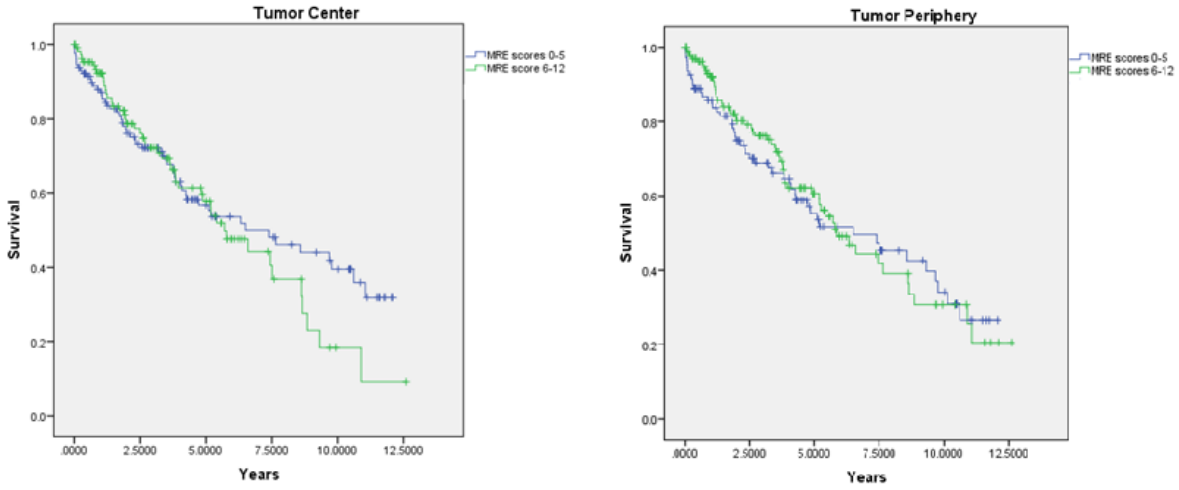


Figure 4-2 Kaplan–Meier survival curves for overall survival according to MRE11 weighted scores for the tumour periphery and tumour centre

4.4.5 Uni- and multi-variate analysis according to MRE11 expression in the TC

As we did in the case of ATM, we next sought to define additional features of disease in the cohort using uni- and multi-variate analyses able to model the effects of different confounding variables. Firstly, we did not find any significant differences in survival according to gender, age, tumour stage, neoadjuvant status, or adjuvant status. However, we did find that patients whose tumours showed evidence of vascular invasion to have a lower OS than patients whose tumours did not. These associations were maintained in both the high (hazard ratio [HR] = 0.42, $P = 0.01$) and low (HR = 0.55, $P = 0.04$) MRE11-expressing sub-groups.

Patients classed as having high levels of MRE11 expression, who were also positive for metastatic disease, had a lower OS than those who were negative for metastasis (HR=0.09, $P< 0.01$). Patients with high MRE11 expression who had negative lymph node status also had a better OS than patients who had positive lymph node status (HR = 0.48, $P = 0.02$). Furthermore, patients without perineural invasion were found to have a better OS than those with invasion, in both subgroups (low MRE11: HR = 0.36, $P< 0.01$; high MRE11: HR = 0.34, $P< 0.01$). Finally, patients with grade 1–2 disease showed better survival than those with grade 3–4 disease, in both low MRE11 (HR = 0.15, $P> 0.01$) and high MRE11 (HR = 0.02, $P < 0.01$) subgroups, as might be expected.

Table 4.4 depicts the multivariate analysis for MRE11 status in the TC. In these results, the absence of perineural invasion showed a protective effect in patients with low and high MRE11 expression (low MRE11: HR = 0.39, $P < 0.01$; high MRE11: HR = 0.37, $P = 0.05$). Furthermore, metastasis stage was a significant factor in patients with a high MRE11 (M0 vs. M1, HR = 0.09, $P< 0.01$). Finally, in patients with low MRE11 score, tumours graded 1–2 showed a protective effect versus those graded 3 (HR = 0.19, $P = 0.01$).

4.4.6 Uni- and multi-variate analysis according to MRE11 expression in the TP

Following on from our analyses of MRE11 staining intensity in the TC, we next extended our studies to investigate the TP. Table 4.3 shows the results of the univariate analysis for the clinicohistopathological variables associated with MRE11 staining intensity in both the MRE1 *low* and MRE11 *high* subgroups, with regards to the TP.

Table 4-3 Univariate analyses of survival outcomes according to MRE11 score in tumour centre and tumour periphery

	Tumour centre								Tumour periphery							
	MRE11 low				MRE11 high				MRE11 low				MRE11 high			
	HR	95% CI		p	HR	95% CI		p	HR	95% CI		p	HR	95% CI		p
Adjuvant therapy	1.94	0.96	3.92	0.06	2.26	0.93	5.49	0.07	3.28	1.38	7.81	0.01	1.51	0.77	2.99	0.24
(no vs. yes)																
Neoadjuvant therapy	3.81	0.38	37.67	0.25	1.06	0.52	2.15	0.88	1.81	0.76	4.28	0.18	0.73	0.41	1.33	0.31
(no. vs. yes)																
Vascular invasion	0.55	0.31	0.96	0.04	0.42	0.22	0.80	0.01	0.57	0.31	1.04	0.07	0.41	0.23	0.74	<0.01
(no vs. yes)																
Gender	1.15	0.66	2.01	0.61	0.88	0.47	1.65	0.69	1.13	0.63	2.01	0.68	0.99	0.57	1.74	0.98
(male vs. female)																
Age group	0.84	0.49	1.44	0.52	0.63	0.33	1.19	0.15	0.78	0.43	1.40	0.40	0.65	0.37	1.14	0.13
(≥70 years vs. <70 years)																
Grade	0.15	0.04	0.48	0.00	0.02	0.00	0.23	<0.01	0.02	0.06	0.81	0.02	0.01	0.00	0.09	<0.01
(1–2 vs. 3)																
Tumour stage	0.63	0.35	1.14	0.13	0.53	0.26	1.08	0.08	0.58	0.30	1.10	0.10	0.73	0.40	1.33	0.31
(T1–2 vs. T3–4)																
Metastasis stage	0.48	0.15	1.56	0.22	0.09	0.04	0.21	<0.01	0.29	0.10	0.83	0.02	0.15	0.07	0.33	<0.01
(M0 vs. M1)																
Node stage	0.87	0.50	1.51	0.63	0.48	0.27	0.87	0.02	0.69	0.39	1.24	0.21	0.68	0.40	1.16	0.16
(N0 vs. N1–2)																
Perineural invasion	0.36	0.20	0.65	0.00	0.34	0.16	0.75	0.01	0.53	0.27	1.03	0.06	0.35	0.19	0.65	<0.01
(negative vs. positive)																

Patients with low a MRE11 score that did not undergo adjuvant treatment were found to have a shorter OS than those who received radiation treatment (HR = 3.28, $P = 0.01$), supportive of the positive effects of this treatment. Furthermore, patients in the high MRE11 subgroup, and with evidence of vascular invasion, also had a lower OS than those without invasion (HR = 0.41, $P< 0.01$), again in-keeping with the aggressive disease phenotypes.

As might be expected, the OS was lower in patients with metastatic disease, in both the low (HR = 0.29, $P = 0.02$) and high MRE11 (HR = 0.15, $P< 0.01$) subgroups. Patients found to have a high MRE11 and were also positive for perineural invasion also had a lower OS than those who were negative for invasion (HR = 0.35, $P< 0.01$). Patients with grade 1–2 disease showed better survival than those with grade 3–4 for both the low MRE11 (HR = 0.02, $P = 0.02$) and high MRE11 (HR = 0.01, $P < 0.01$) subgroups.

Subsequently, we carried out multivariate analyses considering MRE11 status in the TP, the results of which are presented in Table 4.4. Being negative for perineural invasion (HR = 0.24, $P = 0.01$) and metastasis also had a protective effect in the high MRE11 subgroup (HR = 0.19,

$P < 0.01$), and patients in both high and low subgroups had a lower OS if they did not undergo adjuvant treatment (low MRE11, HR=3.05, p=0.01; high MRE11, HR=3.02, p=0.02), further emphasising the importance of radiotherapy to cancer treatment.

Table 4-4 Multivariate analysis of tumour centre and tumour periphery MRE11 scores for survival outcomes

	MRE11 Low				MRE11 High			
	HR	95% CI		p-value	HR	95% CI		p-value
Tumour periphery								
Adjuvant	3.05	1.27	7.34	0.01	3.02	1.19	7.62	0.02
(no vs. yes)								
Perineural invasion	0.64	0.28	1.49	0.30	0.24	0.09	0.67	0.01
(absence vs. presence)								
Metastasis stage	0.31	0.04	2.56	0.28	0.19	0.07	0.49	0.00
(M0 vs. M1)								
Tumour centre								
Perineural invasion	0.39	0.20	0.76	0.01	0.37	0.14	1.00	0.05
(absence vs. presence)								
Metastasis stage	1.05	0.29	3.77	0.94	0.09	0.04	0.23	0.00
(M0 vs. M1)								
Grade	0.19	0.05	0.70	0.01	0.34	0.04	3.34	0.36
(1-2 vs. 3)								

4.4.7 Lymph node-positive and adenoma-positive subgroups

The study population used included 115 patients who showed evidence for lymph node involvement. Amongst these participants, the mean MRE11 score was found to be 5.00 (SD = 3.00). In this subgroup of patients, MRE11 score (high [6–12] versus low [0–5]) was not found to be associated with OS in either the univariate or multivariate analysis (HR = 1.21, $P = 0.48$). The immunostaining method used for adenomas was the standard IHC procedure. However as so few tissue cores of these samples were available we did not include them in our data analysis as any interpretation will not be meaningful with such a small number.

4.5 Results – ATM and MRE11 combinatorial marker panel

Given the associations so far described between ATM, MRE11 and various features of rectal

cancer, we next sought to determine whether ATM and MRE11 could be used in combination to predict rectal cancer outcomes better than either marker alone.

The results arising from this approach have been recently published. Citation: Ho, V., Chung, L., Revoltar, M., Lim, S.H., Tut, T.G., Abubakar, A., Henderson, C.J., Chua, W., Ng, W., Lee, M., et al. (2016). MRE11 and ATM expression levels predict rectal cancer survival and their association with radiotherapy response. PLoS One. *11,* e0167675.

4.5.1 Patient Populations

The patient samples used for the study of ATM and MRE11 in combination is the same as described for MRE11, and is shown in Table 4.1.

4.5.2 Establishment of the combined two-protein biomarker panel

Firstly, ATM and MRE11 expression at the protein level in the TC were analysed in a forward and reverse binary logistic regression analysis, using a dataset of immune-histochemical scoring derived from 257 tumour samples and 255 normal tissue (NCT) samples. The final biomarker model gave an average receiver operator characteristic area under the curve (ROC-AUC) value of 0.849 for both proteins combined.

Similarly, ATM and MRE11 protein expression levels in the TP (tumour, *n* = 258; normal, *n* = 255) were also tested (see Table 4.5), with the model giving an average ROC-AUC value of 0.837. The sensitivity and specificity of the two-protein combined panel were 80.9% and 70.3% for the TC, and 61.6% and 48.8% for the TP, respectively.

Table 4-5 Performance of the two protein panel classification models

Model	**Tumour**	**Normal**	**Sensitivity (%)**	**Specificity (%)**	**Overall (%)**	**ROC-AUC**
Combined TC*	257	255	80.9	70.3	75.6	0.849
Combined TP†	258	255	61.6	48.8	55.2	0.837
*Tumour centre; †Tumour periphery						

4.5.3 Relationship between ATM and MRE11 protein expression and clinicohistopathological features

Given the promising results, we subsequently examined the association between ATM and MRE11 expression levels and clinicohistopathological variables independently.

In order to do this, ATM protein expression was first analysed in 259 central and 260 peripheral tumour cores. Samples from the TC were found to display high levels of expression in 84% of cases, and low expression was detected in 16% of samples. The samples from the TP showed high expression in 79% of samples, and low expression in 21%. When both sites were compared, a higher level of ATM expression was found in the TC (mean (M) = 6.84) when compared to the TP (M = 5.28). This difference was found to be statistically significant ($P< 0.001$).

Within the cohort, ATM expression was found in 199/228 (87%) of adjacent normal mucosa, and 204/250 (82%) of distal normal mucosa samples. Furthermore, ATM expression was found to be higher in the normal mucosa taken from near the tumour (M = 6.43) compared to samples taken from sites more distal to the tumour (M = 4.25). Again, this observation was found to be statistically significant ($P< 0.001$). When comparing tumours and healthy tissue, ATM expression in both the TC and TP was found to significantly higher than in distal normal mucosa ($P< 0.001$ and $P< 0.005$, respectively), although expression was not significantly different to that in the adjacent normal mucosa.

Low ATM expression in the TP was found to be associated with older patient age at the time of surgery ($P = 0.013$) and higher histologic grade ($P = 0.044$) (refer to Table 3.2). We did not find there to be a correlation with sex, TNM category, vascular invasion, or perineural invasion. In addition, expression of ATM in the TC was not associated with any clinicohistopathological variables.

MRE11 expression in the TC was high in 45% of cases, and low in 52%. MRE11 expression in the TP was high in 56%, and low in 44% of samples, and the mean weighted scores were significantly different between TC and TP (5.5 versus 5.8, $P< 0.05$, by paired t-test).

The mean MRE11 score was 4.2 for both adjacent and distal normal tissue, a score that was significantly different to that for both the TC and TP ($P< 0.001$). We did not find any significant associations between MRE11 score in the TC or TP, and clinicohistopathological characteristics (Table 4.6).

Table 4-6 Association between MRE11 expression and clinicohistopathological data

	Tumour centre			Tumour periphery		
	Low	High	P value	Low	High	P value
	n (%)	n (%)		n (%)	n (%)	
Sex						
Male	65.4	68.1	0.69	65.2	66.7	0.9
Female	34.6	31.9		34.8	33.3	
Age						
<71	48.5	42.9	0.38	42.6	48.6	0.38
71+	51.5	57.1		57.4	51.4	
Tumour stage						
T1–2	33.1	33.6	1	30.4	34.8	0.5
T3–4	66.9	66.4		69.6	65.2	
Node stage						
Negative	49.6	55.5	0.38	46.4	59.3	0.06
Positive	50.4	44.5		53.6	40.7	
Metastasis stage						
M0	94.3	90.6	0.32	94.1	91.6	0.61
M1	5.7	9.4		5.9	8.4	
Grade						
1–2	90.4	94.1	0.35	92.2	92.4	1
3	9.6	5.9		7.8	7.6	
Vascular invasion						
Absent	74.4	76.5	0.77	72.8	78.2	0.38
Present	25.6	23.5		27.2	21.8	
Perineural invasion						
Absent	81.2	87.4	0.23	81.6	85.2	0.5
Present	18.8	12.6		18.4	14.8	

4.5.4 Association between combined expression levels with pathological variables and prognosis

Next, we examined the possible association of clinicohistopathological characteristics and survival outcomes with the ATM/MRE11 two-protein combined panel. Figure 4.3A shows representative immunohistochemical staining of high and low/absent ATM and MRE11 expression levels in rectal cancer tissues. Results from the combined marker analysis differed from our initial single biomarker studies of ATM and MRE11 alone, in that a high combined expression of ATM and MRE11 was significantly associated with a number of clinicohistopathological variables. These included neoadjuvant status (*P*=0.001 for TC), TRG (*P*=0.03 for TC, *P*=0.011 for TP), age (*P*=0.02 for TP), and nodal stage (*P*=0.042 for TP) (Table 4.7), demonstrating the informative power of this approach.

Table 4-7 Associations between the combined expression of ATM and MRE11 in the TC and TP and clinicohistopathological data

		Combined TC			Combined TP		
		Low (%)	**High (%)**	***P* value**	**Low (%)**	**High (%)**	*P* value
Sex	Male	62.5	67.3	0.562	65.3	66.8	0.381
	Female	37.5	32.7		34.7	33.2	
Age	<70	43.8	46.3	0.747	42.9	46.0	0.020
	>70	56.2	53.7		57.1	54.0	
Tumour stage	T1–2	50.0	29.5	0.496	38.3	31.9	0.877
	T3–4	50.0	70.5		61.7	68.1	
Node stage	Negative	60.0	52.4	0.840	44.7	55.8	0.042
	Positive	40.0	47.6		55.3	44.2	
Metastasis stage	M0	97.8	91.6	0.282	100	91.1	0.252
	M1	2.2	8.4		0	8.9	
Grade	1–2	93.8	92.1	0.508	89.8	92.9	0.131
	3	6.2	7.9		10.2	7.1	
Vascular invasion	Absent	82.6	74.6	0.231	77.1	75.6	0.380
	Present	17.4	25.4		22.9	24.4	
Perineural invasion	Absent	84.8	83.6	0.928	85.4	83.3	0.500
	Present	15.2	16.4		14.6	16.7	
Adjuvant	No	70.0	69.5	0.483	65.0	70.9	0.55

therapy	Yes	30.0	30.5		35.0	29.1	
Neoadjuvant therapy	No	65.2	81.1	**0.001**	71.7	79.9	0.64
	Yes	34.8	18.9		28.3	20.1	
Tumour regression grade	0–2	31.6	8.0	**0.030**	33.3	6.0	0.011
	3	68.4	92.0		66.7	94.0	

By Kaplan Meier survival analysis, a high score for the two-protein combined expression panel in the TC was significantly associated with worse OS (*P*=0.003) and DFS (*P*=0.035) (Figure. 4.3 B and C). However, no significant survival difference was seen amongst high or low groups for expression in the TP (OS, *P*=0.208 and DFS, *P*=0.748; Figure.4.3 D and E, respectively).

By univariate Cox regression analysis, a high two-protein combined status in the TC was significantly associated with reduced DFS (combined TC high versus low: HR=1.944, 95%CI 0.037–3.645, *P*=0.038) (Table 4.8). Finally, by multivariate Cox analysis (adjusted for combined expression of ATM and MRE11, and perineural invasion), the two-protein combination panel (HR = 2.178, 95%CI 1.115–4.256, *P*=0.023) as well as perineural invasion (HR=2.183, 95%CI 1.222– 3.899, *P*=0.008) remained significantly associated with DFS (Table 4.8).

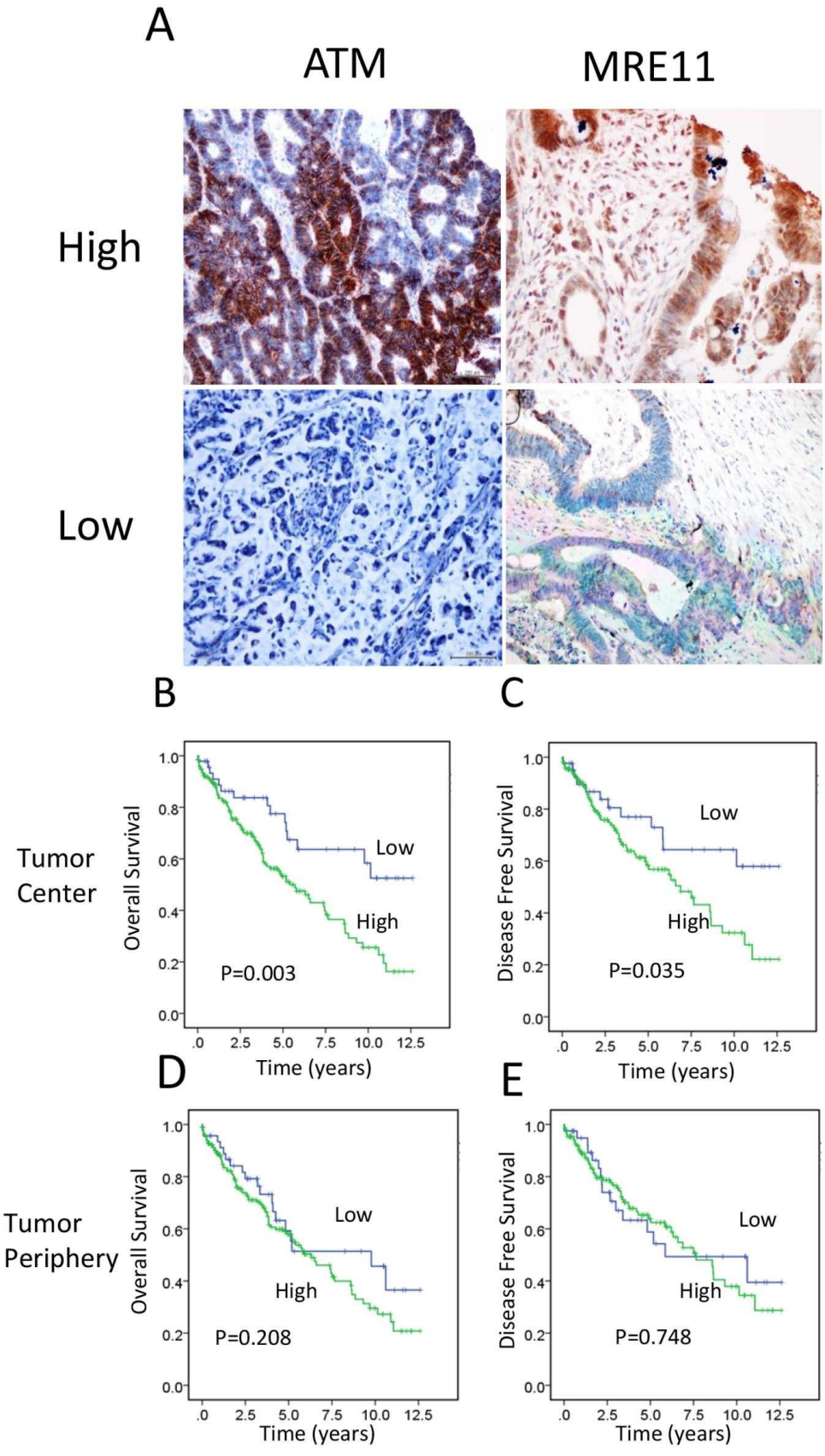

A
ATM
MRE11
High
Low
B
C
Tumor Center
Overall Survival
Disease Free Survival
Low
High
P=0.003
P=0.035
Time (years)
D
E
Tumor Periphery
P=0.208
P=0.748

Figure 4-3 Association between combined protein expression levels of ATM and MRE11 in rectal cancer tissues and survival

(A) Representative immunohistochemical staining of ATM and MRE11 in rectal cancer tissues. Scale bar=200μm. Staining for each protein was scored as high or low, as described in Chapter 2.
(B) Overall survival according to combined protein expression levels of ATM and MRE11 was determined by Kaplan−Meier survival analyses and compared using the log-rank test. Patients with high combined protein expression levels of ATM and MRE11 in the TC (greenline) showed significantly worse OS than those with low expression (blueline; P=0.003).
(C) Similarly, patients with high expression levels in the TC (greenline) exhibited worse disease-free survival than those with low expression (blueline; P=0.035).
(D) (E)When measured in the TP, no significant survival difference was seen between the high- and low-expression groups for overall survival (P=0.208) or disease-free survival (P=0.748).

Table 4-8 Cox regression analyses of combined TC expression level with clinicohistopathological variables and DFS. From: Ho et al (2016)

		Univariate			Multivariate		
	n (%)	HR	95% CI	P Value	HR	95%	P Value
Combined TC							
High	81.3	1.944	1.037–3.645	**0.038**	2.178	1.115–4.256	**0.023**
Low	18.7						
Tumor stage							
T1–2	33.5	1.504	0.899–2.517	0.12			
T3–4	66.5						
Node stage							
Negative	53.5	1.132	0.703–1.825	0.609			
Positive	46.5						
Grade							
1–2	92.4	1.617	0.699–3.739	0.261			
3	7.6						
Vascular invasion							
Absent	76.2	1.164	0.637–2.129	0.662			
Present	23.8						
Perineural invasion							
Absent	83.8	2.334	1.310–4.157	**0.004**	2.183	1.222–3.899	**0.008**
Present	16.2						
Adjuvant therapy							
No	91.0	0.605	0.343–1.067	0.082			
Yes	9.0						
Neoadjuvant therapy							
No	77.7	1.088	0.630–1.878	0.762			
Yes	22.3						
LN-negative† combined TC	53.5				1.343	0.591–3.052	0.481
LN-positive† combined TC	46.5				3.474	1.054–11.451	**0.041**

Abbreviations: HR, hazard ratio; CI, confidence intervals; TC, tumor center; LN, lymph node;
†denotes interaction

doi:10.1371/journal.pone.0167675.t003

4.5.5 Prognostic implications of two-protein combined panel in LN-positive and neoadjuvant radiotherapy subgroups

As noted above, the DFS interval of rectal cancer patients overexpressing of the two-protein combined panel was significantly worse than that of patients showing lower expression. When patients were grouped according to LN involvement, high expression of the two-protein combined panel was again associated with decreased DFS in patients with LN-positive tumours (P = 0.029; Figure 4.4B), but not in those with LN-negative tumours (P = 0.480; Figure 4.4A). By multivariate Cox analysis, expression of the two-protein combined panel in the TC in the LN positive subgroup significantly correlated with DFS (HR = 3.474, 95% CI 1.054–11.451, P = 0.041) (Table 4.8).

We next sought to determine any associations with neo-adjuvant radiotherapy in our cohort. The OS estimates in the subgroup that received neo-adjuvant radiotherapy are shown in Figure 4.4 C, demonstrating that a higher combined expression level is significantly associated with worse OS (P=0.024).Similarly, the DFS of patients that underwent radiotherapy and overexpressed the two-protein combined panel was also

significantly worse than that of patients with lower expression levels (Figure 4.4D, *P*=0.028).These results demonstrate that the MRE11/ATM two-protein panel has specific potential as a predictive marker of the tumour response to radiotherapy.

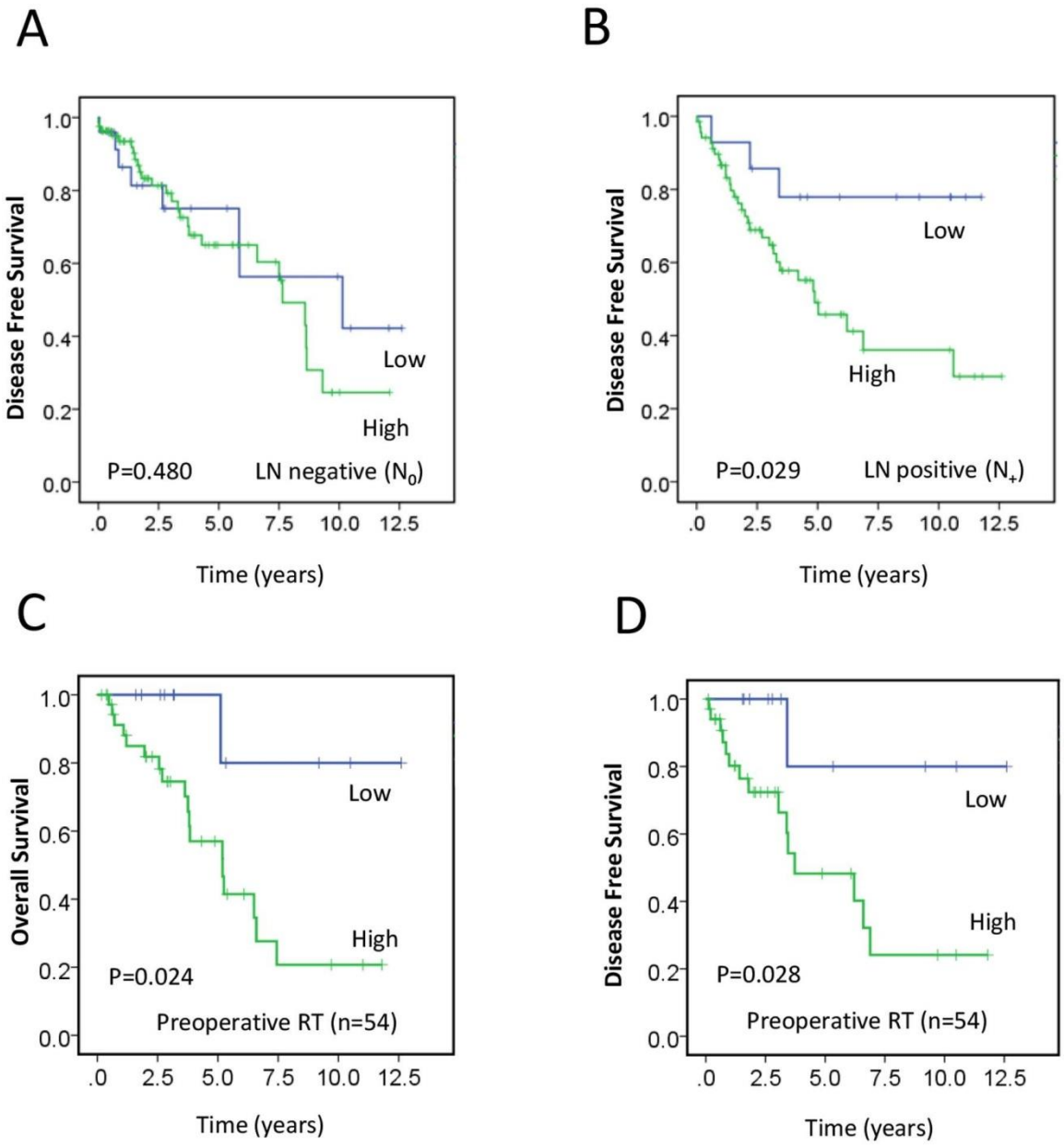


Figure 4-4Kaplan-Meier survival analysis of ATM/MRE11

(A) and (B) respectively show survival curves of high (greenline) and low (blueline) ATM/MRE11 two-protein expression groups in lymph node (LN) negative and LN positive rectal cancers. These show the effect of LN status on the association between expression levels and disease-free survival (DFS).
(C) and (D) respectively show overall survival (OS) and DFS survival curves of 54 patients who received preoperative radiotherapy in terms of high (greenline) and low (blueline) two protein expression groups. Worse OS (*P*=0.024) and DFS (*P*=0.028) were seen with high expression.

4.5.6 ATM and MRE11 expression in relation to TRG

Finally, to investigate the relationship between TRG and tissue levels of ATM and MRE11 (and the two- protein combined panel), univariate analysis was carried out using the Mann−Whitney U test. A significant association was found between increasing TRG, and the TC expression levels of MRE11 (P=0.015, Figure 4.5B) and that of the combined two-protein panel (P=0.011, Figure 4.5C).

The discriminatory power of each protein biomarker, and that of the combinatorial panel, were further characterised using an ROC-AUC analysis of good response (TRG 0-2) versus poor response (TRG3) groups. The average ROC-AUC was found to be 0.745 for the combined panel, compared with 0.618 for ATM, and 0.711 for MRE11 proteins (Figure 4.5 D). These results suggest that the combinatorial biomarker provides excellent discrimination between good and poor tumour responses after radiotherapy.

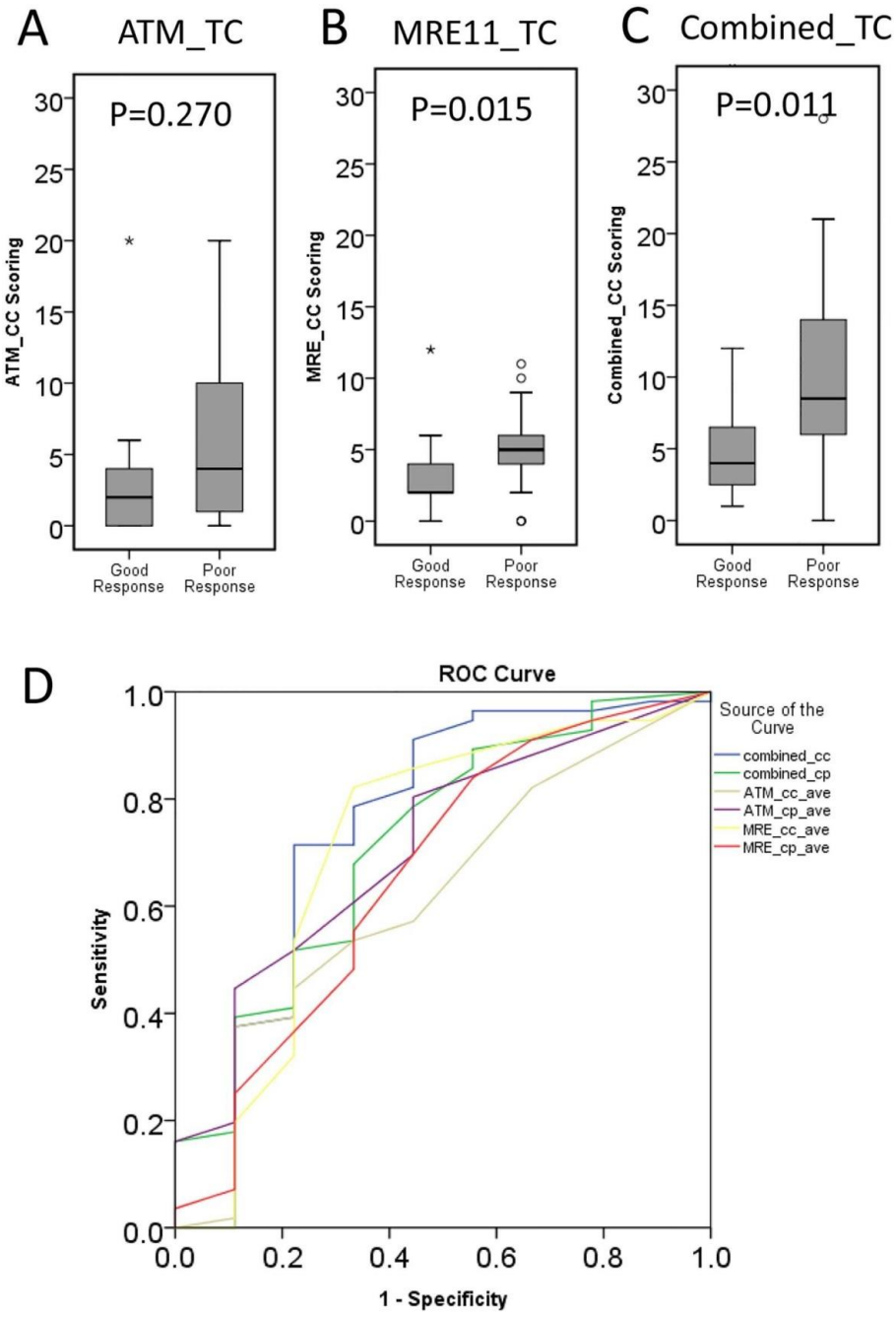


Figure 4-5 ROC-AUC analysis of ATM, MRE11, and combined protein panel expressions with tumour regression grade (TRG)

(A−C) Box plots show levels of ATM, MRE11, and their combined expression in the TC categorized by TRG as 0–2 (good response) or 3 (poor response). The association between protein expression with TRG was examined by Mann−Whitney U test (ATM, P = 0.27; MRE11, P = 0.015; and combined proteins, P = 0.011).
(D) Receiver operating characteristic (ROC) curve analysis comparing the performance of ATM and MRE11 alone with the combined 2-protein panel. It should be noted that cc is actually referring to TC and cp is referring to TP

4.6 Concluding remarks

In contrast to what we discovered with regards to ATM expression in the previous chapter, MRE11 expression alone was not found to affect the survival outcomes of patients with rectal cancer.

The level of MRE11 protein in the TC or TP, however, was found to influence patient outcomes when participants were positive for perineural invasion, metastasis, or high-grade disease. Accordingly, knowledge of MRE11 gene/protein expression status may affect clinical decision-making, and improve patient welfare.

The ATM/MRE11 combinatorial panel developed as part of this study merits further investigation as a clinical predictive marker of the tumour response to neoadjuvant radiotherapy.We accept the limitation of a small sample size of pretreatment biopsies in our study rendering any interpretation of pretreatment data meaningless. Future studies however would evaluate the results of pretreatment biopsies in a predictive model to determine cases responding optimally to radiotherapy.

In conclusion, our data suggest that optimal rectal cancer management may benefit from tailored treatment based on biomarker expression.

4.7 References

Boisvert, F.M., Dery, U., Masson, J.Y., and Richard, S. (2005). Arginine methylation of MRE11 by PRMT1 is required for DNA damage checkpoint control. Genes Dev. *19*, 671–676.

Bressan, D.A., Baxter, B.K., and Petrini, J.H. (1999). The Mre11-Rad50-Xrs2 protein complex facilitates homologous recombination-based double-strand break repair in Saccharomyces cerevisiae. Mol. Cell. Biol. *19*, 7681–7687.

Chapkin, R.S., Davidson, L.A., Ly, L., Weeks, B.R., Lupton, J.R., and McMurray, D.N. (2007). Immunomodulatory effects of (n-3) fatty acids: putative link to inflammation and colon cancer. J Nutr *137*, 200S–204S.

Choudhury, A., Nelson, L.D., Teo, M.T.W., Chilka, S., Bhattarai, S., Johnston, C.F., Elliott, F., Lowery, J., Taylor, C.F., Churchman, M., et al. (2010). MRE11 expression is predictive of cause-specific survival following radical radiotherapy for muscle-invasive bladder cancer. Cancer Res. *70*, 7017–7026.

Chubb, D., Broderick, P., Dobbins, S.E., Frampton, M., Kinnersley, B., Penegar, S., Price, A., Ma, Y.P., Sherborne, A.L., Palles, C., et al. (2016). Rare disruptive mutations and their contribution to the heritable risk of colorectal cancer. Nat. Commun. *7*, 11883.

Damiola, F., Pertesi, M., Oliver, J., Le Calvez-Kelm, F., Voegele, C., Young, E.L., Robinot, N., Forey, N., Durand, G., Vallée, M.P., et al. (2014). Rare key functional domain missense substitutions in MRE11A, RAD50, and NBNcontribute to breast cancer susceptibility: results from a Breast Cancer Family Registry case-control mutation-screening study. Breast Cancer Res. *16*, R58.

Deng, Y., Guo, X., Ferguson, D.O., and Chang, S. (2009). Multiple roles for MRE11 at uncapped telomeres. Nature *460*, 914–918.

De Jager, M., Trujillo, K.M., Sung, P., Hopfner, K.P., Carney, J.P., Tainer, J.A., Connelly, J.C., Leach, D.R.F., Kanaar, R., and Wyman, C. (2004). Differential arrangements of conserved building blocks among homologs of the Rad50/Mre11 DNA repair protein complex. J. Mol. Biol. *339*, 937–949.

Elefsinioti, A., Bellaire, T., Wang, A., Quast, K., Seidel, H., Braxenthaler, M., Goeller, G., Christianson, A., Henderson, D., and Reischl, J. (2016). Key factors for successful data integration in biomarker research. Nat. Rev. Drug Discov. *15*, 369–370.

Giannini, G., Ristori, E., Cerignoli, F., Rinaldi, C., Zani, M., Viel, A., Ottini, L., Crescenzi, M., Martinotti, S., Bignami, M., et al. (2002). Human MRE11 is inactivated in mismatch repair-deficient cancers. EMBO Rep. *3*, 248–254.

Heng, T.S.P., and Painter, M.W. (2008). The Immunological Genome Project: networks of gene expression in immune cells. Nat. Immunol. *9*, 1091–1094.

Ho, V., Chung, L., Revoltar, M., Lim, S.H., Tut, T.G., Abubakar, A., Henderson, C.J., Chua, W., Ng, W., Lee, M., et al. (2016). MRE11 and ATM expression levels predict rectal cancer survival and their association with radiotherapy response. PLoS One. *11*, e0167675.

Hoeijmakers, J.H. (2001). Genome maintenance mechanisms for preventing cancer. Nature *411*, 366–374.

Hopfner, K.P., Karcher, A., Craig, L., Woo, T.T., Carney, J.P., and Tainer, J.A. (2001). Structural biochemistry and interaction architecture of the DNA double-strand break repair Mre11 nuclease and Rad50-ATPase. Cell *105*, 473–485.

Ihara, K., Yamaguchi, S., Ueno, N., Tani, Y., Shida, Y., Ogata, H., Domeki, Y., Okamoto, K., Nakajima, M., Sasaki, K., et al. (2015). Expression of DNA double-strand break repair proteins predicts the response and prognosis of colorectal cancer patients undergoing oxaliplatin-based chemotherapy. Oncol. Rep. *35*, 26676960.

Inagaki, A., Roset, R., and Petrini, J.H.J. (2016). Functions of the MRE11 complex in the development and maintenance of oocytes. Chromosoma *125*, 151–162.

Kim, E.R., and Chang, D.K. (2014). Colorectal cancer in inflammatory bowel disease: The risk, pathogenesis, prevention and diagnosis. World J. Gastroenterol. *20*, 9872–9881.
Kim, H., Cho, D.Y., Choi, D.H., Oh, M., Shin, I., Park, W., Huh, S.J., Nam, S.J., Lee, J.E., and Kim, S.W. (2017). Frequency of pathogenic germline mutation in CHEK2, PALB2, MRE11, and RAD50 in patients at high risk for hereditary breast cancer. Breast Cancer Res. Treat. *161*, 95–102.
Kondo, T., Kobayashi, J., Saitoh, T., Maruyama, K., Ishii, K.J., Barber, G.N., Komatsu, K., Akira, S., and Kawai, T. (2013). DNA damage sensor MRE11 recognizes cytosolic double-stranded DNA and induces type I interferon by regulating STING trafficking. Proc. Natl. Acad. Sci. *110*, 2969–2974.
Koppensteiner, R., Samartzis, E.P., Noske, A., Von Teichman, A., Dedes, I., Gwerder, M., Imesch, P., Ikenberg, K., Moch, H., Fink, D., et al. (2014). Effect of MRE11 loss on PARP-inhibitor sensitivity in endometrial cancer In Vitro. PLoS One *9*. e100041.
Lamarche, B.J., Orazio, N.I., and Weitzman, M.D. (2010). The MRN complex in double-strand break repair and telomere maintenance. FEBS Lett. *584*, 3682–3695.
Lanzano, L., Hernandez, I., Castello, M., Gratton, E., Diaspro, A., and Vicidomini, G. (2015). Encoding and decoding spatio-temporal information for super-resolution microscopy. Nat Comms. *6*, 6701.
Miquel, C., Jacob, S., Grandjouan, S., Aimé, A., Viguier, J., Sabourin, J.-C., Sarasin, A., Duval, A., and Praz, F. (2007). Frequent alteration of DNA damage signalling and repair pathways in human colorectal cancers with microsatellite instability. Oncogene *26*, 5919–5926.
Nascimbeni, R., Di Fabio, F., Di Betta, E., Mariani, P., Fisogni, S., and Villanacci, V. (2005). Morphology of colorectal lymphoid aggregates in cancer, diverticular and inflammatory bowel diseases. Mod. Pathol. *18*, 681–685.
Pavelitz, T., Renfro, L., Foster, N.R., Caracol, A., Welsch, P., Lao, V.V., Grady, W.B., Niedzwiecki, D., Saltz, L.B., Bertagnolli, M.M., et al. (2014). MRE11-deficiency associated with improved long-term disease free survival and overall survival in a subset of stage III colon cancer patients in randomized CALGB 89803 trial. PLoS One *9*. e108483.
Peltomaki, P. (2001). Deficient DNA mismatch repair: a common etiologic factor for colon cancer. Hum. Mol. Genet. *10*, 735–740.
Quaiser, A., Constantinesco, F., White, M.F., Forterre, P., and Elie, C. (2008). The Mre11 protein interacts with both Rad50 and the HerA bipolar helicase and is recruited to DNA following gamma irradiation in the archaeon Sulfolobus acidocaldarius. BMC Mol. Biol. *9*, 25.
Rebbeck, T.R., Mitra, N., Domchek, S.M., Wan, F., Friebel, T.M., Tran, T. V., Singer, C.F., Tea, M.K.M., Blum, J.L., Tung, N., et al. (2011). Modification of BRCA1-associated breast and ovarian cancer risk by BRCA1-interacting genes. Cancer Res. *71*, 5792–5805.
Schiller, C.B., Seifert, F.U., Linke-Winnebeck, C., and Hopfner, K.-P. (2014). Structural Studies of DNA End Detection and Resection in Homologous Recombination. Cold Spring Harb. Perspect. Biol. a017962.
Schwartz, R.A., Palacios, J.A., Cassell, G.D., Adam, S., Giacca, M., and Weitzman, M.D. (2007). The Mre11/Rad50/Nbs1 Complex Limits Adeno-Associated Virus Transduction and Replication. J. Virol. *81*, 12936–12945.
Sipos, F., and Muzes, G. (2011). Isolated lymphoid follicles in colon: Switch points between inflammation and colorectal cancer? World J. Gastroenterol. *17*, 1666–1673.
Stracker, T.H., and Petrini, J.H.J. (2011). The MRE11 complex: starting from the ends. Nat. Rev. Mol. Cell Biol. *12*, 90–103.
Su, A.I., Wiltshire, T., Batalov, S., Lapp, H., Ching, K.A., Block, D., Zhang, J., Soden, R., Hayakawa, M., Kreiman, G., et al. (2004). A gene atlas of the mouse and human protein-encoding transcriptomes. Proc. Natl. Acad. Sci. *101*, 6062–6067.
Tjalsma, H., Boleij, A., and Kato, I. (2012). Streptococcus bovis and colorectal cancer. In

Bacteria and Cancer, pp. 61–78.
Vilar, E., Bartnik, C.M., Stenzel, S.L., Raskin, L., Ahn, J., Moreno, V., Mukherjee, B., Iniesta, M.D., Morgan, M.A., Rennert, G., et al. (2011). MRE11 deficiency increases sensitivity to poly(ADP-ribose) polymerase inhibition in microsatellite unstable colorectal cancers. Cancer Res. *71*, 2632–2642.
West, S., Gromak, N., and Proudfoot, N.J. (2004). Human 5′ → 3′ exonuclease Xrn2 promotes transcription termination at co-transcriptional cleavage sites. Nature *432*, 522–525.
Williams, R.S., Moncalian, G., Williams, J.S., Yamada, Y., Limbo, O., Shin, D.S., Groocock, L.M., Cahill, D., Hitomi, C., Guenther, G., et al. (2008). Mre11 Dimers Coordinate DNA End Bridging and Nuclease Processing in Double-Strand-Break Repair. Cell *135*, 97–109.
Wu, C., MacLeod, I., and Su, A.I. (2013). BioGPS and MyGene.info: Organizing online, gene-centric information. Nucleic Acids Res. *41*.
Zdobnov, E.M., Tegenfeldt, F., Kuznetsov, D., Waterhouse, R.M., Simao, F.A., Ioannidis, P., Seppey, M., Loetscher, A., and Kriventseva, E. V. (2017). OrthoDB v9.1: Cataloging evolutionary and functional annotations for animal, fungal, plant, archaeal, bacterial and viral orthologs. Nucleic Acids Res. *45*, D744–D749.
Zhong, Z., Jiang, W., Cesare, A., Neumann, A., Wadhwa, R., and Reddel, R. (2007). Disruption of telomere maintenance by depletion of the MRE11/RAD50/NBS1 complex in cells that use alternative lengthening of telomeres. J Bio Chem *292*, 29314–29322.

CHAPTER 5
THE ASSOCIATION BETWEEN TUMOUR NBS1 EXPRESSION AND SURVIVAL OUTCOMES IN RECTAL CANCER

5.1 Introduction

This chapter concerns the DNA DSB repair protein, NBS1, another critical member of the MRN complex discussed previously (Kobayashi et al., 2004); emphasising specifically the role of chromosomal integrity mechanisms in rectal oncogenesis.

Taking a similar approach to what has been described in previous Chapters, NBS1 protein expression will be interrogated by immunohistology in tumour samples obtained from rectal cancer patients. By so doing – and remembering that results relating to RAD50 will be covered in Chapter Six – a more complete picture of how this complex/pathway contributes to rectal cancer pathogenesis will be clarified, as discussion of which will follow in Chapter Seven. Furthermore, the information here presented is relevant to the results of a combinatorial panel (composed of MRE11, NBS1 and RAD50) in the same clinical context, which is presented in Chapter Six.

In this chapter, different members of the same complex are studied in tandem, although the statistical benefits and increased biological information associated with the use of combinatorial biomarker panels has furthered methods to determine panels *de novo* (Milward et al., 2014), from hypothesis-free networks built from genome-wide datasets (Tanić and Beck, 2017).

5.1.1 NBS1: biochemistry and mechanisms of action

At present, NBS1 is also commonly referred to as *Nibrin*, and has also previously been known as cell cycle regulatory protein p95, as well as Nijmegen breakage syndrome protein 1 (from where it got its now-common abbreviated name). The varied nomenclature history is demonstrative of the long-held interest in this gene/protein, and its notable association with human clinical diseases (Zhou and Elledge, 2000). Indeed, defects in the *NBN* gene (located on human chromosome 8q21) encoding the NBS1 protein, lead to aberrant DNA repair and chromosomal instability (Wu, 2016), which has effects across the entire organism.

As with the other members of the MRN complex, NBS1 is highly conserved (Zdobnov et al., 2017). *NBN* can trace its origins beyond the beginning of eukaryotic life (Speir et al., 2016). Important studies have shown the role of this protein in organising and maintaining DNA integrity and function in the plant, *Arabidopsis thaliana*, where NBS1 was found to participate in DNA recombination during the very early stages of meiosis; which it achieves in tandem with ATM (Waterworth et al., 2007).

In-keeping with its best-known role across taxa, the majority of NBS1 at any one time is found in the cell's nucleus, notably in nuclear dots (or PML bodies) interspersed between chromatin folds (Naka et al., 2002). Thus, it is located in close proximity to potential lesions that arise during DNA replication, and in association with many of the co-factors involved in its function (Misteli and Soutoglou, 2009).

A primary component of nuclear dots has been shown to be the protein, sp100, which in response to interferons (such as those released during viral infection), drive transcription (Misteli and Soutoglou, 2009), and presumably additional mechanisms exist to regulate PML body utility. Sp100 has also been shown to recruit NBS1 to nuclear dots (Naka et al., 2002). Interestingly, the Human Herpes Virus protein, ICP0, has been found to disrupt nuclear dot formation (Hutchinson et al., 2002). As will be discuss below, ICP0 is also a key regulator of the MRN complex, and NBS1. Several common human variants in NBS1 also lead to nuclear foci formation disruption (Tauchi et al., 2001).

NBS1 is also associated with telomeres and is known to move into the cytoplasm when required, such as during heat shock (Seno and Dynlacht, 2004). Again, and in common with other members of the MRN complex, NBS1 has been shown to have crucial roles in maintaining telomere integrity, cell-cycle progression, and the regulation of meiosis, aside from its role in DSB repair (Zhang et al., 2006). In the context of DSBs, ATM first phosphorylates members of the MRN complex (Lee and Paull, 2007), including NBS1, as well as H2AX. These phospho-residues on H2AX are detected by the BRCT and FHA domains of NBS1 (Kobayashi et al., 2002). Activated NBS1 is also known to further regulate DNA damage signalling through effects on ATM, ATR and PRKDC, and is also involved in bringing MRE11 and RAD50 into the proximity of DSBs in a H2AX-dependent manner (Enriquez-Rios et al., 2017). Furthermore, NBS1 is also a potent cell cycle regulator, and has been shown to regulate both the G1 and G2 checkpoints (Komatsu, 2016).

In mammals, complete abrogation of the gene encoding NBS1 leads to embryonic lethality (Zhu et al., 2001), as shown here in mice, and in agreement with the hypothesis that in individuals with *NBN* mutations, some level of gene and protein function remains (i.e. the mutants are still compatible with maintaining cell viability up to a point). Genes displaying embryonic lethality illustrate the absolute requirement of said element for life; being irreplaceable in certain circumstances; which may be genetically or developmentally determined, or driven by the environment. Furthermore, disruption of different segments of the gene have been associated with female infertility due to oogenesis failure, again highlighting its important role in germ-cell development (Kang et al., 2002).

According to the respected IntAct database of the EBI EMBL, which is routinely updated (Orchard et al., 2014), and which quantifies protein-protein binary interactions at the sub-cellular level, the proteins most associated with changes in NBS1 expression are H2AX, TCOF, MDC1 and VRK1. We know about some of the mechanisms of H2AX already; TCOF may be involved in the maturation of rRNA (something that needs to be curtailed when halting cell growth, as occurs during infection (Zhang and Kuspa, 2009)); MDC1 recruits and anchors proteins to DSB lesions (Stucki and Jackson, 2006); and VRK1 has been shown to promote the stability of p53 (Vega et al., 2004); to name but a few functions of these associated co-factors, all of which are central to safeguarding cellular genomic DNA and the cell cycle.

These interactions and the functions of NBS1 are mediated by its structure. The primary isoform of the human protein is 754 amino acids in length, and has a mass of approximately

85 kDa(The UniProt Consortium, 2015). The N-terminal region of the protein is characterised, first by a forkhead-associated domain (FHA) and secondly by a BRCT domain. The FHA domain recognises phosphor-peptides (which widely regulate cellular physiology, and this domain has been reported to be essential for T cell development (Weng et al., 2015)), as does the BRCT domain, which is known to be important in proteins responsible for cell cycle control, and is widely known to participate in cancer development through studies of breast cancer, where it was named (Yu, 2003).

Part of the BRCT domain also confers the capacity for NBS1 to interact with SP100 (the major component of nuclear dots), whilst the middle section of the protein is involved in mediating interactions with the central metabolic regulators mTOR, MAPKAP1 and RICTOR. Indeed, residues 221-402 of NBS1 mediate the interaction of the protein with the mTOR/RICTOR/SIN1 complex (Wang et al., 2013), which drives AKT activity – a key regulator of cell growth and proliferation, as well as glucose utilisation and apoptosis.

At the C-terminal end of the protein lie the nuclear localisation signal sequence, and the EEXXXDDL motif, which mediates the interaction of NBS1 with ATM (Passananti and Fanciulli, 2007). NBS1 has been shown to activate ATM under various conditions (Difilippantonio et al., 2005), and provides a critical link between the phosphatase and the MRN complex. Antibody-binding epitopes have also been discovered at the C-terminal end, which may regulate the function of NBS1. Indeed, the increasing understanding of intracellular antibody responses has opened a number of interesting avenues, and antibody-coated pathogens, for example, in the intracellular space, may be detected by NBS1 (McEwan et al., 2013).

As mentioned earlier, and like the proteins ATM and MRE11, NBS1 has also been associated with viral infectivity in humans and primates. In an interesting study from 2016 study by Lou and colleagues, an intrinsically disordered region of NBS1 was found to represent a species-specific barrier to Herpes Simplex Virus 1 in primates; i.e. some primates carried a variant allowing for HSV infection (Lou et al., 2016). Further analysis of the variants that promoted infection revealed that primate NBS1 interacted with viral ICP0 via a region of structural disorder in NBS1. Thus, at least two outcomes are possible. The human protein senses the virus to halt infectivity (and some variants may be better at this than others), or the virus hijacks the protein via such a disordered region. Both are likely to occur, given the pathogenic arms race and the reported associations of other MRN complex members with anti-viral responses, both discussed previously.

The authors of this study extend their analyses to a further 1,237 mammalian proteins known to bind to viruses, and found an enrichment for similar disordered domains in these proteins, compared to non-interacting factors (Lou et al., 2016). It is tempting to speculate that what we here refer to as disorder, actually represents fine variability in the structural methods required to identify pathogens (as has been described for the T-cell receptor that can selectively responsd to peptide antigens of 8 amino acids) by such putative pattern recognition receptors.

Related studies have explored this area further, and it has been shown that the interaction between Herpes viruses and the MRN complex is profound and also includes the viral protein ICP8, which is known to also modulate the function of RAD50, MRE11, and PRKDC (Balasubramanian et al., 2010). Many authors have suggested that the MRN complex is required for Herpes virus infectivity, presumably by means distinct from the mechanisms in primates, discussed above. Further high-resolution structures of these components, and genome-wide CRISPR screens (Puschnik et al., 2017), are likely to provide additional information in this respect. Given the ubiquitous nature of viruses, bacteriophages and other micro-organisms over evolutionary time, it is likely that we are only beginning to understand the means by which our cells sense foreign material and relay that information to other cellular components; maybe sometimes we need to do as the virus tells us, and the protective effects of viral infections against a range of ailments are well established (Furman et al., 2015; Staras et al., 2006). Further functional analysis of the NBS1 protein showed that it has important transcription factor binding domains, allowing it to regulate gene expression (Ashburner et al., 2000). Indeed, positive and negative regulation of gene expression is essential to maintain cell cycle control.

Hence, NBS1 has diverse, important roles in all eukaryotic cells. Before moving on to the role of NBS1 in cancer, and rectal cancer in particular, it is useful to consider the clinical manifestations of Nijmegen breakage syndrome (NBS), from where the protein gets its name, as it is well described, systemic and insightful into the protein we here study. Thankfully, genetic testing has for decades been helping heterozygous parents (carrying *NBN* mutations) inform their future parenthood choices, as well as helping diagnose new cases (Varon et al., 2000).

In the autosomal recessive, congenital disorder, NBS (also known as ataxia telangiectasia variant 1), patients who all carry mutations in *NBN*, display several significant clinical features (Varon et al., 2002). For example, patients are typically characterised by a short stature, microcephaly, immunosuppression (and an increased infectious disease burden), cognitive dysfunction, vitiligo, and a common skeletal morphology leading to distinctive facial features (large forehead, large ears, small mandible and prominent nose) and many other conditions - including an increased risk to cancer (Berardinelli et al., 2013).

The inheritance pattern further demonstrates the redundancy in the DSB repair pathway and the potent tumour suppressor effect of NBS1, although it should be noted that heterozygous patients display heightened radiation sensitivity, and have been found to have fertility defects, but in the absence of the typical facial deformity of homozygous patients (Warcoin et al., 2009).

At present, estimates suggest that patients carrying some of the most penetrant lesions in NBS1 are 50 times more likely to develop cancer than non-affected individuals (Kondratenko et al., 2007). Strikingly, more than 40% of patients with a common deletion in the gene (described below) develop cancer before the age of 21 (Kondratenko et al., 2007). The life expectancy is estimated to lie between 30-40 years with routine medical monitoring and access to the latest

treatment, although estimates vary widely and many patients do not reach adulthood (The International Nijmegen Breakage Syndrome Study Group, 2000).

Historically, it appears that a large proportion of patients with Nijmegen breakage syndrome are of Slavonic origin, with many carrying a typical 5 nucleotide deletion (657-661 delACAAA) in the gene. Additional truncating mutations have been found in other patients with distinct haplotypes (International Nijmegen Breakage Syndrome Study Group, 2000). These truncated variants are less able to efficiently repair DNA DSB and participate in other essential cellular processes, and thus dysregulation arises. In a recent study looking at the Czech population, the frequency of heterozygotes for this mutation was found to be 1: 158-170 people (Seeman et al., 2004; Tauchi et al., 2002); with slightly lower rates being reported in other affected populations. Why this variant was maintained within this and other populations is unknown, but it may have served a selective advantage epistatically or in respect to the environment.

In Nijmegen breakage syndrome patients, microcephaly and decreased body growth is apparent from parturition in the majority of affected individuals, and normally becomes a progressive microcephaly where the brain remains further stunted in proportion to the rest of the body (Digweed and Sperling, 2004; Varon et al., 2000). By the age of 3, the majority of patients show the characteristic facial morphology associated with the condition. Notably, however, intellectual development does not appear to be aberrant before the age of 2, perhaps as during this time the body is investing energy into rapid tissue growth and environmental adaptation (i.e. in immunity; see Dopico et al., 2015), and not into the formation of social contacts, upon which many diagnoses of intellectual disability are made. Furthermore, given that the majority of infants do not begin to show preference for phonemes in their native language until approximately 6-7 months of age, defects in speech are also likely to be missed by parents and professional healthcare practitioners (Purves et al., 2001). Thus, it is likely that many of the defects and mechanisms leading to later observable phenotypes in these patients begin in utero as mutations progressively accumulate. Patients with *NBN* mutations are highly predisposed to the development of non-Hodgkin lymphoma, with many patients acquiring the disease during their teenage years (Seemanová et al., 2007).

Notably, affected individuals also show profound defects in immunity, presumably because many haematopoietic lineages undergo rapid, extensive clonal expansion and turnover in response to homeostasis and pathogens, and also due to genomic recombination in adaptive immunity. Given the high amount of DSBs generated during V(D)J recombination during thymocyte development and early B lymphocyte development (Murphy et al., 2008), NBS1 defects are likely to impact the process. Indeed, NBS1 and H2AX (introduced previously) have been shown to form nuclear foci that co-localise with the T cell receptor alpha locus in response to RAG-mediated cleavage at designated VDJ junctions (Chen, 2000). Furthermore, given that many specialised immunological lineages (such as those that show memory to pathogens) depend upon tissue-resident stem cells to replenish the surveying lymphocyte pool (Gattinoni et al., 2011), a similar clinical picture may emerge in these patients, as has been found to occur in multiple myeloma patients; where seamlessly endless putative cancer cells are generated,

with many clones showing hallmark cancer features (such as dysregulated proliferation), but with protracted periods of sub-clinical disease (Guedez and Stetler-Stevenson, 2010). Furthermore, in NBS patients, decreased IgG and IgA levels – which require high-fidelity DNA recombination in co-operation with RAG genes and plasma cell population expansion to arise – contribute to an increased risk of respiratory infections such as pneumonia, bronchitis and sinusitis; further demonstrating the link between the processes. Increased infection will lead to increased host cell death, mutation accumulation, and cancer predisposition.

Aside from Nijmegen Breakage Syndrome, NBS1 mutations have been associated with the development of aplastic anaemia, in which the bone marrow fails to produce adequate numbers of erythrocytes and haematopoietic cells. Consequently, patients exhibit pancytopenia and marrow hypoplasia (Shimada et al., 2004).

5.1.2 Associations between NBS1 and cancer

As discussed above the case of Nijmegen breakage syndrome patients, mutations within *NBN* are known to be associated with diverse types of cancer. In addition to non-Hodgkin lymphoma, and rectal cancer (which will be discussed in detail below), *NBN* mutations or defective function, have been associated with several different types of brain tumour, such as glioma and medulloblastoma (Ciara et al., 2010; Piekutowska-Abramczuk et al., 2010). Mutations in this locus have also been associated with the development of tumours in skeletal muscle, such as rhabdomyosarcomas (Seemanová et al., 2007).

NBN mutations also represent an important contribution to breast cancer, and expression of NBS1 has been shown to be a negative prognostic marker in lung and pancreatic cancer (Uhlen et al., 2017).

Polymorphisms in the NBS1 locus have been associated with colorectal cancer development, with a SNP in the 3'-untranslated region (rs2735383C/G) being the most associated with disease development (Li et al., 2015). Perhaps variation in this region underpins differences in NBS1 expression, impinging upon the ability of the protein to efficiently repair DNA damage. Indeed, follow-up work from the association revealed the rs2735383 C allele to have reduced binding affinity to the transcriptional regulator has-miR-509-5p, reducing expression (Li et al., 2015). This fits in with the studies we have previously discussed, in which decreased DSB repair proteins are associated with increased cancer risk.

A 2013 meta-analysis attempted to combine the available genetic association data concerning NBS1 and cancer to date. The work, based on more than 60 peer-reviewed publications and nearly 40,000 patients, showed rs275383 to be the most associated markers with cancer susceptibility, with an OR of 1.12 (Gao et al., 2013). The variants, I171V, 657del5 and R215W were also associated with cancer risk, albeit with each of them having low penetrance. Furthermore, studies from Uhrhammer and colleagues in France, found human 8q21.3 to harbour a gene with abundant allelic imbalance and with a tumour suppressive effect in colorectal cancer, namely NBS1 (Uhrhammer et al., 2000), although this is disputed (Varon et

al., 2002). This allelic imbalance has since been confirmed in other studies (Varon et al., 2002), suggesting that deleterious alleles contribute to cancer development; remembering that heterozygosity is rarely sufficient to cause disease.

In a Polish population (Slav populations being predisposed to lesions in this region), the R215W variant was found to be the most associated with colorectal cancer development, with the aforementioned 657del5 mutation being most associated with melanoma and non-Hodgkin lymphoma (Steffen et al., 2004). Furthermore, in this study, those with the 657del5 mutation were more likely to develop secondary tumours. The 657del5 mutation has also been found in colorectal and stomach cancer patients (di Masi and Antoccia, 2008), and deep-sequencing of additional patients is likely to reveal additional germ-line and somatic variants associated with cancer development.

Given the robust and reproducible associations reported between NBS1 and cancer, genetic testing of this locus has long been in the clinic, helping people avoid cancer risk factor and develop the psychological mechanisms which may be required to cope with such a condition. Indeed, the 657del5 mutation is routinely screened for in people with a history of breast cancer (Heikkinen et al., 2006). In the future, it is hoped that high-resolution, individual genetic data will allow for prophylactic consultation and disease prevention.

As part of the MRN complex, expression of NBS1 has also been described as an important prognostic marker. In a cohort of gastric cancer patients, Altan and colleagues report that high expression levels of the complex are associated with poor prognosis and chemoresistance (Altan et al., 2016). Similarly, work by Lee, Park and Lee show that NBS1 expression in epithelial ovarian cancer tissue was associated with advanced stage, high grade disease, and patients with high expression survived on average 48 months less than those with low/negative expression, demonstrating the applicability of NBS1 to inform upon the disease course (Lee et al., 2015). Furthermore, NBS1 expression levels also predicted the recurrence of epithelial ovarian cancers after treatment, further illustrating how NBS1 is associated with detrimental clinicohistopathological features.

Additional work by Wokolorczyk and colleagues has also shown worse survival in prostate cancer patients carrying the 657del5 mutation in *NBS1* (Wokolorczyk et al., 2012). Ehlers and Harbour has shown NBS1 expression to also be a useful prognostic marker in uveal melanoma, with high expression again being associated with worse survival (Ehlers and Harbour, 2005). As we have seen in other examples, the combinatorial panels of MRN complex proteins have been associated with the response to radiotherapy (Söderlund et al., 2007), although published studies exploiting NBS1 alone as a marker are lacking.

Several studies have also reported that different *NBN* mutations can influence chemotherapy success, as has been shown in non-small cell lung cancer, where platinum-based therapeutics were affected by rs1805794 and rs13312840, both of which lie within NBS1 (Xu et al., 2012), although why this occurs remains unknown.

5.2 Hypothesis and study overview

Given the role of the MRN complex in cancer, as discussed, and the evidence suggesting NBS1 could have a pathological role in rectal cancer, we here investigated whether expression of this protein in rectal cancer tissue was associated with patient survival outcomes and the response to radiotherapy.

To achieve our aims, we used immunohistochemical techniques and validated reagents to survey NBS1 expression in primary pathological samples, as described previously.

5.3 Results

Whilst convincing associations have been demonstrated by our studies with regards to ATM, MRE11 (and will be shown for RAD50 in the following Chapter), NBS1 proved to be the least informative marker in our studies; where expression levels in rectal tumours had weaker predictive value with respect to our known clinicohistopathological variables.

Nevertheless, demonstrating null associations is as important as demonstrating positive ones, especially when studying associations between different members of the same pathway/process; provided, of course, that the study is sufficiently powered statistically to detect the differences hypothesised to exist. With this in mind, and as a limited number of robust associations with respect to NBS1 were discovered in our relatively small study, large, independent rectal cancer cohorts should be studied to further delineate the mechanistic and clinical associations for *NBS1*.

5.3.1 Patient populations

A total of 266 patients were included in this study, similarly to what has been described in previous Chapters. In total, 176 (66.2%) patients were male, whilst 90 (33.8%) were female. The median age of participants whose samples were used in the study was 72 years (range: 35–100 years) (Table 5.1).

Of the 246 patients involved, 77 (31.3%) were treated with radiotherapy, and from these, 55 (71.4%) received preoperative therapy. During the time of sample collection, patients were followed for a median period of 3.16 years (range: 0−12.6 years), and the median time to death in the cohort was found to be 2.5 years after surgery (ranging between 0−11.1 years).

Table 5-1 Patient cohort characteristics for the NBS1 study

	All Patients (%)
Total, *n*	266
Age median	72
Sex	
Male	176/266 (66.2)
Female	90/266 (33.8)
Tumour stage	
T1–2	88/260 (33.8)
T3–4	172/260 (66.2)
Node stage	
N0	140/259 (54.1)
N1–2	119/259 (45.9)
Metastasis stage	
M0	223/240 (92.9)
M1	17/240 (7.1)
Histological Grade	
1–2	246/266 (92.5)
3	20/266 (7.5)
Vascular invasion	
Absent	201/263 (76.4)
Present	62/263 (23.6)
Perineural invasion	
Absent	220/263 (83.7)
Present	43/263 (16.3)
Radiotherapy	
Total	77/246 (31.3)
Neoadjuvant	55/77 (71.4)
Adjuvant	22/77 (28.6)

5.3.2 NBS1 expression in tumour sections

In this part of our investigation, NBS1 protein was detected using immunohistochemistry in a total of 260 samples. Following the staining scoring guidelines described in Chapter Two, analysed samples were taken from the tumour centre and periphery.

In samples taken from the tumour centre, 239/260 (92%) showed positive staining for NBS1, whilst twenty-one samples (8%) did not show evidence of NBS1 expression. When samples taken from the periphery of the tumour were considered, again 239 samples stained positive for NBS1, whilst 21 were negative. Staining intensity was also found to be similar between both the tumour centre and periphery (Figure 5.1).

As the majority of samples stained positive for NBS1, this is predicted to reduce our ability to spot differences between sub-groups of patients in the study.

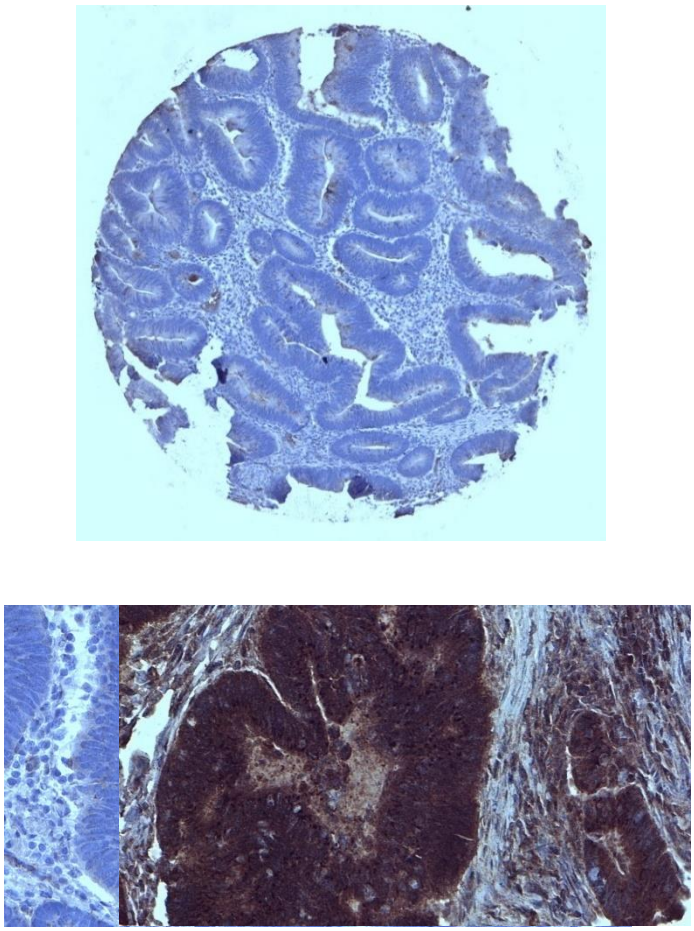


Figure 5-1 Representative examples of NBS1 staining (in rectal cancers) are shown.

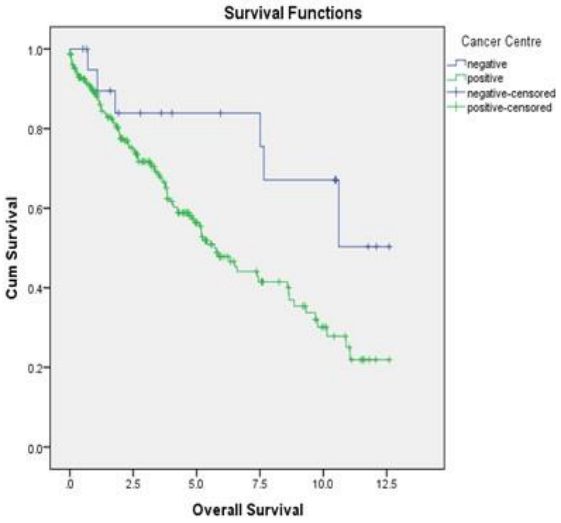
The left two images show negative NBS1 staining, whilst the right-hand panels show positive NBS1. Examples of positive and negative staining as shown for comparable anatomical sections of tumours on each row.

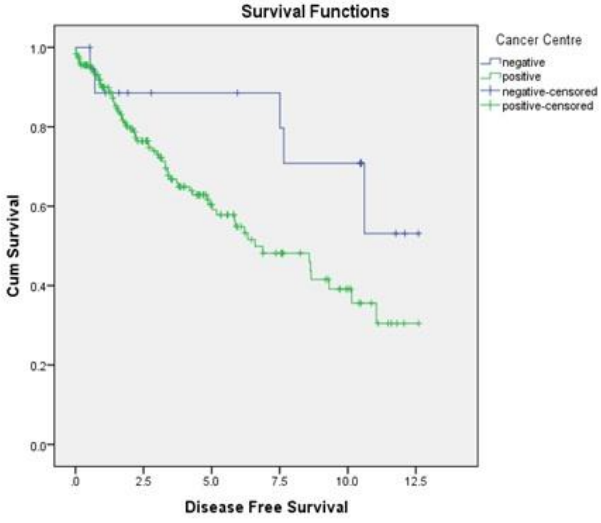
5.3.3 Association between NBS1 expression and clinicohistopathological features and prognosis

Although fewer and less *statistically robust* associations were reported for NBS1, by Kaplan-Meier survival analysis we did discover that positive expression in the tumour centre was associated with decreased OS (P = 0.02). Furthermore, we report a marginal association between NBS1 expression at this location with DFS (P = 0.059); as shown in Figure 5.2. In contrast, we did not observe any similar significant associations with NBS1 protein expression in the tumour periphery (Figure 5.3).

Furthermore, univariate cox regression analyses demonstrated that positive expression of NBS1 in the cancer centre, was significantly associated with reduced OS (p = 0.025), but not with worse DFS (p = 0.066). These analyses also allowed us to interrogate the interacting effects of clinicohistopathological variables, and revealed that advanced tumour stage (P = 0.009), vascular invasion (P = 0.001), perineural invasion (P < 0.001) and adjuvant radiotherapy (P = 0.01) were also associated with worse overall patient survival (Table 5.2).

Furthermore, multivariate Cox regression analyses demonstrated that NBS1 expression in the tumour centre, vascular invasion, perineural invasion, and adjuvant therapy, remained associated with overall survival in our cohort, when interacting variables were taken into account. Please see Table 5.3.

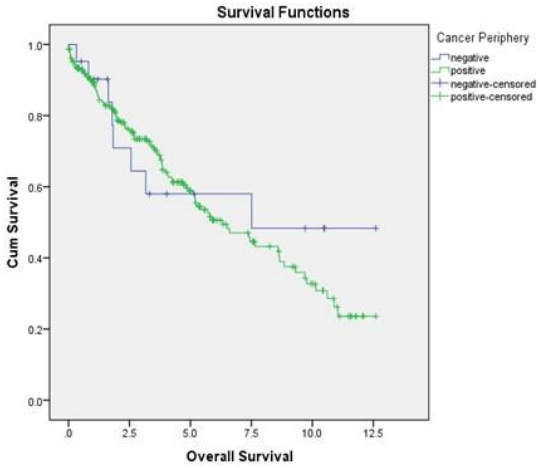


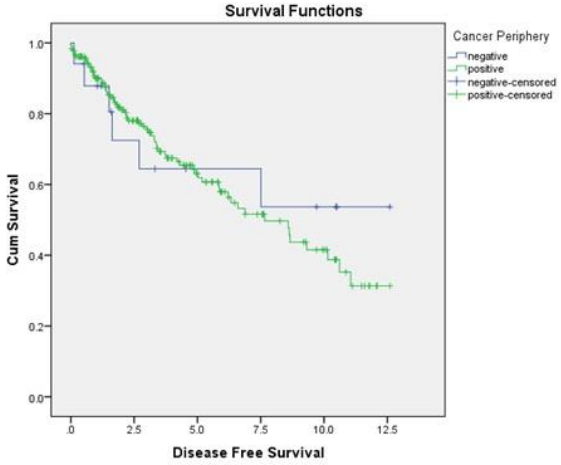


Overall Comparisons			
	Chi-square	DF	Sig.
Log Rank (Mantel-Cox)	5.400	1	0.020
Test of equality of survival distributions for the different levels of cancer centre expression.			

Overall Comparisons			
	Chi-square	DF	Sig.
Log Rank (Mantel-Cox)	3.576	1	0.059
Test of equality of survival distributions for the different levels of cancer centre expression.			

Figure 5-2 Associations between NBS1 expression in the tumour centre and patient survival are shown. Results are from Kaplan-Meier Survival Analysis, as previously described.





Overall Comparisons			
	Chi-square	DF	Sig.
Log Rank (Mantel-Cox)	0.352	1	0.553
Test of equality of survival distributions for the different levels of cancer periphery expression.			

Overall Comparisons			
	Chi-square	DF	Sig.
Log Rank (Mantel-Cox)	0.181	1	0.671
Test of equality of survival distributions for the different levels of cancer periphery expression.			

Figure 5-3 Associations between NBS1 expression in the tumour periphery and patient survival are shown. A Kaplan-Meier Survival Analysis was again used.

Table 5-2 Univariate analysis of Nbs1 with disease-free survival and overall survival

	DFS			OS		
Variables	HR	95% CI	*P* Value	HR	95% CI	*P* Value
Nbs1 TC						
positive	2.370	0.944-5.950	.066	2.586	1.127-5.936	.025
negative						
Nbs1 TP						
positive	1.199	0.518-2.778	.671	1.244	0.604-2.562	.554
negative						
Age						
≤72	1.279	0.785-2.086	0.323	1.334	0.897-1.984	.155
>72						
Sex						
Male	1.074	0.657-1.754	0.776	1.092	0.734-1.625	.664
Female						
Tumour stage						
T1–2	1.643	0.983-2.747	.058	1.796	0.897-1.984	.009
T3–4						
Node stage						
Negative	1.138	0.709-1.827	.593	1.454	0.987-2.140	.058
Positive						
Grade						
1–2	1.646	0.712-3.804	0.244	1.561	0.836-2.916	.162
3						
Vascular invasion						
Absent	1.188	0.650-2.171	.575	2.030	1.340-3.075	.001
Present						
Perineural invasion						
Absent	2.534	1.310-4.157	0.001	2.48	1.594-3.859	<0.001
Present						
Adjuvant therapy						
No	0.65	0.373-1.134	0.129	0.506	0.301-0.850	0.01
Yes						
Neoadjuvant therapy						
No	1.147	0.672-1.957	0.616	0.63	0.556-1.427	0.63
Yes						

Table 5-3 Multivariate analysis of NBS1 with overall survival

	Multivariate		
Variables	HR	95%	*P* Value
Nbs1 TC	2.396	1.016-5.650	0.046
Vascular invasion	1.805	1.057-3.083	0.031
Perineural invasion	2.537	1.408-4.573	0.002
Adjuvant therapy	0.290	0.165-0.507	<0.001

5.4 Summary

Although limited associations were discovered with respect to NBS1 expression in our samples, we do report that rectal cancers express significant levels of NBS1 protein. In-keeping with a role for the DDR pathway in maintaining genome integrity in the face of oncogenesis, and that the cancer has a high mutation burden.

In addition, we are able to report marginal associations between NBS1 expression in the tumour centre and patient overall survival (as well as the expected clinicohistopathological associations reported; i.e. perineural invasion (PNI)-positivity), although additional samples need to be studied to confirm this observation; given the marginal statistical association.

Patients in whom NBS1 staining was absent, perhaps due to disease or treatment stage, appear to drive the statistical association between high expression and overall survival. Sub-sets of rectal cancer patients with mutations in individual components of the DDR/MMR protein machinery would be informative for learning about disease development and pathway redundancy; such samples are likely to become increasingly available as more genomes get sequenced. Furthermore, the relative redundancy of different DDR proteins remains a controversial issue.

These results suggest that NBS1 is a less informative marker than ATM, MRE11 and RAD50 with respect to rectal cancer prognosis and disease pathophysiology, although, as noted, additional patient samples are required to convincingly demonstrate this.

Further exploration NBS1 expression utility in this context will be explored in Chapter Six, where the protein is described as forming part of a three-component biomarker panel with other MRN complex members.

5.5 References

Altan, B., Yokobori, T., Ide, M., Bai, T., Yanoma, T., Kimura, A., Kogure, N., Suzuki, M., Bao, P., Mochiki, E., et al. (2016). High Expression of MRE11-RAD50-NBS1 Is Associated with Poor Prognosis and Chemoresistance in Gastric Cancer. Anticancer Res. *36*, 5237–5247.

Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P., Dolinski, K., Dwight, S.S., Eppig, J.T., et al. (2000). Gene ontology: Tool for the unification of biology. Nat. Genet. *25*, 25–29.

Balasubramanian, N., Bai, P., Buchek, G., Korza, G., and Weller, S.K. (2010). Physical interaction between the herpes simplex virus type 1 exonuclease, UL12, and the DNA double-strand break-sensing MRN complex. J. Virol. *84*, 12504–12514.

Berardinelli, F., Masi, A., and Antoccia, A. (2013). NBN Gene Polymorphisms and Cancer Susceptibility: A Systemic Review. Curr. Genomics *14*, 425–440.

Chen, H.T. (2000). Response to RAG-Mediated V(D)J Cleavage by NBS1 and gamma-H2AX. Science. *290*, 1962–1964.

Ciara, E., Piekutowska-Abramczuk, D., Popowska, E., Grajkowska, W., Barszcz, S., Perek, D., Dembowska-Bagińska, B., Perek-Polnik, M., Kowalewska, E., Czajńska, A., et al. (2010). Heterozygous germ-line mutations in the NBN gene predispose to medulloblastoma in pediatric patients. Acta Neuropathol. *119*, 325–334.

Difilippantonio, S., Celeste, A., Fernandez-Capetillo, O., Chen, H.T., San Martin, B.R., Van Laethem, F., Yang, Y.P., Petukhova, G. V., Eckhaus, M., Feigenbaum, L., et al. (2005). Role of Nbs1 in the activation of the Atm kinase revealed in humanized mouse models. Nat. Cell Biol. *7*, 675–685.

Digweed, M., and Sperling, K. (2004). Nijmegen breakage syndrome: Clinical manifestation of defective response to DNA double-strand breaks. DNA Repair (Amst). *3*, 1207–1217.

di Masi, A., and Antoccia, A. (2008). NBS1 Heterozygosity and Cancer Risk. Curr. Genomics *9*, 275–281.

Dopico, X.C., Evangelou, M., Ferreira, R.C., Guo, H., Pekalski, M.L., Smyth, D.J., Cooper, N., Burren, O.S., Fulford, A.J., Hennig, B.J., et al. (2015). Widespread seasonal gene expression reveals annual differences in human immunity and physiology. Nat. Commun. *6*, 7000.

Ehlers, J.P., and Harbour, J.W. (2005). NBS1 expression as a prognostic marker in uveal melanoma. Clin Cancer Res *11*, 1849–1853.

Enriquez-Rios, V., Dumitrache, L.C., Downing, S.M., Li, Y., Brown, E.J., Russell, H.R., and McKinnon, P.J. (2017). DNA-PKcs, ATM, and ATR Interplay Maintains Genome Integrity during Neurogenesis. J. Neurosci. *37*, 893–905.

Furman, D., Jojic, V., Sharma, S., Shen-Orr, S.S., L. Angel, C.J., Onengut-Gumuscu, S., Kidd, B.A., Maecker, H.T., Concannon, P., Dekker, C.L., et al. (2015). Cytomegalovirus infection enhances the immune response to influenza. Sci. Transl. Med. *7*, 281ra43-281ra43.

Gao, P., Ma, N., Li, M., Tian, Q.B., and Liu, D.W. (2013). Functional variants in NBS1 and cancer risk: Evidence from a meta-analysis of 60 publications with 111 individual studies. Mutagenesis *28*, 683–697.

Gattinoni, L., Lugli, E., Ji, Y., Pos, Z., Paulos, C.M., Quigley, M.F., Almeida, J.R., Gostick, E., Yu, Z., Carpenito, C., et al. (2011). A human memory T cell subset with stem cell-like

properties. Nat. Med. *17*, 1290–1297.
Guedez, L., and Stetler-Stevenson, W.G. (2010). The prognostic value of TIMP-1 in multiple myeloma. Leuk. Res. *34*, 576–577.
Heikkinen, K., Rapakko, K., Karppinen, S.-M., Erkko, H., Knuutila, S., Lundán, T., Mannermaa, A., Børresen-Dale, A.-L., Borg, A., Barkardottir, R.B., et al. (2006). RAD50 and NBS1 are breast cancer susceptibility genes associated with genomic instability. Carcinogenesis *27*, 1593–1599.
Hutchinson, I., Whiteley, A., Browne, H., and Elliott, G. (2002). Sequential localization of two herpes simplex virus tegument proteins to punctate nuclear dots adjacent to ICP0 domains. J. Virol. *76*, 10365–10373.
The, I. (2000). Nijmegen breakage syndrome. The International Nijmegen Breakage Syndrome Study Group. Arch Dis Child *82*, 400–406.
Kang, J., Bronson, R.T., and Xu, Y. (2002). Targeted disruption of NBS1 reveals its roles in mouse development and DNA repair. EMBO J. *21*, 1447–1455.
Kobayashi, J., Tauchi, H., Sakamoto, S., Nakamura, A., Morishima, K. ichi, Matsuura, S., Kobayashi, T., Tamai, K., Tanimoto, K., and Komatsu, K. (2002). NBS1 localizes to gamma-H2AX foci through interaction with the FHA/BRCT domain. Curr. Biol. *12*, 1846–1851.
Kobayashi, J., Antoccia, A., Tauchi, H., Matsuura, S., and Komatsu, K. (2004). NBS1 and its functional role in the DNA damage response. DNA Repair (Amst). *3*, 855–861.
Komatsu K. (2016). NBS1 and multiple regulations of DNA damage response. J Radiat Res. *57(Suppl 1)*:i11–ii7.
Kondratenko, I., Paschenko, O., Polyakov, A., and Bologov, A. (2007) Nijmegen Breakage Syndrome. In: Shurin M.R., Smolkin Y.S. (eds) Immune-Mediated Diseases. *Advances in Experimental Medicine and Biology*, vol 601. Springer, New York, NY
Lee, J.-H., and Paull, T.T. (2007). Activation and regulation of ATM kinase activity in response to DNA double-strand breaks. Oncogene *26*, 7741–7748.
Lee, Y.-K., Park, N.-H., and Lee, H. (2015). Clinicopathological values of NBS1 and DNA damage response genes in epithelial ovarian cancers. Exp. Mol. Med. *47*, e195.
Li, J.T., Zhong, B.Y., Xu, H.H., Qiao, S.Y., Wang, G., Huang, J., Fan, H.Z., and Zhao, H.C. (2015). Associations between NBS1 polymorphisms and colorectal cancer in Chinese population. PLoS One *10*. e0132332.
Lou, D.I., Kim, E.T., Meyerson, N.R., Pancholi, N.J., Mohni, K.N., Enard, D., Petrov, D.A., Weller, S.K., Weitzman, M.D., and Sawyer, S.L. (2016). An Intrinsically Disordered Region of the DNA Repair Protein Nbs1 Is a Species-Specific Barrier to Herpes Simplex Virus 1 in Primates. Cell Host Microbe *20*, 178–188.
McEwan, W.A., Tam, J.C.H., Watkinson, R.E., Bidgood, S.R., Mallery, D.L., and James, L.C. (2013). Intracellular antibody-bound pathogens stimulate immune signaling via the Fc receptor TRIM21. Nat. Immunol. *14*, 327–336.
Milward, E.A., Moscato, P., Riveros, C., and Johnstone, D.M. (2014). Beyond statistics: a new combinatorial approach to identifying biomarker panels for the early detection and diagnosis of Alzheimer's disease. J. Alzheimers. Dis. *39*, 211–217.
Misteli, T., and Soutoglou, E. (2009). The emerging role of nuclear architecture in DNA repair and genome maintenance. Nat. Rev. Mol. Cell Biol. *10*, 243–254.
Murphy, K., Travers, P., Walport, M., and Janeway, C. (2008). *Janeway's immunobiology*.

New York: Garland Science.
Naka, K., Ikeda, K., and Motoyama, N. (2002). Recruitment of NBS1 into PML oncogenic domains via interaction with SP100 protein. Biochem. Biophys. Res. Commun. *299*, 863–871.
Orchard, S., Ammari, M., Aranda, B., Breuza, L., Briganti, L., Broackes-Carter, F., Campbell, N.H., Chavali, G., Chen, C., Del-Toro, N., et al. (2014). The MIntAct project - IntAct as a common curation platform for 11 molecular interaction databases. Nucleic Acids Res. *42*, D358-D363.
Passananti, C., and Fanciulli, M. (2007). The anti-apoptotic factor Che-1/AATF links transcriptional regulation, cell cycle control, and DNA damage response. Cell Div. *2*, 21.
Piekutowska-Abramczuk, D., Ciara, E., Popowska, E., Grajkowska, W., Dembowska-Bagińska, B., Kowalewska, E., Czajńska, A., Perek-Polnik, M., Roszkowski, M., Syczewska, M., et al. (2010). The frequency of NBN molecular variants in pediatric astrocytic tumors. J. Neurooncol. *96*, 161–168.
Purves, D., Augustine, G., Fitzpatrick, D., Katz, L., LaMantia, A., McNamara, J. and Williams, S. (2001). The Development of Language: A Critical Period in Humans. In: *Neuroscience 2nd edition*, Sinauer Associates, Sunderland, MA.
Puschnik, A.S., Majzoub, K., Ooi, Y.S., and Carette, J.E. (2017). A CRISPR toolbox to study virus-host interactions. Nat. Rev. Microbiol. *15*, 351–364.
Seeman, P., Gebertová, K., Paděrová, K., Sperling, K., and Seemanová, E. (2004). Nijmegen breakage syndrome in 13% of age-matched Czech children with primary microcephaly. Pediatr. Neurol. *30*, 195–200.
Seemanová, E., Jarolim, P., Seeman, P., Varon, R., Digweed, M., Swift, M., and Sperling, K. (2007). Cancer risk of heterozygotes with the NBN founder mutation. J. Natl. Cancer Inst. *99*, 1875–1880.
Seno, J.D., and Dynlacht, J.R. (2004). Intracellular Redistribution and Modification of Proteins of the Mre11/Rad50/Nbs1 DNA Repair Complex Following Irradiation and Heat-Shock. J. Cell. Physiol. *199*, 157–170.
Shimada, H., Shimizu, K., Mimaki, S., Sakiyama, T., Mori, T., Shimasaki, N., Yokota, J., Nakachi, K., Ohta, T., and Ohki, M. (2004). First case of aplastic anemia in a Japanese child with a homozygous missense mutation in the NBS1 gene (I171V) associated with genomic instability. Hum. Genet. *115*, 372–376.
Söderlund, K., Stål, O., Skoog, L., Rutqvist, L.E., Nordenskjöld, B., and Askmalm, M.S. (2007). Intact Mre11/Rad50/Nbs1 Complex Predicts Good Response to Radiotherapy in Early Breast Cancer. Int. J. Radiat. Oncol. Biol. Phys. *68*, 50–58.
Speir, M.L., Zweig, A.S., Rosenbloom, K.R., Raney, B.J., Paten, B., Nejad, P., Lee, B.T., Learned, K., Karolchik, D., Hinrichs, A.S., et al. (2016). The UCSC Genome Browser database: 2016 update. Nucleic Acids Res. *44*, D717–D725.
Staras, S.A., Dollard, S.C., Radford, K.W., Flanders, W.D., Pass, R.F., and Cannon, M.J. (2006). Seroprevalence of cytomegalovirus infection in the United States, 1988-1994. Clin. Infect. Dis. *43*, 1143–1151.
Steffen, J., Varon, R., Mosor, M., Maneva, G., Maurer, M., Stumm, M., Nowakowska, D., Rubach, M., Kosakowska, E., Ruka, W., et al. (2004). Increased cancer risk of heterozygotes with NBS1 germline mutations in Poland. Int. J. Cancer *111*, 67–71.
Stucki, M., and Jackson, S.P. (2006). γH2AX and MDC1: Anchoring the DNA-damage-

response machinery to broken chromosomes. DNA Repair (Amst). *5*, 534–543.
Tanić, M., and Beck, S. (2017). Epigenome-wide association studies for cancer biomarker discovery in circulating cell-free DNA: technical advances and challenges. Curr. Opin. Genet. Dev. *42*, 48–55.
Tauchi, H., Kobayashii, J., Morishima, K.I., Matsuura, S., Nakamura, A., Shiraishi, T., Ito, E., Masnada, D., Delia, D., and Komatsu, K. (2001). The forkhead-associated domain of NBS1 is essential for nuclear foci formation after irradiation but not essential for hRAD50??hMRE11??-NBS1 complex DNA repair activity. J. Biol. Chem. *276*, 12–15.
Tauchi, H., Matsuura, S., Kobayashi, J., Sakamoto, S., and Komatsu, K. (2002). Nijmegen breakage syndrome gene, NBS1, and molecular links to factors for genome stability. Oncogene *21*, 8967–8980.
Uhlen, M., Fagerberg, L., Hallstrom, B.M., Lindskog, C., Oksvold, P., Mardinoglu, A., Sivertsson, A., Kampf, C., Sjostedt, E., Asplund, A., et al. (2015). Tissue-based map of the human proteome. Science,*347*, 1260419–1260419.
Uhlen, M., Zhang, C., Lee, S., Sjöstedt, E., Fagerberg, L., Bidkhori, G., Benfeitas, R., Arif, M., Liu, Z., Edfors, F., et al. (2017). A pathology atlas of the human cancer transcriptome. Science. *357*, eaan2507.
Uhrhammer, N., Bay, J., Gosse-Brun, S., Kwiatkowski, F., Rio, P., Daver, A., and Bignon, Y. (2000). Allelic imbalance at NBS1 is frequent in both proximal and distal colorectal carcinoma. Oncol Rep *7*, 427–431.
UniProt Consortium (2015). UniProt: a hub for protein information. Nucleic Acids Res. *43*, D204-12.
Varon, R., Seemanova, E., Chrzanowska, K., et al. (2000). Clinical ascertainment of Nijmegen breakage syndrome (NBS) and prevalence of the major mutation, 657del5, in three Slav populations. Eur J Hum Genet. *8*, 900–902.
Varon, R., Gosse-Brun, S., Bignon, Y.J., Sperling, K., and Uhrhammer, N. (2002). Nijmegen breakage syndrome gene (NBS1) is not the tumor suppressor gene at 8q21.3 involved in colorectal carcinoma. Oncol. Rep. *9*, 709–711.
Vega, F.M., Sevilla, A., and Lazo, P.A. (2004). p53 Stabilization and Accumulation Induced by Human Vaccinia-Related Kinase 1. Mol. Cell. Biol. *24*, 10366–10380.
Wang, J.Q., Chen, J.H., Chen, Y.C., Chen, M.Y., Hsieh, C.Y., Teng, S.C., and Wu, K.J. (2013). Interaction between NBS1 and the mTOR/Rictor/SIN1 Complex through Specific Domains. PLoS One. *8*, e65586.
Warcoin, M., Lespinasse, J., Despouy, G., D'Enghien, C.D., Laugé, A., Portnoï, M.F., Christin-Maitre, S., Stoppa-Lyonnet, D., and Stern, M.H. (2009). Fertility defects revealing germline biallelic nonsense NBN mutations. Hum. Mutat. *30*, 424–430.
Waterworth, W.M., Altun, C., Armstrong, S.J., Roberts, N., Dean, P.J., Young, K., Weil, C.F., Bray, C.M., and West, C.E. (2007). NBS1 is involved in DNA repair and plays a synergistic role with ATM in mediating meiotic homologous recombination in plants. Plant J. *52*, 41–52.
Weinstein, J.N., Collisson, E.A., Mills, G.B., Shaw, K.R., Ozenberger, B.A., Ellrott, K., et al. (2013). Cancer Genome Atlas Research Network The Cancer Genome Atlas Pan-Cancer analysis project. Nat Genet. *45*, 1113–20.
Weng, J.H., Hsieh, Y.C., Huang, C.C.F., Wei, T.Y.W., Lim, L.H., Chen, Y.H., Ho, M.R., Wang, I., Huang, K.F., Chen, C.J., et al. (2015). Uncovering the Mechanism of Forkhead-

Associated Domain-Mediated TIFA Oligomerization That Plays a Central Role in Immune Responses. Biochemistry *54*, 6219–6229.
Wokolorczyk, D., Kluzniak, W., Cybulski, C., and Lubinski, J. (2012). NBS1 Mutation and prognosis of prostate cancer. Hered Cancer Clin Pr. *10*, A29 Supplement.
Wu, Z.-H. (2016). Phenotypes and genotypes of the chromosomal instability syndromes. Transl. Pediatr. *5*, 79–83.
Xu, J.-L., Hu, L.-M., Huang, M.-D., Zhao, W., Yin, Y.-M., Hu, Z.-B., Ma, H.-X., Shen, H.-B., and Shu, Y.-Q. (2012). Genetic Variants of NBS1 Predict Clinical Outcome of Platinum-based Chemotherapy in Advanced Non-small Cell Lung Cancer in Chinese. Asian Pacific J. Cancer Prev. *13*, 851–856.
Yu, X. (2003). The BRCT Domain Is a Phospho-Protein Binding Domain. Science.*302*, 639–642.
Zdobnov, E.M., Tegenfeldt, F., Kuznetsov, D., Waterhouse, R.M., Simao, F.A., Ioannidis, P., Seppey, M., Loetscher, A., and Kriventseva, E. V. (2017). OrthoDB v9.1: Cataloging evolutionary and functional annotations for animal, fungal, plant, archaeal, bacterial and viral orthologs. Nucleic Acids Res. *45*, D744–D749.
Zhang, C., and Kuspa, A. (2009). Transcriptional down-regulation and rRNA cleavage in Dictyostelium discoideum mitochondria during Legionella pneumophila infection. PLoS One. *4*, e5706.
Zhang, Y., Zhou, J., and Lim, C.U. (2006). The role of NBS1 in DNA double strand break repair, telomere stability, and cell cycle checkpoint control. Cell Res. *16*, 45–54.
Zhou, B.B.S., and Elledge, S.J. (2000). The DNA damage response: Putting checkpoints in perspective. Nature *408*, 433–439.
Zhu, J., Petersen, S., Tessarollo, L., and Nussenzweig, A. (2001). Targeted disruption of the Nijmegen breakage syndrome gene NBS1 leads to early embryonic lethality in mice. Curr. Biol. *11*, 105–109.

CHAPTER 6
THE ASSOCIATION BETWEEN RAD50 EXPRESSION AND SURVIVAL OUTCOMES IN RECTAL CANCER, AND THE USE OF A COMBINATORIAL MRN COMPLEX EXPRESSION BIOMARKER PANEL IN THE SAME CONTEXT

6.1 Introduction

This chapter presents the results on the study of the expression of RAD50, the MRN complex ATPaseand a major activator and modulator of the complex's activity in relation to rectal cancer radiotherapy response and survival.

This Chapter will also cover our experiments concerned with the development and application of a combinatorial three-marker panel, composed of MRE11, RAD50 and NBS1 staining, to our rectal cancer samples.

By using this combined approach, an increased amount of information can be captured concerning the status of the DSB repair machinery in every cancer sample. Given the marked cancer-associated variation in many of these genes, being able to monitor how the entire pathway is working is important. For example, using the three-marker panel, patients with reduced ATM expression but normal MRN complex expression could be distinguished, allowing different therapeutics to be used when genetics in considered concomitantly.

6.1.1 RAD50: biochemistry and mechanism of action

In humans, the DNA repair protein RAD50 (RAD50) is encoded by a gene of the same name that is found on chromosome 5 (5q31.1) (Figure 6.1). RAD50 is a member of the structural maintenance of chromosome (SMC) family (Ball and Yokomori, 2001), and being an ATPase, is thought to be the primary means by which the MRN complex exploits energy stored in ATP to facilitate strand repair (Hopfner et al., 2000a); although the precise mechanistic details of this function remain to be determined (Kinoshita et al., 2009). SMC proteins are also well known to participate in chromatin condensation, and high-order organisation of genomic DNA (Strunnikov, 1998).

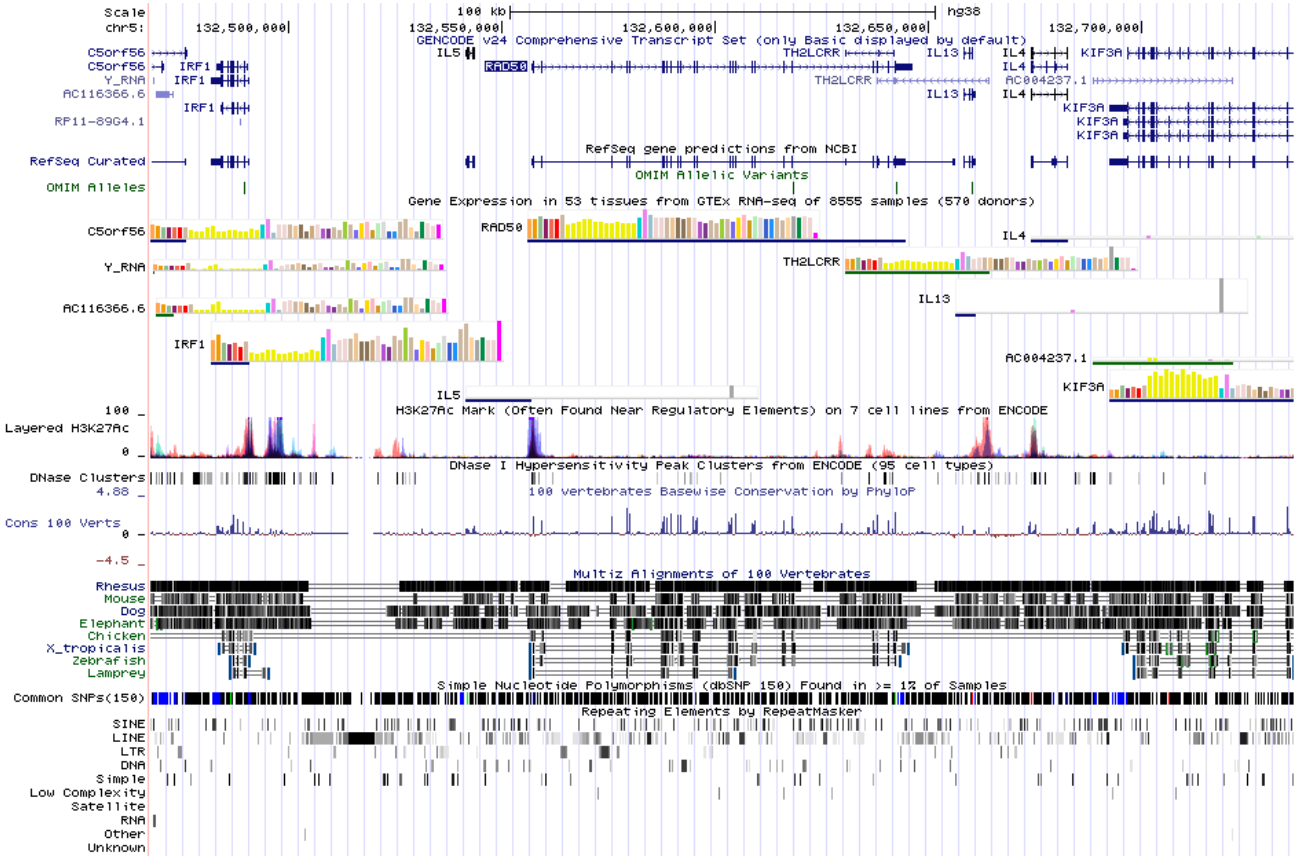


Figure 6-1The genomic context of RAD50 is shown via graphical output from the UCSC Genome Browser

(Speir et al., 2016). As can be seen by the blue tracks (position marked on the right-hand side of the image), the RAD50 gene is located in a region rich in cytokine genes, such as *IL-4*,*IL-5* and *IL-13*, as well as the immune activator, interferon response factor 1 (*IRF1*), and the microtubule motor involved in nucleic acid and protein transport, *KIF3A*. This genomic organisation may allow for concomitant regulation of the DDR and inflammation, especially in cells like neutrophils, eosinophils and basophils, who produce such cytokines and proliferate rapidly in response to inflammatory triggers. Exonic regions in the blue tracks are represented by vertical dissecting lines/rectangles; depending on exon size. Furthermore, the start of the RAD50 gene is marked by a H3K27AC modification in 7 cell lines analysed as part of the ENCODE dataset (Ecker et al., 2012), allowing for the gene's frequent transcriptional regulation. Conservation of the exonic RAD50 sequence between species is illustrated in the black tracks towards the bottom of the image, noting that very high levels of intronic conservation is present between *Rhesus macaques* and *Homo sapiens*.

RAD50 shows marked homology to proteins in highly diverse species (Hopfner et al., 2000b). In yeast, for example, RAD50 poses numerous catalytic activities which are essential for DSB repair to take place (Anderson et al., 2001). In-keeping with the well described roles of other members of the MRN complex and fertility, disruption of RAD50 in *Arabidopsis* leads to plant

sterility, and experiments exploiting the DSB-inducing agent, methyl methane sulphonate, demonstrated the protein to be involved in DSB repair in plants (Gallego et al., 2001). Furthermore, homologues of RAD50 have been found in prokaryotic archaea, where it again helps mediate the response to gamma radiation and homologous recombination repair (De Jager et al., 2004). Abrogation of the gene in mice can lead to embryonic lethality/deformity and a high sensitivity to ionising radiation (Luo et al., 1999), where defects predispose to premature ageing, cancer and other systemic abnormalities.

The major isoform of the protein is 1,312 amino acids in length (approximately 154 kDa weight), and in contrast to the other members of the MRN complex (which is formed by two heterodimers of RAD50/MRE11 and a single NBS1), RAD50 shows a long, repeating, internal coiled-coil domain composed of five alpha helices between 29-291 amino acids long, that allows the N- and C-termini of the protein to form a globular structure (The UniProt Consortium, 2015). Indeed, a protein with a coiled-coil domain is required for MRN complex function (Hohl et al., 2010). Crucially, this folding is dependent on the acquisition of a zinc ion by RAD50 (which is achieved via the Zinc-hook located between residues 635-734 – separating the large internal coiled regions), forming a metal-bound homodimer with a newly-formed ABC ATPase head region (Hopfner et al., 2002); mutations at both the extreme C- and N-terminal ends of the protein can interfere with ATP hydrolysing ability. This region of the protein is essential for mediating the MRE11 interaction, whilst the zinc-hook contains two cysteine residues which bind the zinc ion, with each residue contributing one bond across the homodimer, which forms a V-shaped rod composition as a whole (Park et al., 2017).

Present consensus states that RAD50's ability to change conformation from *open* to *closed* in the presence of free DNA ends, is essential for tethering DNA strands together for repair to take place – zinc-mediated tethering of DNA ends is thought to depend upon RAD50 (Bhaskara et al., 2007; Connelly and Leach, 2002). Thus, RAD50 is thought to be essential for allowing for sequence homology searches and ligation to take place.

RAD50 is also responsible for curtailing the nuclease activity of MRE11 and for activating DNA ligase (Trujillo and Sung, 2001) – co-ordinating and fine-tuning the repair. Additional reports suggest that RAD50 is involved in stimulating ATM (Deshpande et al., 2017), as other members of the MRN complex have been reported to do. Thus, RAD50 is central to DNA repair and starting cell cycle arrest processes. Aside from its functions in DSB repair, RAD50 – like the other members of the MRN complex – is involved in telomere maintenance, and meiosis (Vannier et al., 2006). For plants PSH1 was found to be essential for transport of RAD50 from the cytoplasm to execute meiotic recombination (Ronceret et al., 2009). Although the mechanisms in humans are probably different, RAD50 is still thought to be essential so similar meiotic processes, and has been associated with infertility (Handel and Schimenti, 2010). One of few studies in mammalian cells in respect to meiosis has shown that the MRN complex co-opts Ctp1 to carry out meiotic DNA DSB repair (Ma et al., 2015). Mutations in the nuclease Ctp1 is also associated with inherited disorders such as Seckel syndrome, which is characterised by dwarfism, microcephaly and abnormal skeletal morphology, similar to what has been discussed for NBS1 mutations (Qvist et al., 2011). This study also found the PI3K

kinase, ATR (involved in activating the DNA damage checkpoint), to be a risk locus for Seckel syndrome.

Unsurprisingly, RAD50 displays a predominantly nuclear localisation, although it can be re-targeted as required by different cellular states, such as infection or temperature stress. For example, when a cell becomes infected and foreign dsDNA is present in the cytoplasm, RAD50 forms a complex with the innate immune system adaptor protein, CARD9) (Roth et al., 2014). Together, they activate NF-kappa-B, which drives IL-1-beta production – a central molecule across innate and adaptive immunity and the hypothalamic pyretic axis.

RAD50 has also been documented by several reports to be a member of the BRCA1-associated genome surveillance complex (BASC), which consists of a group of proteins involved in DNA damage repair and cell cycle control (Wang et al., 2000). Amongst the proteins found in the BASC complex are the TSGs and DNA damage response proteins, such as MSH2, MSH6, MLH1, ATM and BLM, as well as MRE11 and NBS1. Similar to the amounts of MRN complex that has been described in nuclear dots, this spatial organisation likely allows for DNA damage sensing and repair to take place more quickly; being compartmentalised in close proximity to lesions and having the inherent ability to modify chromatin structure to execute repair.

Bioinformatic approaches that rely on vast databases and chemical and biophysical predictions are useful first steps when considering protein function, as many features may have been seen before. Indeed, the multiple associated Gene Ontology functions attributed to RAD50 included: ATPase activity; adenylate kinase activity; DNA binding, G-quadruplex binding; protein binding; metal ion binding; and single- and double-stranded telomeric DNA binding, demonstrating the range of activities and capabilities of the protein, and agreeing with what has been reported in the literature (Ashburner et al., 2000). The same functional analysis of the protein suggests that it is involved with telomere capping, nucleic acid phosphodiester bond hydrolysis, and mediating Pi3K-dependent signal transduction.

Given its wide range of functions, it is also not surprising to learn that lesions in this gene are associated with debilitating clinical diseases. Mutations in *RAD50* are known to cause Nijmegen breakage syndrome-like conditions. For example, Waltes and colleagues describe in detail a patient previously diagnosed as having NBS (mental retardation, facial malformation, microcephaly) (Waltes et al., 2009). Unlike true NBS patients, this patient did not harbour lesions in the *NBN* gene, but rather was compound heterozygous for mutations in *RAD50* that led to the generation of an unstable protein. Cells from the patient displayed increased radiation sensitivity, abnormal MRN complex foci formation, chromosomal instability, defective ATM activity, and defects in G1/S cell cycle regulation; all of which could be rescued by wild-type RAD50 *in vitro*. The fact that very few of such cases have been reported in the medical literature is likely due to the highly penetrant and deleterious nature of mutations within or around *RAD50*; remembering that its deletion leads to mouse embryo death.

Common polymorphisms in RAD50 are also associated with a range of human traits as can be seen by data from genome-wide associatgion studies in the NHGRI-EBI GWAS Catalogue (MacArthur et al., 2017). RAD50 variants are strongly associated with the levels of circulating eosinophils and basophils; the strongest being rs2706345, an intronic variant in RAD50. Additional intronic variants in the RAD50-IL13 region have been associated with atopic dermatitis, asthma and inflammatory skin disease (MacArthur et al., 2017), although the various cytokines present in the RAD50 region would appear to be the most likely genetic cause of that biology (Figure 6.1). Indeed, other SNPs in the region (such as rs2040704) have been reported to increase IgE titres more than 13%, again suggesting inflammatory pathology linked to IL-13, an important mediator of allergic inflammation (Weidinger et al., 2008). Various cis- and trans-regulatory processes and regulatory roles have been widely reported at other loci (Davison et al., 2012).

Another SNP in the region (within a RAD50 intron), rs13164856, marks a haplotype encompassing RAD50, C5ORF56 and IRF1, that is associated with polycystic ovary syndrome (Day et al., 2015), perhaps suggesting defects in germ cell maintenance and reproductive tract function.

6.1.2 Associations between RAD50 and cancer

Thus, compared to the other members of the DNA DSB repair pathway discussed thus far, it seems that less is known concerning RAD50 and cancer development.

Perhaps the relative scarcity of information concerning RAD50 and cancer is due to increased redundancy in the MRN complex's zinc-dependent ATP-hydrolysing component, compared to the ATM (a protein)-binder NBS1, or the nucleic acid-editing MRE11. Along these lines, Kish and DiRuggiero (2008) have reported that Rad50 is not essential for Mre11-dependent repair of DNA DSBs in *Halobacteria*. After exposing bacteria deficient in either Mre11 or Rad50 (or both) to gamma radiation, the alkylating agent N-methyl-N'-nitro-N-nitroguanidine, and UV-C, it was discovered that cells lacking Mre11 showed a reduced rate of DNA repair, whereas those lacking Rad50 were not different to controls (Kish and DiRuggiero, 2008). The same is not thought to occur in eukaryotes, but the problem has not been fully explored, and many homologous proteins to RAD50 are encoded in the human genome; such as the nuclear-located A8K3I2 (or cDNA FLJ75532), which a quick bioinformatics search reveals it also harbours internal coil-coil domains and ATPase activity (The UniProt Consortium, 2015).

The links between RAD50 and cancer predisposition (Gasch et al., 2016; Hood and Rowen, 2013) were established from a study of Nordic populations. A truncation mutation in RAD50, 687delT (originating in Finland), has been found to predispose to breast cancer (Heikkinen et al., 2006). Additional predisposing variation in this clinical context has also been described in other patients, including the IVS3-I G>A mutation that interferes with RAD50 splicing.

It has been reported that a lack of RAD50 (and NBS1 and MRE11) typified epithelial ovarian carcinoma, where the authors relied on histology to describe their findings (Brandt et al., 2017), adding to similar findings in cystadenomas (Ali-Fehmi et al., 2010). Perhaps in this instance, RAD50 function was impaired by the cancer, and post-translational mechanisms must also be considered when describing the mechanisms of oncogenesis (Dwek et al., 2001). Furthermore, neither cancer heterogeneity, sample or disease collection variables (Baker, 2016), or the current genomic resolution based on limited numbers of patients (Devine and Smith, 1998; Spencer et al., 2009), can be overlooked as limitations in such kinds of studies.

The case of RAD50 (a gene that one would assume would have deleterious cancer-predisposing mutations associated with it) demonstrates the importance of individual germ-line and tumour DNA/RNA analysis to determine how individual cancers develop and progress (Auslander et al., 2016). The same way that the majority of cases suffering from an outbreak of infectious disease might have sub-clinical presentations, only determining the precise base pair sequence of the host and tumour will allow for the determination of cancer causality conclusively. The human genetics field has long been troubled by the problem of missing heritability (Manolio et al., 2009), and the search for low-frequency deleterious variants that predispose to complex polygenic diseases is on-going – with mixed success thus far (Chubb et al., 2016; Hunt et al., 2013; Nejentsev et al., 2009).

6.2 Hypothesis and study overview

The aim of the experiment described in this chapter is to determine the relationship between RAD50 expression in rectal tumour tissue and clinicopathological features including survival in a cohort of Australian patients. Although relatively fewer deleterious germ-line or somatic variants in RAD50 have been associated with cancer (compared to ATM, MRE and NBS1), the critical role of this protein in powering DSB repair makes it possible that its expression levels and activity could be indicative of disease state and treatment response.

The primary interest is to determine whether RAD50 expression in tumours can be used as a corollary marker for radiation sensitivity in rectal cancer patients.

Furthermore, expression of all three MRN complex proteins – MRE11, RAD50 and NBS1 – as a combinatorial biomarker panel is investigated for the same purposes; hypothesising that additional metrics of the DNA DSB response will increase the predictive power of any marker alone, especially when different genomic lesions in different patients can influence a single component of the pathway.

* The results presented in this chapter have recentlybeen published. Citation: **Ho, V**., Chung, L., Singh, A., Lea, V., Revoltar, M., Lim, SH., Tut, TG., Ng, W., Lee, M., de Souza, P., Shin, J and Lee, CS. (2017). Early postoperative low expression of RAD50 in rectal cancer patients associates with disease-free survival. Cancers *9*, 16.

6.3 Results

6.3.1 Patient Populations

A total of 266 patients were included in this study, similarly to what has been described in previous Chapters. In total, 176 (66.2%) patients were male, whilst 90 (33.8%) were female. The median age of participants whose samples were used in the study was 72 years (range: 35–100 years) (Table 6.1).

Of the 246 patients involved, 77 (31.3%) were treated with radiotherapy, and from these, 55 (71.4%) received preoperative therapy. During the time of sample collection, patients were followed for a median period of 3.16 years (range: 0−12.6 years), and the median time to death in the cohort was found to be 2.5 years after surgery (ranging between 0−11.1 years).

Table 6-1 Patient characteristics for the RAD50 study

	All Patients (%)
Total, *n*	266
Age median	72
Sex	
Male	176/266 (66.2)
Female	90/266 (33.8)
Tumour stage	
T1–2	88/260 (33.8)
T3–4	172/260 (66.2)
Node stage	
N0	140/259 (54.1)
N1–2	119/259 (45.9)
Metastasis stage	
M0	223/240 (92.9)
M1	17/240 (7.1)
Histological Grade	
1–2	246/266 (92.5)
3	20/266 (7.5)
Vascular invasion	
Absent	201/263 (76.4)
Present	62/263 (23.6)
Perineural invasion	
Absent	220/263 (83.7)
Present	43/263 (16.3)
Radiotherapy	
Total	77/246 (31.3)
Neoadjuvant	55/77 (71.4)
Adjuvant	22/77 (28.6)

6.3.2 Association between RAD50 expression and clinicohistopathological features and prognosis

The associations between postoperative RAD50 expression and the clinicohistopathological characteristics available to use for this study are presented in Table 6.2.

Table 6-2 Associations between RAD50 protein expression in the tumour centre and tumour periphery and clinicohistopathological data

		Tumour center			**Tumour periphery**		
		Low (%)	High (%)	*P* value	Low (%)	High (%)	*P* value
Sex	Male	26.3	73.7	0.98	29.6	70.4	0.42
	Female	26.1	73.9		34.4	65.6	
Age	≤72	28.8	71.2	0.39	30.6	69.4	0.82
	>72	24.1	75.9		31.9	68.1	
Tumour stage	T1–2	23.5	76.5	0.65	32.1	67.9	0.68
	T3–4	26.2	73.8		29.6	70.4	
Node stage	Negative	26.5	73.5	0.67	31.1	68.9	0.95
	Positive	24.1	75.9		30.8	69.2	
Metastasis stage	M0	25.5	74.5	0.72	31.5	68.5	0.09
	M1	29.4	70.6		11.8	88.2	
Histological Grade	1–2	26.1	73.9	0.88	31.3	68.7	0.98
	3	27.8	72.2		31.6	68.4	
Vascular invasion	Absent	26.3	73.7	0.74	31.9	68.1	0.50
	Present	24.2	75.8		27.4	72.6	
Perineural invasion	Absent	24.7	75.3	0.40	32.2	67.8	0.28
	Present	31.1	68.9		23.8	76.2	
Adjuvant therapy	No	25.2	74.8	0.83	24.8	75.2	**0.04**
	Yes	26.6	73.4		38.8	61.2	
Neoadjuvant therapy	No	23.8	76.2	0.14	27.1	72.9	0.16
	Yes	34.1	65.9		37.3	62.7	
Tumour regression grade	0–1	50	50	0.1**4**	50	50	0.22
	2-3	28.3	71.7		32.8	67.2	
MSH6	Negative	50.0	50.00	0.43	0	100	0.35
	Positive	25.5	74.5		30.8	69.2	
PMS2	Negative	55.6	44.4	**0.04**	44.4	55.6	0.36

	Positive	25.5	74.5		30.1	69.9	

Low RAD50 protein expression levels are found in the tumour periphery (TP) to be significantly associated with adjuvant therapy treatment in the cohort ($P = 0.04$). No significant differences were observed for age, sex, histological tumour stage, lymph node involvement, metastasis, vascular invasion, or perineural invasion, between patients with low or high RAD50 expression.

Kaplan–Meier survival analysis demonstrated that low RAD50 expression levels in the tumour centre (TC) were significantly associated with worse DFS ($P = 0.016$; Figure 6.2), whereas the association between RAD50 expression and OS was of borderline significance ($P = 0.056$; Figure 6.2B). No significant difference in survival was seen between patients with high or low RAD50 expression in the TP (DFS, $P = 0.295$; OS, $P = 0.695$) (Figures 6.2C and 6.2D). These results suggest that RAD50 expression serves to limit the progression of cancer, as reduced levels are associated with worse patient outcomes, although the effect was not observed in the tumour periphery. Representative immunohistochemical staining of high and low RAD50 expression in rectal cancer tissues is shown in Figures 6.3 A and 6.3 B.

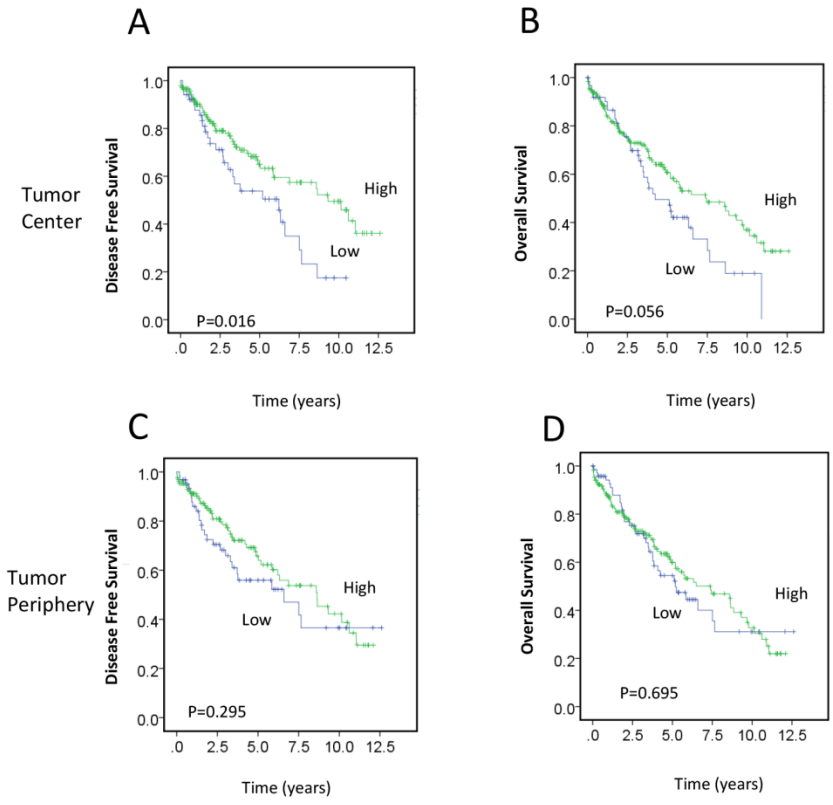


Figure 6-2 Association between postoperative RAD50 expression in the TC and TP and survival

(A-D) Kaplan–Meier survival analysis illustrating DFS (A, C) and OS (B, D) of patients with RAD50 expression in the TC (A, B) and TP (C, D). Blue lines represent patients with low RAD50 expression and green lines represent patients with high RAD50 expression.

The status of the MMR pathway in patient samples was investigated by evaluating the association of MMR proteins with RAD50 expression. All cases were found to be positive for MLH1 and MSH2 expression, and therefore, none of the cases were classified as MSI-high (MMR-negative). Furthermore, expression of MSH6 and PMS2 was found to be negative in 2/253 (0.8%) and 9/253 (3.6%) cases, respectively. Additionally, no significant associations between RAD50 and MSH6 expression, in either the TC or TP, were found. Finally, low post-operative RAD50 expression was significantly associated with PMS2 expression in the TC ($P = 0.04$), but not in the TP ($P = 0.36$) (Table 6.2). The MRE11-RAD50-NBS1 (MRN) complex is known to interact with the mismatch repair system and indeed is dependent on mismatch repair activity for its formation (Mirzoeva et al, 2006). This lends support to an intriguing interaction between RAD50 and PMS2 that warrants further investigation.

Low expression levels of RAD50 were found in the TC (HR = 0.552, 95% CI 0.339–0.899, $P = 0.017$; by univariate Cox regression) was significantly associated with reduced DFS in patients (Table 6.3). Additionally, using multivariate Cox analysis, we found that RAD50 expression (HR = 0.567, 95% CI 0.345–0.931, $P = 0.025$) and perineural invasion (HR = 2.364, 95% CI 1.343–4.162, $P = 0.003$) remained significantly associated with DFS (Table 6.3). However, in multivariate Cox analysis relating to OS, perineural invasion (HR = 1.701, 95% CI 1.036–2.792, $P = 0.036$) remained significantly associated with OS, but not with tumour RAD50 expression (HR = 0.712, 95% CI 0.462–1.095, $P = 0.122$) (Table 6.3).

Finally, by means of a Kaplan–Meier analysis of DFS to compare RAD50 low and high expression groups with tumours classified according to their perineural invasion status, low levels of RAD50 expression in rectal cancer tissues were significantly associated with perineural invasion (Figure 6.3C and 6.3D). These results indicate that low RAD50 in the context of PNI is a marker of poor prognosis, as shown above for DFS and OS.

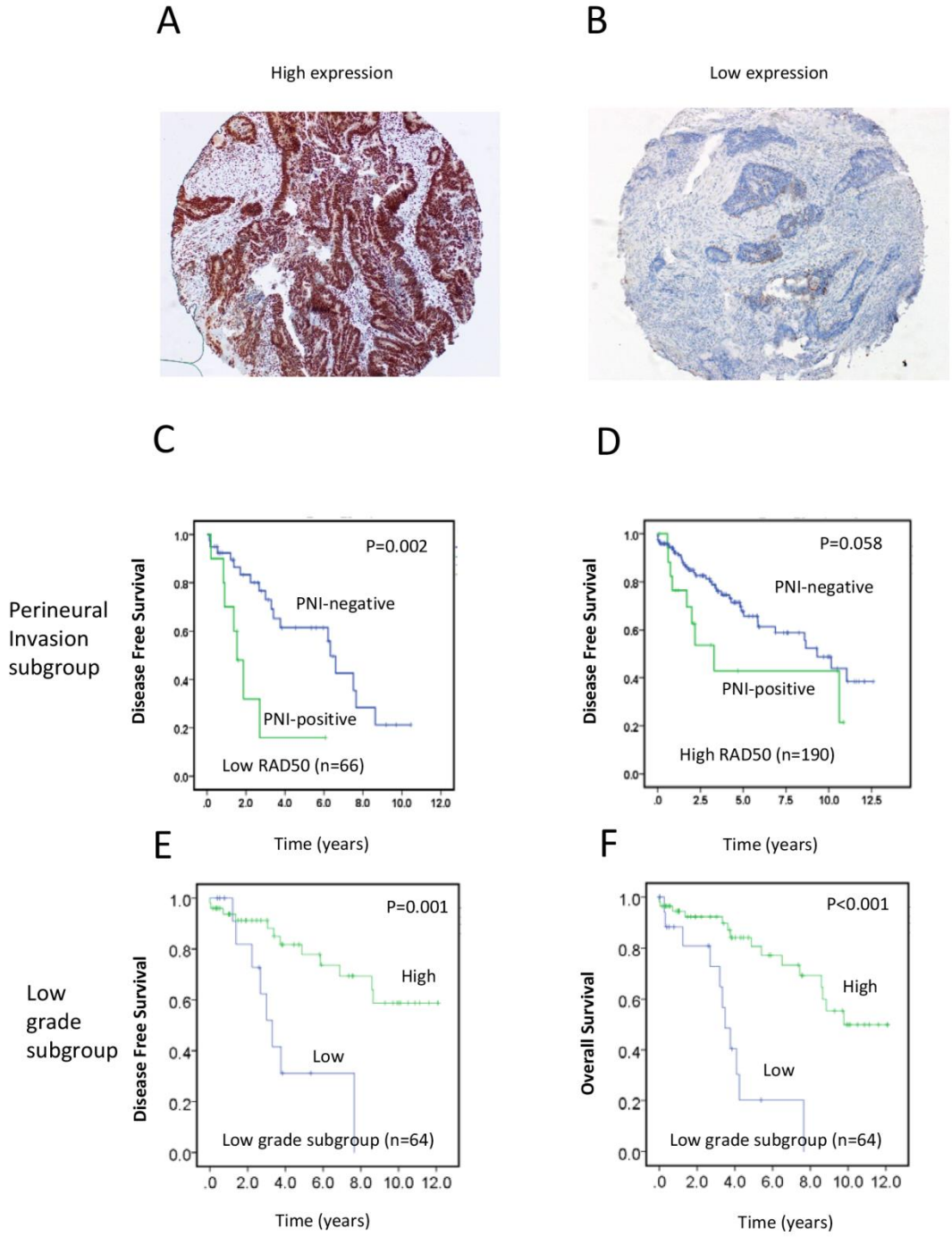


Figure 6-3 Correlation between RAD50 expression and perineural invasion and survival in early stage rectal cancers

(Panels A, B) Images showing representative immunohistochemical staining of RAD50 in rectal cancer samples (high versus low expression shown).
(Panels C, D) Kaplan–Meier survival analysis of DFS in the low RAD50 expression group (C) and the high RAD50 expression (D) group, with (green line) or without (blue line) perineural invasion.
(Panels E, F) Kaplan–Meier survival analysis illustrating the relationship of RAD50 expression with DFS (E) and OS (F) in low-grade (G1–2) with the early tumour stage (T1-2) subgroup. The analyses were divided into four subgroups including the low-grade with early stage tumours (G1-2, T1-2, n = 64), high-grade with early stage tumours (G3, T1-2, n = 4), low-grade with late stage tumours (G1-2, T3-4, n = 113) and high-grade with late stage tumours (G3, T3-4, n = 8).

Table 6-3 Cox regression analyses of postoperative RAD50 with overall survival.

Variables	Univariate			Multivariate		
	HR	**95% CI**	***p* Value**	**HR**	**95%**	***p* Value**
RAD50						
Low	0.674	0.448–1.012	0.047	0.712	0.462–1.095	0.122
High						
Age						
≤72	1.334	0.897–1.984	0.155			
>72						
Sex						
Male	1.092	0.734–1.625	0.664			
Female						
Tumor stage						
T1–2	1.796	1.158–2.786	0.001	1.382	0.858–2.226	0.183
T3–4						
Node stage						
Negative	1.454	0.987–2.140	0.058			
Positive						
Grade						
1–2	1.561	0.836–2.916	0.162			
3						
Vascular invasion						
Absent	2.03	1.340–3.015	0.001	1.365	0.848–2.196	0.200
Present						
Perineural invasion						
Absent	2.48	1.594–3.859	0.000	1.701	1.036–2.792	0.036
Present						
Adjuvant therapy						
No	0.506	0.301–0.850	0.09			
Yes						
Neoadjuvant therapy						
No	0.891	0.550–0.427	0.63			
Yes						

HR, hazard ratio; CI, confidence interval; RAD50: DNA repair protein RAD50 homolog.

Table 6-4 Cox regression analyses of post-operative RAD50 with disease-free survival

	Univariate			**Multivariate**		
	HR	95% CI	*P*-value	HR	95%	*P*-value
RAD50						
Low High	0.552	0.339–0.899	**0.017**	0.567	0.345–0.931	**0.025**
Age ≤72 >72	1.279	0.785-2.086	0.323			
Sex Male Female	1.074	0.657-1.754	0.776			
Tumour stage						
T1–2	1.643	0.983–2.747	0.058			
T3–4						
Node stage						
Negative	1.198	0.709–1.827	0.593			
Positive						
Histological Grade						
1–2	1.646	0.712–3.804	0.244			
3						
Vascular invasion						
Absent	1.888	0.650–0.2171	0.575			
Present						
Perineural invasion						
Absent	2.534	0.373–1.134	**0.001**	2.364	1.343–4.162	**0.003**
Present						
Adjuvant therapy						
No	0.65	0.373–1.134	0.129			
Yes						
Neoadjuvant therapy						
No	1.147	0.672–0.957	0.616			
Yes						
T1–2, G1–2,† RAD50				0.218	0.084-0.570	**0.002**

T3–4, G3, † RAD50				0.401	0.065-2.471	0.324
HR, hazard ratio; CI, confidence interval; TC, tumour centre, †denotes interaction						

6.3.3 RAD50 Expression as a Putative Prognostic Factor for Early Stage Rectal Cancer

The DFS of rectal cancer patients showing a low level of RAD50 expression was found to be significantly worse than for patients exhibiting high levels of RAD50 expression.
When patients were grouped into early tumour stage (T1–2) and low-grade (G1–2) subgroups, a low expression of RAD50 was found to be associated with decreased DFS ($P = 0.001$) (Figure 6.3E), indicating that RAD50 expression may be a useful prognostic biomarker for the early tumour stage and low-grade subgroups.

Similarly, low RAD50 expression in early tumour stage and low-grade tumour subgroups was significantly associated with worse OS ($P < 0.001$) (Figure 6.3F). Additional Cox regression analyses confirmed that expression of RAD50 in early tumour stage and low-grade subgroups significantly correlated with DFS (HR = 0.218, 95% CI 0.084–0.570, $P = 0.002$) (Table 6.4).

6.3.4 Results of the combinatorial MRE11-NSB1-RAD50 biomarker panel

6.3.4.1 Patient populations

A total of 265 patients were included for this part of the study, with patient characteristics being listed in Table 6.5. In total, one hundred and seventy-six (66.4%) volunteers were male, whilst 89 (33.6%) were female. The median age of the cohort was 71 years (ranging between 35–100 years). 77 out of 246 patients (31.3%) were treated with radiotherapy, and 55 of these (71.4%) also received preoperative therapy. Patients were followed for a median period of 3.16 years (ranging between 0−12.6 years), and the median time to death was 2.5 years after surgery (ranging between 0−11.1 years).

Table 6-5 MRN combinatorial panel patient characteristics

	All Patients (%)	Preoperative Radiotherapy Group
Total, n	265	55
Age median	71	66
Gender		
Male	176 (66.4)	37 (67.3)
Female	89 ((33.6)	18 (32.7)
Tumour stage		
T1-2	87/260 (33.4)	17/55 (30.9)
T3-4	173/260 (66.6)	38/55 (69.1)
Node stage		
N0	140/259 (54.1)	29/55 (52.7)
N1-2	119/259 (45.9)	26/55 (47.3)
Metastasis stage		
M0	223/240 (92.9)	53/54 (98.1)
M1	17/240 (7.1)	1/54 (1.9)
Grade		
1-2	245/265 (92.5)	51/55 (92.7)
3	20/265 (7.5)	4/55 (7.3)
Vascular invasion		
Absent	201/263 (76.4)	47/55 (85.5)
Present	62/263 (23.6)	8/55 (14.5)
Perineural invasion		
Absent	220/263 (83.7)	41/55 (74.5)
Present	43/263 (16.3)	14/55 (25.5)
Radiotherapy		
Total	77/246 (31.3)	-
Neoadjuvant	55/77 (71.4)	-
Adjuvant	22/77 (28.6)	0/55 (0)
Recurrence		
Absent	131/213 (61.5)	25/46 (54.3)
Present	82/213 (28.5)	21/46 (45.7)
Tumour regression grade		
0-2 (good response)	-	9/55 (16.4)
3 (poor response)	-	46/55 (83.6)

6.3.4.2 Establishment of a putative biomarker panel of the MRN complex

In order to determine the utility of this combinatorial MRN biomarker in rectal cancer, MRE11, RAD50 and NBS1 protein expression levels in the TC were tested in a forward and reverse binary logistic regression analysis; using a dataset of immunohistochemical scoring derived from 262 tumour samples and 258 normal tissue samples.

The final biomarker model gave an average receiver operator characteristic area under the curve (ROC-AUC) value of 0.870 when the three proteins of the MRN complex were combined. Similarly, MRE11, RAD50 and NBS1 protein expression levels in the TP (tumour, n = 261; normal, n = 258) were also evaluated, with the model giving an average ROC-AUC value of 0.862.

The sensitivity and specificity values for the MRN combinatorial panel were 89.0% and 77.2% for TC, and 78.2% and 77.6% for TP, respectively (see Table 6.5).

Table 6-6 Performance of the MRN three proteins combined classification models

Model	Tumour	Normal	Sensitivity (%)	Specificity (%)	Overall (%)	ROC-AUC
Combined TC*	262	258	89.0	77.2	83.1	0.870
Combined TP†	261	258	78.2	77.6	77.9	0.862
Tumour centre; †Tumour periphery						

6.3.4.3 Association between the MRN combined expression and clinicohistopathological features

The association between the three-protein combined expression levels and the clinicohistopathological characteristics is summarised in Table 6.7.

Table 6-7 Associations between the MRN combined expression in the tumour center and tumour periphery and clinicohistopathological data

		Tumour Center			Tumour Periphery		
		Low (%)	High (%)	P value	Low (%)	High (%)	P value
Sex	Male	62.5	66.8	0.567	80.0	64.6	0.121
	Female	37.5	33.2		20.0	35.4	
Age	≤70	43.8	46.5	0.751	36.0	47.1	0.290
	>70	56.2	53.5		64.0	52.9	
Tumour stage	T1–2	50.0	30.0	**0.009**	44.0	32.5	0.246
	T3–4	50.0	70.0		56.0	67.5	
Node stage	Negative	60.0	53.1	0.395	52.0	54.5	0.811
	Positive	40.0	46.9		48.0	45.5	
Metastasis stage	M0	97.8	91.7	0.147	100	92.1	0.153
	M1	2.2	8.3		0	7.9	
Grade	1–2	93.8	92.2	0.707	88.0	92.9	0.376
	3	6.2	7.8		12.0	7.1	
Vascular invasion	Absent	82.6	75	0.27	84	75.5	0.343
	Present	17.4	25		16	24.5	
Perineural invasion	Absent	84.8	83.8	0.869	92	83.1	0.250
	Present	15.2	16.2		8	16.9	
Adjuvant therapy	No	70.0	69.4	0.945	52.4	71.4	0.072
	Yes	30.0	30.6		47.6	28.6	
Neoadjuvant therapy	No	65.2	80.9	**0.021**	68	78	0.205
	Yes	34.8	19.1		32	22	
Tumour regression grade	0–2	31.6	11.3	**0.042**	30	14.5	0.223
	3	68.4	88.7		70	85.5	
MSH6	Negative	0	100	0.518	0	100	0.663
	Positive	17.3	82.7		8.7	91.3	
PMS2	Negative	11.1	88.9	0.669	0	100	0.345
	Positive	16.5	83.5		9.1	90.9	

High expression levels of the MRN combinatorial panel were found in the TC and were significantly associated with the histological tumour stage (P = 0.009), and with a TRG of 0 to 2 (responders) (P = 0.042), indicating that MRN-mediated DNA damage repair may mediate an important DNA damage response. No significant differences were associated with patient age, sex, lymph node involvement, metastasis, vascular invasion, or perineural invasion and low and high MRN complex protein expression.

Subsequent Kaplan–Meier survival analyses demonstrated that a high score for the combinatorial MRN complex expression in the TC was significantly associated with a worse DFS in patients (P = 0.021; Figure 6.4A), and OS (P = 0.002; Figure 6.4B). No significant differences in survival were seen between patients with high or low MRN complex protein expression levels in the TP (DFS, P = 0.646; OS, P = 0.251; Figures 6.4C and 6.4D, respectively).

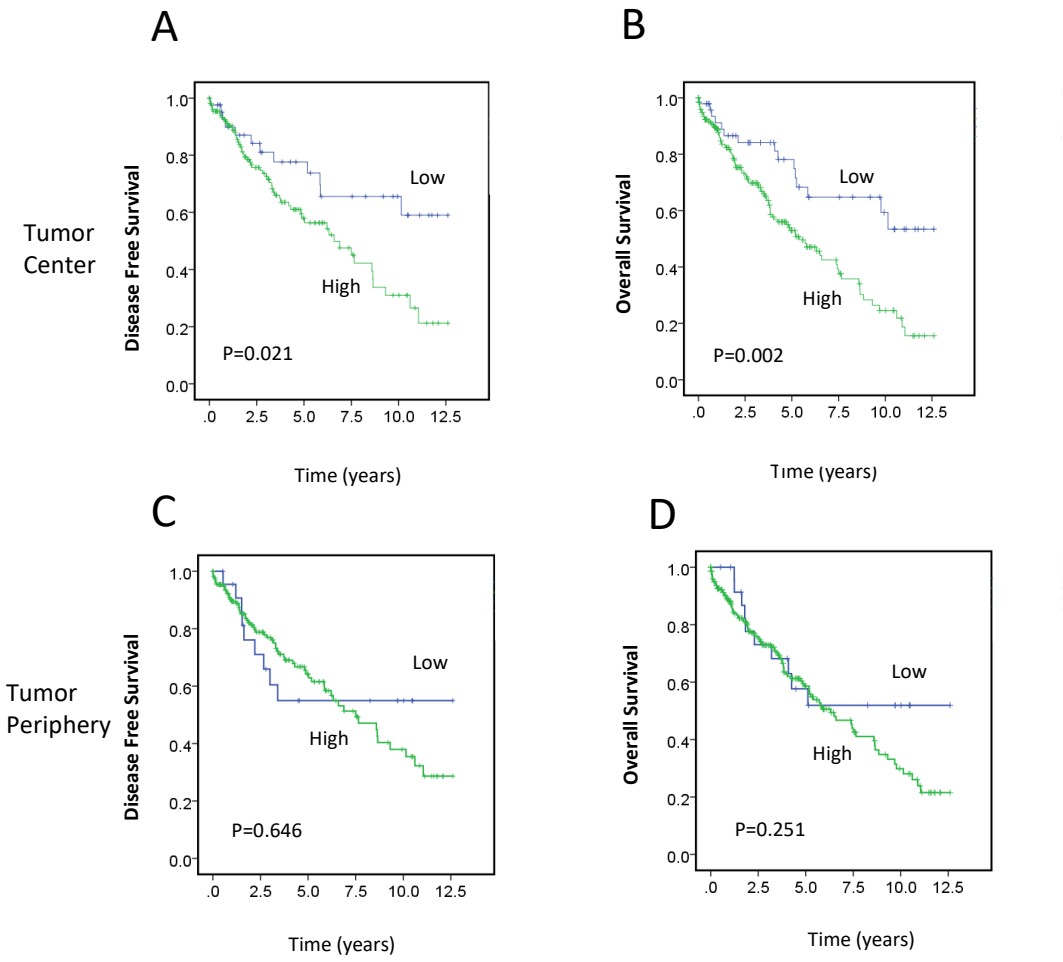


Figure 6-4 Association between MRN complex proteins expression in the TC and TP and survival

(A-D) Kaplan–Meier survival analysis illustrating DFS (A, C) and OS (B, D) of patients with the MRN combined expression levels in the TC (A, B) and TP (C, D). Blue lines represent patients with low MRN combined expression and green lines represent patients with high MRN combined expression.

Using univariate Cox regression analysis, we found that high expression of the combined three protein panel in the TC (HR = 2.069, 95% CI 1.102–3.882, *P* = 0.024) was significantly associated with reduced DFS (Table 6.8). Additionally, multivariate Cox analysis (adjusted for the combined three-protein expression of MRN complex and perineural invasion) demonstrated that MRN complex expression (HR = 2.114, 95% CI 1.096–4.078, *P* = 0.026) and perineural invasion (HR = 2.16, 95% CI 1.209–3.859, *P* = 0.009) remained significantly associated with DFS (Table 6.8), implying that those markers together are strongly prognostic for disease-free survival in rectal cancer patients.

Table 6-8 Cox regression analyses of MRN combined TC expression with clinicohistopathological variables

		Univariate			Multivariate		
	n (%)	HR	95% CI	P Value	HR	95%	P Value
MRN combined TC*							
High	81.9	2.069	1.102-3.882	**0.024**	2.114	1.096-4.078	**0.026**
Low	18.1						
Tumour stage							
T1–2	33.6	1.501	0.897-2.512	0.122			
T3–4	66.4						
Node stage							
Negative	54.3	1.44	0.976-2.126	0.066			
Positive	45.7						
Grade							
1–2	92.5	1.537	0.823-2.872	0.178			
3	7.5						
Vascular invasion							
Absent	76.3	1.167	0.638-2.134	0.617			
Present	23.7						
Perineural invasion							
Absent	84.0	2.334	1.310-4.157	**0.004**	2.16	1.209-3.859	**0.009**
Present	16.0						
Adjuvant therapy							

No	69.5	0.602	0.341-1.063	0.08			
Yes	30.5						
Neoadjuvant therapy							
No	78.0	0.855	0.529-1.381	0.521			
Yes	22.0						
LN-negative† combined TC	54.3				1.339	0.589-3.042	0.486
LN-positive† combined TC	45.7				3.472	1.051-11.454	**0.047**
Three marker combined expression in the tumour center; † denotes interaction							
HR, hazard ratio; CI, confidence interval; TC, tumour center; LN, lymph node							

6.3.4.4 Association between MRN combinatorial panel and MMR protein expression

All cases were positive for MLH1 and MSH2 expression in this cohort, and therefore, none of the cases were classified as MSI-high (MMR-negative).

The expression of MSH6 and PMS2 was negative in 2/256 (0.8%) and 9/252 (3.6%) cases, respectively, and no significant associations between combined MRN protein expression and MSH6 or PMS2 expression, in either TC or TP samples, were found (Table 6.7).

6.3.4.5 Correlation of the MRN combined expression with neoadjuvant radiotherapy

Disease-free and overall survival outcomes were analysed in the 55 patients that had previously received pre=operative radiotherapy. Of these 55 patients, thirty-seven (67.3%) of them were male, and 18 (33.6%) were female (Table 6.5). The DFS estimates in the subgroup receiving neo-adjuvant radiotherapy are shown in Figure 6.6A, demonstrating that a higher combined expression level of the MRN complex is significantly associated with worse DFS in this cohort (P = 0.024).

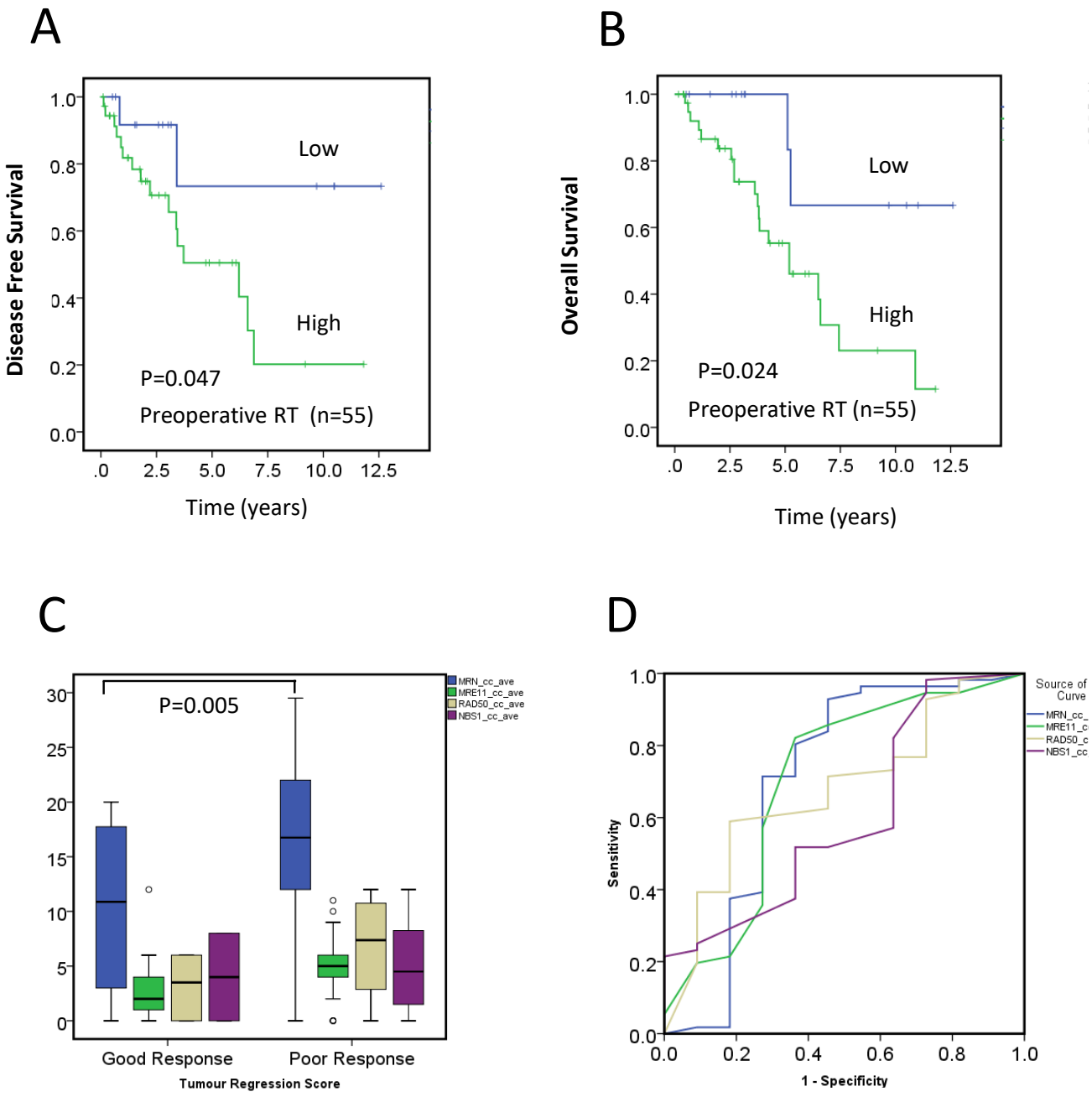


Figure 6-5 Relationship between preoperative MRN in rectal cancer tissues and survival and MRN combined expression in relation to TRG

(A, B) Kaplan–Meier survival analysis illustrating DFS (A) and OS (B) in preoperative radiotherapy patient groups with low (blue line) and high (green line) MRN complex panel expression.
(C) Box plot shows levels of MRE11 (green), RAD50 (yellow), NBS1 (purple), and their combined expression (blue) in the TC categorized by TRG as 0-2 (good response) or 3 (poor response). The association of between protein expression with TRG was examined by Mann-Whitney U test.
(D) Receiver operating characteristics (ROC) curve analysis comparing the performance of MRE11, RAD50 and NBS1 alone with the MRN combined three-protein panel.

Similarly, when considering the overall survival of patients receiving radiotherapy, overexpression of MRN complex proteins was also associated with significantly worse OS than that of patients with lower expression (Figure 6.4B, $P = 0.028$). These results suggest that the MRN three-protein combined panel has potential as a predictive marker of the tumour response to radiotherapy.

As expected, multivariate analyses in patients that had received preoperative radiotherapy (see Table 6.9) revealed that a higher histological grade was strongly correlated with decreased overall survival (HR = 7.275, 95% CI 1.842−28.730, $P = 0.005$). High expression of the three MRN complex proteins was also significantly associated with worse OS (HR = 4.196, 95% CI 0.968−18.191, $P = 0.045$).

Table 6-9 Multivariate analysis of MRN combined expression with overall survival in patients who received preoperative radiotherapy

	Multivariate		
	HR	95%	P Value
MRN combined TC expression	4.196	0.968-18.191	0.045
Grade	7.275	1.842-28.730	0.005
Sex	3.017	1.199-7.592	0.019

6.3.4.6 MRN complex proteins expression in relation to TRG

TRG provides a valuable tool to assist in clinical oncology decision making. In the TRG subgroup study, we used a univariate analysis by the Mann−Whitney U test to explore the relationship between TRG and expression of MRN complex proteins.

A significant association was found between increasing TRG and the expression level of the combinatorial MRN panel in the TC ($P = 0.005$, Figure 6.5C), and MRE11 and RAD50 proteins ($P = 0.046$ and $P = 0.01$, respectively), but we did not find there to be a significant association between TRG scores and NBS1 expression.

Using a ROC-AUC analysis of *good* versus *poor* histological tumour response among patients treated preoperatively with radiotherapy, the average ROC-AUC was 0.725 for the combined panel, 0.711 for MRE11, 0.677 for RAD50 and 0.602 for NBS1 alone, respectively (Figure 6.5D).

Together, these results strongly suggest that the combined three-protein biomarker panel may provide better levels of discrimination between *good* and *poor* tumour responses after radiotherapy, than either marker alone.

6.3.4.7 Prognostic implications of MRN complex proteins in LN-positive subgroup

The DFS of rectal cancer patients showing overexpression of the combinatorial MRN panel was significantly worse than that of patients found to have lower expression.

When patients were grouped according to the status of LN involvement, high expression of the MRN complex proteins was associated with decreased DFS, and worse OS in patients with LN-positive tumours ($P = 0.029$ for DFS, Figure 6.6B; $P = 0.020$ for OS; Figure 6.6D), but not in those with LN-negative tumours ($P = 0.485$ for DFS, Figure 6.6A; $P = 0.073$ for OS; Figure 6.6C).

By multivariate Cox regression analysis, expression of the MRN combined panel in TC in the LN positive subgroup was significantly correlated with DFS (HR = 3.474, 95% CI 1.054–11.451, $P = 0.041$) (Table 6.8).

These results suggest that the three-protein combined expression of the MRN complex may be associated with lymph node involvement in relation to patient survival.

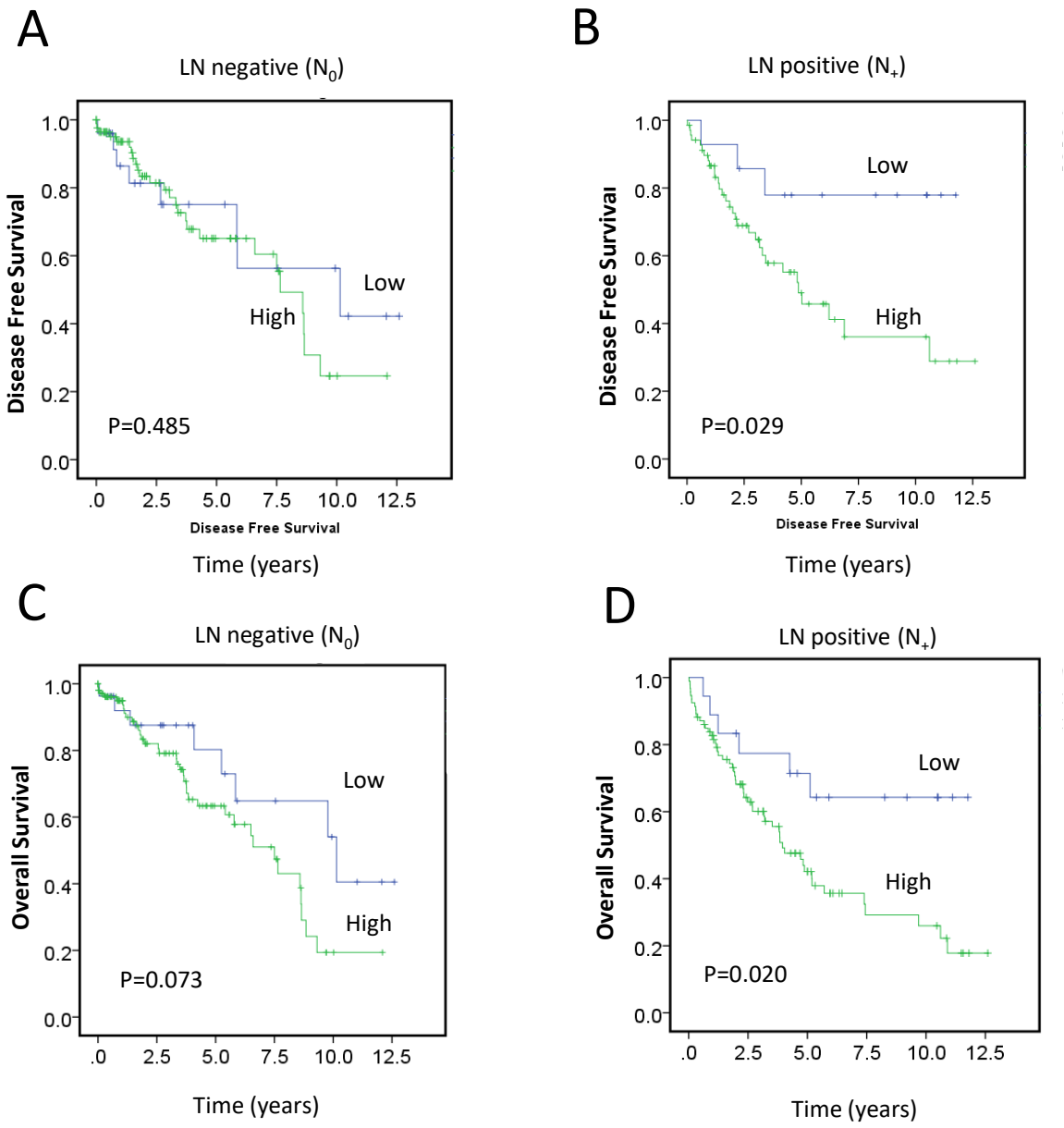


Figure 6-6 Kaplan-Meier survival analysis of MRN combined expression according to lymph node involvement

(A-D) respectively show survival curves of high (green line) and low (blue line) MRN combined expression groups in lymph node (LN) negative and LN positive rectal cancers. These show the effect of LN status on the association between expression levels and DFS (A, B) or OS (C, D).

6.4 References

Ali-Fehmi, R., Chatterjee, M., Ionan, A., Levin, N.K., Arabi, H., Bandyopadhyay, S., Shah, J.P., Bryant, C.S., Hewitt, S.M., O'Rand, M.G., et al. (2010). Analysis of the expression of human tumor antigens in ovarian cancer tissues. Cancer Biomark. *6*, 33–48.

Anderson, D.E., Trujillo, K.M., Sung, P., and Erickson, H.P. (2001). Structure of the Rad50 x Mre11 DNA repair complex from Saccharomyces cerevisiae by electron microscopy. J. Biol. Chem. *276*, 37027–37033.

Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P., Dolinski, K., Dwight, S.S., Eppig, J.T., et al. (2000). Gene ontology: Tool for the unification of biology. Nat. Genet. *25*, 25–29.

Auslander, N., Yizhak, K., Weinstock, A., Budhu, A., Tang, W., Wang, X.W., Ambs, S., and Ruppin, E. (2016). A joint analysis of transcriptomic and metabolomic data uncovers enhanced enzyme-metabolite coupling in breast cancer. Sci. Rep. *6*, 29662.

Baker, M. (2016). Reproducibility: Respect your cells! Nature *537*, 433–435.

Ball, A.R., and Yokomori, K. (2001). The structural maintenance of chromosomes (SMC) family of proteins in mammals. Chromosom. Res. *9*, 85–96.

Bhaskara, V., Dupré, A., Lengsfeld, B., Hopkins, B.B., Chan, A., Lee, J.H., Zhang, X., Gautier, J., Zakian, V., and Paull, T.T. (2007). Rad50 Adenylate Kinase Activity Regulates DNA Tethering by Mre11/Rad50 Complexes. Mol. Cell *25*, 647–661.

Brandt, S., Samartzis, E.P., Zimmermann, A.-K., Fink, D., Moch, H., Noske, A., and Dedes, K.J. (2017). Lack of MRE11-RAD50-NBS1 (MRN) complex detection occurs frequently in low-grade epithelial ovarian cancer. BMC Cancer *17*, 44.

Chubb, D., Broderick, P., Dobbins, S.E., Frampton, M., Kinnersley, B., Penegar, S., Price, A., Ma, Y.P., Sherborne, A.L., Palles, C., et al. (2016). Rare disruptive mutations and their contribution to the heritable risk of colorectal cancer. Nat. Commun. *7*, 11883.

Connelly, J.C., and Leach, D.R.F. (2002). Tethering on the brink: The evolutionarily conserved Mre11-Rad50 complex. Trends Biochem. Sci. *27*, 410–418.

Davison, L.J., Wallace, C., Cooper, J.D., Cope, N.F., Wilson, N.K., Smyth, D.J., Howson, J.M.M., Saleh, N., Al-Jeffery, A., Angus, K.L., et al. (2012). Long-range DNA looping and gene expression analyses identify DEXI as an autoimmune disease candidate gene. Hum. Mol. Genet. *21*, 322–333.

Day, F.R., Hinds, D.A., Tung, J.Y., Stolk, L., Styrkarsdottir, U., Saxena, R., Bjonnes, A., Broer, L., Dunger, D.B., Halldorsson, B. V., et al. (2015). Causal mechanisms and balancing selection inferred from genetic associations with polycystic ovary syndrome. Nat. Commun. *6*, 8464.

De Jager, M., Trujillo, K.M., Sung, P., Hopfner, K.P., Carney, J.P., Tainer, J.A., Connelly, J.C., Leach, D.R.F., Kanaar, R., and Wyman, C. (2004). Differential arrangements of conserved building blocks among homologs of the Rad50/Mre11 DNA repair protein complex. J. Mol. Biol. *339*, 937–949.

Deshpande, R.A., Lee, J.H., and Paull, T.T. (2017). Rad50 ATPase activity is regulated by DNA ends and requires coordination of both active sites. Nucleic Acids Res. *45*, 5255–5268.

Devine, O.J., and Smith, J.M. (1998). Estimating sample size for epidemiologic studies: The impact of ignoring exposure measurement uncertainty. Stat. Med. *17*, 1375–1389.

Dwek, M. V, Ross, H.A., and Leathem, A.J. (2001). Proteome and glycosylation mapping identifies post-translational modifications associated with aggressive breast cancer. Proteomics *1*, 756–762.

Ecker, J.R., Bickmore, W.A., Barroso, I., Pritchard, J.K., Gilad, Y., and Segal, E. (2012). Genomics: ENCODE explained. Nature *489*, 52–55.

Gallego, M.E., Jeanneau, M., Granier, F., Bouchez, D., Bechtold, N., Charles, and White, I. (2001). Disruption of the Arabidopsis RAD50 gene leads to plant sterility and MMS sensitivity. Plant J. *25*, 31–41.
Gasch, A.P., Payseur, B.A., and Pool, J.E. (2016). The Power of Natural Variation for Model Organism Biology. Trends Genet. *32*, 147–154.
Handel, M.A., and Schimenti, J.C. (2010). Genetics of mammalian meiosis: Regulation, dynamics and impact on fertility. Nat. Rev. Genet. *11*, 124–136.
Heikkinen, K., Rapakko, K., Karppinen, S.-M., Erkko, H., Knuutila, S., Lundán, T., Mannermaa, A., Børresen-Dale, A.-L., Borg, A., Barkardottir, R.B., et al. (2006). RAD50 and NBS1 are breast cancer susceptibility genes associated with genomic instability. Carcinogenesis *27*, 1593–1599.
Ho, V., Chung, L., Singh, A., Lea, V., Revoltar, M., Lim, SH., Tut, TG., Ng, W., Lee, M., de Souza, P., Shin, J and Lee, CS. (2017). Early postoperative low expression of RAD50 in rectal cancer patients associates with disease-free survival. Cancers *9*, 163.
Hohl, M., Kwon, Y., Galván, S.M., Xue, X., Tous, C., Aguilera, A., Sung, P., and Petrini, J.H.J. (2010). The Rad50 coiled-coil domain is indispensable for MRE11 complex functions. Nat. Struct. Mol. Biol. *18*, 1124–1131.
Hood, L., and Rowen, L. (2013). The human genome project: big science transforms biology and medicine. Genome Med. *5*, 79.
Hopfner, K.P., Karcher, A., Shin, D.S., Craig, L., Arthur, L.M., Carney, J.P., and Tainer, J.A. (2000a). Structural biology of Rad50 ATPase: ATP-driven conformational control in DNA double-strand break repair and the ABC-ATPase superfamily. Cell *101*, 789–800.
Hopfner, K.P., Karcher, A., Shin, D., Fairley, C., Tainer, J.A., and Carney, J.P. (2000b). Mre11 and Rad50 from Pyrococcus furiosus: Cloning and biochemical characterization reveal an evolutionarily conserved multiprotein machine. J. Bacteriol. *182*, 6036–6041.
Hopfner, K.P., Craig, L., Moncalian, G., Zinkel, R.A., Usui, T., Owen, B.A.L., Karcher, A., Henderson, B., Bodmer, J.L., McMurray, C.T., et al. (2002). The Rad50 zinc-hook is a structure joining Mre11 complexes in DNA recombination and repair. Nature *418*, 562–566.
Hunt, K.A., Mistry, V., Bockett, N.A., Ahmad, T., Ban, M., Barker, J.N., Barrett, J.C., Blackburn, H., Brand, O., Burren, O., et al. (2013). Negligible impact of rare autoimmune-locus coding-region variants on missing heritability. Nature *498*, 232–235.
Kinoshita, E., Van Der Linden, E., Sanchez, H., and Wyman, C. (2009). RAD50, an SMC family member with multiple roles in DNA break repair: How does ATP affect function? Chromosom. Res. *17*, 277–288.
Kish, A., and DiRuggiero, J. (2008). Rad50 is not essential for the Mre11-dependent repair of DNA double-strand breaks in Halobacterium sp. strain NRC-1. J. Bacteriol. *190*, 5210–5216.
Luo, G., Yao, M.S., Bender, C.F., Mills, M., Bladl, A.R., Bradley, A., and Petrini, J.H. (1999). Disruption of mRad50 causes embryonic stem cell lethality, abnormal embryonic development, and sensitivity to ionizing radiation. Proc. Natl. Acad. Sci. U. S. A. *96*, 7376–7381.
Ma, L., Milman, N., Nambiar, M., and Smith, G.R. (2015). Two separable functions of Ctp1 in the early steps of meiotic DNA double-strand break repair. Nucleic Acids Res. *43*, 7349–7359.
MacArthur, J., Bowler, E., Cerezo, M., Gil, L., Hall, P., Hastings, E., Junkins, H., McMahon, A., Milano, A., Morales, J., et al. (2017). The new NHGRI-EBI Catalog of published genome-wide association studies (GWAS Catalog). Nucleic Acids Res. *45*, D896–D901.
Manolio, T.A., Collins, F.S., Cox, N.J., Goldstein, D.B., Hindorff, L.A., Hunter, D.J., McCarthy, M.I., Ramos, E.M., Cardon, L.R., Chakravarti, A., et al. (2009). Finding the missing heritability of complex diseases. Nature *461*, 747–753.
Mirzoeva O.K., Kawaguchi T., Pieper R.O. (2006). The Mre11/Rad50/Nbs1 complex interacts

with the mismatch repair system and contributes to temozolomide-induced G2 arrest and cytotoxicity. Mol. Cancer Ther. *5*, 2757–2766.
Nejentsev, S., Walker, N., Riches, D., Egholm, M., and Todd, J.A. (2009). Rare variants of IFIH1, a gene implicated in antiviral responses, protect against type 1 diabetes. Science *324*, 387–389.
Park, Y.B., Hohl, M., Padjasek, M., Jeong, E., Jin, K.S., Kręzel, A., Petrini, J.H.J., and Cho, Y. (2017). Eukaryotic Rad50 functions as a rod-shaped dimer. Nat. Struct. Mol. Biol. *24*, 248–257.
Qvist, P., Huertas, P., Jimeno, S., Nyegaard, M., Hassan, M.J., Jackson, S.P., and Børglum, A.D. (2011). CtIP mutations cause Seckel and Jawad syndromes. PLoS Genet. *7*, e1002310.
Ronceret, A., Doutriaux, M.-P., Golubovskaya, I.N., and Pawlowski, W.P. (2009). PHS1 regulates meiotic recombination and homologous chromosome pairing by controlling the transport of RAD50 to the nucleus. Proc. Natl. Acad. Sci. U. S. A. *106*, 20121–20126.
Roth, S., Rottach, A., Lotz-Havla, A.S., Laux, V., Muschaweckh, A., Gersting, S.W., Muntau, A.C., Hopfner, K.P., Jin, L., Vanness, K., et al. (2014). Rad50-CARD9 interactions link cytosolic DNA sensing to IL-1β production. Nat. Immunol. *15*, 538–545.
Speir, M.L., Zweig, A.S., Rosenbloom, K.R., Raney, B.J., Paten, B., Nejad, P., Lee, B.T., Learned, K., Karolchik, D., Hinrichs, A.S., et al. (2016). The UCSC Genome Browser database: 2016 update. Nucleic Acids Res. *44*, D717–D725.
Spencer, C.C.A., Su, Z., Donnelly, P., and Marchini, J. (2009). Designing genome-wide association studies: Sample size, power, imputation, and the choice of genotyping chip. PLoS Genet. *5*, e1000477.
Strunnikov, A. V. (1998). SMC proteins and chromosome structure. Trends Cell Biol. *8*, 454–459.
Trujillo, K.M., and Sung, P. (2001). DNA Structure-specific Nuclease Activities in the Saccharomyces cerevisiae Rad50-Mre11 Complex. J. Biol. Chem. *276*, 35458–35464.
UniProt Consortium (2015). UniProt: a hub for protein information. Nucleic Acids Res. *43*, D204-12.
Vannier, J.B., Depeiges, A., White, C., and Gallego, M.E. (2006). Two roles for Rad50 in telomere maintenance. EMBO J. *25*, 4577–4585.
Waltes, R., Kalb, R., Gatei, M., Kijas, A.W., Stumm, M., Sobeck, A., Wieland, B., Varon, R., Lerenthal, Y., Lavin, M.F., et al. (2009). Human RAD50 Deficiency in a Nijmegen Breakage Syndrome-like Disorder. Am. J. Hum. Genet. *84*, 605–616.
Wang, Y., Cortez, D., Yazdi, P., Neff, N., Elledge, S.J., and Qin, J. (2000). BASC, a super complex of BRCA1-associated proteins involved in the recognition and repair of aberrant DNA structures. Genes Dev. *14*, 927–939.
Weidinger, S., Gieger, C., Rodriguez, E., Baurecht, H., Mempel, M., Klopp, N., Gohlke, H., Wagenpfeil, S., Ollert, M., Ring, J., et al. (2008). Genome-wide scan on total serum IgE levels identifies FCER1A as novel susceptibility locus. PLoS Genet. *4*, e1000166.

CHAPTER 7
DISCUSSION

7.1 Introduction

To date, long-term clinical trials (Choudhury et al., 2010) and several meta-analyses (Camma et al., 2000; Colorectal Cancer Collaborative Group, 2001) have confirmed that neoadjuvant radiotherapy plus surgery, reduces the rate of local recurrence, and increases the survival rate of rectal cancer patients, where benefits far outweigh the associated side effects.

However, patients with rectal cancer can have variable responses to radiotherapy (Aschele et al., 2011), therefore, the identification of rectal cancer-specific radio-sensitivity and prognostic markers would enable targeted therapeutic decisions, improve treatment results, and survival rates (Compton et al., 2000). Preliminary laboratory work in finding potential biomarkers can help pave the way for randomised controlled trials which are the gold standard approach to validate predictive markers (Forker et al, 2015).

In the following sections major implications and considerations arising from the findings in this thesis will be outlined, covering each protein of interest individually, prior to reviewing the results arising from the two combinatorial panels.

7.2 Results Discussion

7.2.1 ATM

ATM is a pleiotropic kinase that mediates several pathways to apoptosis, and facilitates cell cycle checkpoint regulation; as well as activating p53, BRCA1 and other TSGs (please see Chapter Three; (Lee and Paull, 2007)). High ATM expression would, therefore, be expected to correlate with increased DNA damage burden and repair, hallmarks of cancer. Indeed, high expression of ATM, the cell cycle protein Ki67, and PRKDC, have been reported in diverse types of malignant tumours, and have been described as prognostic markers. For example, work by Abdel-Fatah and colleagues showed high ATM expression to be predictive of serous cystadenocarcinoma and resistance to platinum-based chemotherapy (Abdel-Fatah et al., 2014). Additionally, in a study of early stage hormone receptor positive breast cancer, high ATM expression predicted a favourable prognosis, suggesting cells are able to repair DNA damage (Feng et al., 2016). These studies add weight to the notion that ATM expression is an important and informative feature of diverse cancers; necessitating its deeper understanding across different clinical contexts.

A significantly higher ATM expression is found in cancer tissues compared with normal healthy mucosa. These observations suggest that, at the point of sample collection, the kinase has a more prominent role in cancer tissue compared to healthy tissue, the malignant DNA tissue has a higher mutation burden than healthy tissue, and that DNA repair is occurring to a greater degree at this site. Theoretically, a loss of ATM expression could facilitate cell division in the face of DNA damage, and may be more important during the early stages of disease, or

in different sections of the tumour.

This result (high ATM expression in tumours) does not agree with other studies that report a loss of ATM expression in colorectal carcinomas (Bai et al., 2004; Grabsch et al., 2006); although this could be due to tissue-specific (rectal versus colonic tumours show different mechanisms, as discussed) and patient cohort differences; including genetic variation (i.e. loss of ATM variants), age and sex distributions, disease stage, and concurrent chemotherapy. As discussed, differential scoring of the TRG could also impact variability between studies. Furthermore, as ATM has a number of diverse roles in the cell, different studies may be detecting the protein in relation to different functions, confusing the clinical picture. Indeed, ATM is predicted to have hundreds of protein targets in the cell (Shiloh and Ziv, 2013), and may contribute to, or restrain cancer, in a number of ways. Unfortunately, the only way to resolve such differences is through additional, more informed studies.

In contrast to the findings of this thesis, high levels of ATM expression in rectal tumours may represent activation of ATM in response to an array of insults, including ionising radiation, mutagenic chemicals, replicative errors, inherited variants, and metabolic by-products; especially reactive oxygen species (Jackson, 2002). Along these lines, it is important to note that although cancers acquire the genetic ability to elude regulatory control, many normal molecular mechanisms remain intact and will function to restrain cell growth under deleterious conditions.

ATM expression was, remarkably, found to be lower in samples taken from the periphery of tumours, compared with samples taken from more central tumour locations. This important distinction emphasises the heterogeneity within the tumour microenvironment (Quail and Joyce, 2013), where different cells carry out different roles (Buccione et al., 2009; Han et al., 2017), and nicely describe the protein's expression patterns within a single rectal tumour. How cancers regulate and co-ordinate the symbiotic development and specialisation of clonal cells within a tumour remains to be fully determined, although hormonal mechanisms and hijacking of adjacent stroma are known to facilitate tumour microenvironment development; by facilitating T-cell exclusion using chemokines, for example (Feig et al., 2013; Flint et al., 2016; Kraman et al., 2010). Indeed, hormonal paraneoplastic syndromes are well appreciated mechanisms tumours use to grow at the expense of the rest of the organism (Pelosof and Gerber, 2010).

It is evident from our ATM immunohistochemistry results that differences in protein expression levels, and their association with clinicohistopathological variables and survival outcomes, exist; especially when comparing staining from the centre versus the periphery of the tumour mass. Effects due to sampling site variation across tumours have been noted previously; with multi-site tumour sampling being effective at capturing a more representative picture of the cancer (Guarch et al., 2016). For example, in an evaluation of biomarkers in ovarian cancer, Permuth-Wey *et al* reported sampling variability in protein expression analysis using tissue microarrays (Permuth-Wey et al., 2009). The authors concluded that, for ovarian cancer at least, more reliable data could be obtained from the

tumour periphery than from the centre, perhaps due to the hypoxic nature of the tumour centre affecting *ex vivo* staining protocols. Alternatively, this could be attributed to optimal exposure to fixatives at the periphery of the processed tissues (Yamashita-Kashima et al., 2014), although the effect of tumour heterogeneity on sampling variation should also be considered. Across our studies, every effort was made to select sample cores from representative areas of pathology or healthy tissue in a blinded manner, although future studies could select a greater number of cores for a large number of sites within the same tumour.

Peripheral tumour cells are generally thought to facilitate invasiveness and metastasis (Quail and Joyce, 2013) – where we found ATM expression to be lowest within the tumour – potentially allowing for more cancerous growth in the absence of other repair mechanisms . Furthermore, rapid, un-restrained growth in peripheral tumour cells leads to vascular insufficiency and hypoxia in the central tumour microenvironment (Gatenby et al., 2007). This could drive the increase in ATM expression, which helps keep cancer cells quiescent in the centre of the tumour, and alive in low-oxygen conditions. Indeed, mTORC signalling under hypoxic conditions is controlled by ATM-dependent phosphorylation of HIF-1-alpha (Cam et al., 2010). Thus, the centrally-elevated levels of ATM could represent hypoxia due to rapid peripheral cell growth, which may indicate invasiveness. Indeed, a loss of ATM has been found to further the progression of pancreatic and breast cancer (Feng et al., 2015; Russell et al., 2015). It is unsurprising, therefore, that ATM has previously been described as an anti-cancer barrier in early tumorigenesis (Bartkova et al., 2005; Reddy et al., 2010), and higher expression in peripheral tumour cells would be predicted to curtail tumour growth (expansion and basement membrane penetration); whilst a dysfunctional DDR pathway would be predicted to allow the cell to evade the safeguards preventing un-restricted clonal expansion. Our results support this, namely that there is insufficient ATM function in the tumour periphery.

Indeed, negative/low peripheral ATM expression was found to be associated with higher grade tumours, and older patient age in this thesis. In older patients, increased ATM may be contributed by a higher mutation burden in these patients, compared to younger patients; along the lines of the DNA damage theory of ageing (Freitas and De Magalhães, 2011), from which it follows that more ATM is needed to repair cumulatively more lesions over time. As the cohort investigated in this thesis consisted mainly of elderly patients (with a mean age between 71 and 72), similar studies need to be carried out in samples taken from younger patients.

The fact that low peripheral ATM expression was associated with a higher grade of tumour further supports that ATM restrains oncogenesis and is defective in these samples. No correlations were found to be present for ATM staining in the tumour centre, highlighting the importance in studying the tumour as a whole. Here, limited sample numbers and TRG score resolution may mask associations. In future, wider tissue and cancer micro-environments should be considered in tandem to better evaluate disease stage, and determine and monitor treatment, which will also help uncover biological associations. Still, these results are of immediate clinical interest, as no previous correlations between ATM expression and clinicohistopathological variables in CRC or rectal cancer had been reported at the time of

writing.

As well as studying ATM expression in this part of our studies, MSI-associated protein expression was also explored in the same samples. Microsatellite instability is a well-recognised mechanism leading to ATM dysfunction (as well as of the MRN complex proteins) (Cortes-Ciriano et al., 2017), and was investigated by evaluating the expression patterns of known mismatch repair proteins - MLH1, MSH2, MSH6, and PMS2.

Although in many instances, MSI has been associated with the loss of ATM function (Ham et al., 2006), any meaningful statistical analysis in the cohort of this thesis was impeded by the extremely low prevalence of MMR deficiencies. Both MLH1 and MSH2 were expressed in 100 percent of samples, and MSH6 and PMS2 to be expressed in more than 99 percent of samples. These results suggest an absence of MSI, and illustrate that mismatch error repair is on-going in these cells, in-keeping with their high mutation burden. It would have been interesting to further add to the literature concerning ATM expression in the context of MSI, something which will no doubt be pursued by future studies.

ATM expression was also evaluated in the context of radiotherapy. The response to radiotherapy was assessed by TRG scoring by two blinded, independent pathologists; to help negate biases. Although it is recognised that classification according to TRG systems in rectal cancer generally show a low concordance rate among pathologists (Chetty et al, 2012), disagreements between our two pathologists were rare. On those rare occasions where there were differences between the pathologists on scoring, a TRG score was provided by a third senior pathologist. When considering the TRG score, a lower value is representative of heightened radio-sensitivity; i.e. no disease detectable – 0. This system remains the most appropriate method of evaluating short-term radiotherapy response in such (historical) histological samples (and guides clinical decision making in living patients), although issues arise due to high scorer subjectivity and poor reproducibility between scorers, as discussed previously (Santos et al., 2013).

There are only few published recommendations for the handling of surgical resection specimens. Some authors consider standard processing protocols are appropriate if the tumour is clearly visible (Chetty et al, 2012). We have adopted the personal approach of Thies and Langer (2013) in embedding the entire tumour bed from the beginning. A graphical outline of their approach to standardising workup and reporting of TRG is presented in Figure 7.1.

Figure 7-1Thies and Langer approach to standardised workup and reporting of TRG

PHOTOGRAPHIC DOCUMENTATION
Photocopy or photograph of resection specimen (orientation and documentation of blocks and of histologically proven residual tumor)
Macroscopic description; tumor size (three-dimensional), distance to resection margins
WORK UP
Inking of the deep (circumferential) resection margin
Complete embedding of the macroscopically identifiable tumor bed, orientated from proximal to distal in 0.5 cm levels. If tumor bed >8 cm, significant regression in unlikely: first take blocks following the longitudinal and vertical largest dimension. If no or less residual tumor embed remaining tumor bed in second step. CRM is included in these blocks
All slides stained by Hematoxylin/eosin, selected blocks by periodic acid-schiff, Elastica van Gieson staining; immunohistochemistry may be helpful for discrimination of histiocytes and alterated tumor cells
If no residual tumor: another three step sections to confirm complete response
Resection margins oral, aboral
Additional macroscopic findings
Lymph node stations. Immunohistochemistry (pan-cytokeratin) if ypN_0
PATHOLOGICAL REPORT SHOULD INCLUDE
UICC ypTNM status (including L, V, Pn)
UICC R-status
Distance to circumferential resection margin (esophagus; rectum)
Grading, typing (according to WHO; additionally Lauren's type for upper GI adenocarinomas)
Histopathological tumor regression grade (e.g., Becker TRG 1a, 1b, 2, 3)

In our studies, tumours displaying low/absent ATM expression in peripheral cells were found to display a better response to radiotherapy; in-keeping with the hypothesis that a lack of ATM hampers DSB repair, and cues apoptosis after DSB induction by radiotherapy. Thus, the mechanisms we hypothesise to be allowing the cancer to grow and metastasise (a lack of cell cycle control in response to low DNA damage repair in the tumour periphery), also lead to its downfall after radiotherapy.

Long-term outcomes of radio-sensitivity were measured by disease-free and overall survival in our cohort patients. Positive responders to radiotherapy are known to benefit from lower rates of local disease recurrence, and improved disease-free survival (Camma et al., 2000; Colorectal Cancer Collaborative Group, 2001); outlining the importance of the treatment modality to rectal cancer.

Survival outcomes in the index cohort were analysed in patients who had received pre-operative radiotherapy (and compared to those that did not), although the study was limited by its statistical power, since only 55 patients in the cohort underwent pre-operative radiotherapy. In spite of this limited sample size, low ATM expression in peripheral tumour cells after radiotherapy was found to be convincingly associated with improved disease-free survival, and the association maintained significance in more stringent multivariate analyses (HR = 34.636 (2.160 − 555.293), $P = 0.012$). Furthermore, low ATM expression patterns in central tumour cells were also found to be significantly associated with improved disease-free survival after radiotherapy, after adjusting for known confounders, as covered previously (HR = 6.948 (1.192 − 40.504), $P = 0.031$). In the context of radiotherapy, cells which retain their ability to repair DNA damage are more likely to repair lesions, regain genomic stability and continue to

proliferate; whereas, cells with an impaired DDR pathway are more prone to undergo apoptosis (Roos and Kaina, 2013).

The findings of low ATM expression and increased DFS presented in this thesis is in direct opposition to the findings of Grabsch et al and Beggs et al, who identified an association between reduced ATM expression in CRC and worse disease-free survival (Beggs et al, 2012; Grabsch et al., 2006). However, their cohorts were not strictly limited to rectal cancer patients, nor stratified according to pre-operative radiotherapy treatment; in this study's subgroup analysis of patients who received adjuvant treatment, a clear association was not described. Grabsch et al, did however, report that reduced ATM expression was a poor prognostic marker in patients who received radiotherapy and/or chemotherapy, but this analysis was limited by small sample size (n = 38). On the other hand, Beggs et al found a trend towards reduced ATM expression and shorter disease-free survival which did not reach statistical significance in either the radiotherapy (n = 68) or chemotherapy (n = 436) group.

The studies by Grabsch et al and Beggs et al show some important differences to those in this thesis, including how patients were stratified, what treatments they received, and when they were analysed. Furthermore, ATM function is known to be affected by platinum-based chemotherapies (Kim et al, 2002). These points highlight how clinical heterogeneity can contribute to discordant results between cohorts and groups, and only additional studies in larger sample sizes will help better understand the role of ATM in these different situations. It suffices to say that, to date, ATM has proved to be an ambiguous prognostic marker in a variety of cancers; although the repeated associations do suggest the protein plays an important role in pathogenesis, even if the directions and mechanisms of action remain to be fully determined.

Despite these previous reports, the findings in this thesis establishes ATM expression as a negative prognostic marker in rectal cancer. Similar results have been noted in cervical cancer and non-small cell lung cancer (Roossink et al., 2012; Xing et al., 2008), whereas, high ATM expression is associated with superior survival in gastric cancer and B-cell chronic lymphocytic leukaemia (Austen et al., 2005; Kang et al., 2008). As discussed, these inconsistencies may be due to differences in cancer pathogenesis, characteristics, treatments, quantification of ATM expression and experimental methodology. For example, Xing et al attributed their ability to correlate reduced ATM expression with improved prognosis in NSCLC, to their measurement of the tumour: normal tissue (T/N) expression ratio, rather than on tumour expression alone. Similarly, the delineation of expression patterns between peripheral and central tumour samples has allowed for more consistent comparison between clinicohistopathological and clinical outcome data, and for novel associations to be discovered.

A major aim of this thesis was to determine whether ATM expression could be used in a predictive manner in rectal cancer. These analyses did not result in the discovery of any correlations between ATM expression and disease-free, or overall survival, in either subgroup; further reinforcing the interpretation that ATM expression predicts response to radiotherapy. Similarly, studies in human cell lines *in vitro* have been shown that the response to radiotherapy can be improved by inhibiting the action of ATM (Collis et al., 2003; Kim et al., 2002; Li et

al., 2006; Lin et al., 2012). Therefore, ATM may not only have a role as a predictive and prognostic marker, but may provide a therapeutic target for enhancing the effectiveness of radiotherapy, especially when its expression is regulated by hypoxia, which a major barrier to the effectiveness of radiotherapy.

There are a few limitations in the index study. Due to the retrospective nature of the study, complete clinicohistopathological data was not available for all patients; the problem of missing data in clinical medicine has been discussed at length elsewhere, and has led to large-scale efforts for developing statistical tools to account for it (Little et al., 2012). Additionally, potential patient and treatment selection biases in such retrospective studies could result in false positive, or false negative, findings; the patients were randomly selected based on chronological receipt of the operative specimens receive in the pathology department.
However main limitations of this study for ATM (and for the other biomarkers) were the small sample sizes of patients undergoing radiation treatment and small numbers of pretreatment tissue obtained. This prohibited an effective comparison of biomarker expression between cases that received neoadjuvant radiotherapy to those that did not receive any radiotherapy. Additionally the disease grade and vascular invasion categories, included groups in which no patients had died or experienced local recurrence, leading to large and unstable hazard ratios. Future prospective studies, or randomised clinical trials enlisting a larger cohort of patients undergoing preoperative radiotherapy and obtaining pretreatment biopsies, would be useful for further defining the causal relationships between ATM expression, TRG, and survival.

Despite these limitations, the study in this thesis demonstrates that ATM is a promising candidate as a clinical biomarker of radio-sensitivity and disease-free and overall survival in rectal cancer.

7.2.2 MRE11

Based on evidence linking MRE11 expression and cancer progression, as discussed (Gupta et al., 2013), w the expression of the MRN complex nucleic acid editor, MRE11, was studied in the context of rectal cancers.

However, although MRE11 appears to be a good candidate for such cancer expression studies, and given the associations discovered for ATM and what has been reported for MRE11, the study in this thesis failed to find any correlation between overall survival and the level of MRE11 expression in the centre or periphery of the rectal tumour samples. Similarly, a study by Sheriden et al, did not find a correlation between MRE11 expression (by immunohistochemistry) and survival or radio-sensitivity after neoadjuvant radiotherapy, in patients with CRCs (Sheridan et al., 2013). As with many studies in the field, however, this one was hampered by a small patient sample size.

However, although MRE11 expression alone was not found to significantly affect the survival outcomes of patients with rectal cancer in our cohort, we found that the level of MRE11 expression in the tumour centre or periphery significantly influenced the survival outcomes of

patients with concurrent perineural invasion, metastasis, or high-grade disease; suggesting that the predictive power of this marker increases as disease progresses, and additional clinicohistopathological variables are available to stratify patients. Specifically, survival outcomes (DFS and OS) were found to be significantly worse amongst: patients with high MRE11 expression in the tumour periphery who also had perineural invasion or metastasis; those with high MRE11 expression in the tumour centre and metastasis; and those with low MRE11 in the tumour centre and high-grade disease. Thus, high MRE11 expression was associated with disease metastasis (more advanced disease, with a greater amount of DNA damage), although low levels in the tumour centre were associated with higher grade disease.

Importantly, several previous studies have reported associations between MRE11 protein expression in tumours, and survival outcomes for patients suffering from different types of malignancies. For example, Choudhury et al demonstrated an independent association between tumour MRE11 expression and cause-specific survival in patients with muscle invasive bladder cancer after radical radiotherapy, but not after surgery - with low tumour MRE11 expression predictive of poor survival (Choudhury et al., 2010). Similarly, low tumour expression of the MRN complex (MRE11/RAD50/NBS1), was associated with a poor response to radiotherapy in patients with early breast cancer (Söderlund et al., 2007). Furthermore, a recently-published systematic review of biomarkers of tumour radio-sensitivity by Forker et al, identified MRE11 as one of the most promising predictive biomarkers in radiotherapy, although much remains to be determined about the mechanisms involved (Forker et al., 2015).

Mutations in proteins involved in DNA repair might also influence the clinical response to chemotherapeutic agents that damage DNA, such as the topoisomerase I poison camptothecin and its derivatives, irinotecan and topotecan(Vilar et al., 2008). In pre-clinical trials, Vilar and colleagues showed that colorectal cancer cell lines that were MRE11/RAD50-deficient, displayed enhanced sensitivity to irinotecan, or a combination treatment of thymidine and CPT/irinotecan (Rodriguez et al., 2008). Therefore, patients with MRE11-deficient tumours might be found to respond better to treatment with topoisomerase I poisons, compared to those with a fully-functional MRN complex; highlighting the importance of known MRE11 status in tumours with regards to treatment.

The loss of MRE11 protein expression also sensitises MSI positive colorectal cancer cells to inhibitors of the DNA repair enzyme, poly(ADP-ribose) polymerase (PARP) (McPherson et al., 2014; Vilar et al., 2011). In this regard, it is interesting that in the study of this thesis, adjuvant radiotherapy showed a protective trend in univariate analysis; in particular, low MRE11 expression in patients who did not receive adjuvant therapy showed significantly shorter overall survival than those who did.

Although further research is required to confirm any prognostic and predictive roles of MRE11 tumour expression in rectal cancer, our findings have important implications for the treatment of rectal cancer patients with perineural invasion, metastasis, or high-grade disease; where MRE11 expression was found to be associated with survival outcomes. In such cases, knowledge of MRE11 status can provide a clearer prognostic picture, and may affect clinical

management decisions.

7.2.3 NBS1

Another critical member of the MRN complex, is Nibrin. Mutations in this gene have been studied for a long time, and much is known about the role of NBS1 in the immune system, fertility, radiation sensitivity, and DSB repair (Difilippantonio et al., 2005; Varon et al., 2000; Warcoin et al., 2009).

NBS1, therefore, presents as an ideal candidate rectal cancer gene, with which it has previously been implicated (di Masi and Antoccia, 2008; Uhrhammer et al., 2000). However, that the study in this thesis showed only few associations in respect of NBS1 expression– although a marginal association between NBS1 expression in the tumour centre and patient overall survival was described. In comparison to the other genes here analysed, NBS1 expression appears to be a worse predictive marker of patient survival in rectal cancer, and was not found to greatly help delineate clinicohistopathological variables.

This may be because NBS1 is simply less important to rectal cancer progression and DSB repair during oncogenesis, relative to the other proteins studied – and there are several plausible reasons to hypothesise that this could be the case. Firstly, NBS1 activity in this context may be compensated for by other proteins, such as ATMIN (Zhang et al., 2012); making NBS1 expression alone less important to DNA repair, and therefore its expression less predictive of lesion repair. It is also worth considering that the high expression of the protein in the majority of tumours also reduces the dynamic range of the assay (Cox, 2012); compounded by the fact that a significant sub-set of patients showed complete absence of expression. Furthermore, there is no somatic or germ-line sequence data from the patients in the study cohort, the genetic integrity of the loci under investigation cannot be determined, which may impart differences on gene function in the tumour or elsewhere.

There are many reasons for why differences were not detected according to NBS1 expression. However, the results do not suggest that NBS1 expression is not important to rectal cancer. The investigation of this gene represents the major limitation of this thesis; due to the statistical statistical limitations (Ioannidis, 2005; Sprent, 2003). This is further impacted upon by patient cohort heterogeneity (i.e. sex and age), and treatment variation between patients, all of which will impact the cancer transcriptome and proteome (Uhlen et al., 2017).

In an ideal scenario, the associations of these proteins with rectal cancer would be investigated in a large cohort of patients that can be strictly classified according to the nature of their clinical history, genetics, specific treatments (dose and time), and other relevant clinicohistopathological features (Süt, 2014). Which is not to say that retrospective studies such as this from patient biobanks and collections are not central for discovering such associations and furthering the case for future, wider-scope studies. Indeed, the results here presented as a whole demonstrate the utility of such approaches.

NBS1 quantification in tumours was useful for building the predictive-three marker panel (Ho

et al, 2018) and further emphasises the utility of analysing different members of the same pathway/complex to glean more information about the disease state (Garcia-Campos et al., 2015).

7.2.4 RAD50

As has been shown for other members of the MRN complex, RAD50 expression has been found to correlate with a good response to treatment (Brandt et al., 2017). Furthermore, an adenovirus targeting RAD50 also showed promise in sensitising nasopharyngeal carcinoma cells to radiotherapy (Chang et al., 2016), further supporting its exploration in this context, and suggesting that RAD50 is central to the repair of radiation-induced DSBs.

To achieve the aims of the thesis, any correlations between the available clinicohistopathological features and treatment type were explored, with long-term (DFS and OS) radiotherapy responses after staining for RAD50. Low levels of RAD50 expression at early tumour stages, and in low-grade tumour sub-groups, was significantly associated with worse disease-free survival and overall survival, defining a relationship between post-operative tumour expression of RAD50 and prognosis. Specifically, the results from univariate and multivariate analyses showed that worse DFS outcomes were associated, not only with low RAD50 expression in the tumour centre, but also with perineural invasion. These data, therefore, indicate that post-operative RAD50 expression predicts long-term survival in rectal cancer, if evaluated alongside other tumour-related clinicohistopathological features, and supports the potential use of RAD50 as an early prognostic biomarker in rectal cancer

The role of RAD50 has also been evaluated in CRC by other studies. Gao et al demonstrated that RAD50 expression is reduced/low in MSI-positive CRCs, and is not associated with clinicohistopathological patient characteristics (Gao et al., 2008a). Conversely, RAD50 expression was increased in early stage primary MSS CRCs, and interestingly the oncogenic RAD50 frameshift mutation $(A)^9$ occurred in MSI, but not in MSS CRCs; suggesting that RAD50 might play different roles in these CRC phenotypic subtypes. The findings from this study, which show high or low levels of RAD50 to define different disease stages, highlight the complex roles of the protein during disease, and outline the importance of large sample sizes and accurate clinical denominators to identify robust, reproducible conclusions.

Importantly, RAD50 was also described as a prognostic biomarker for colorectal mucinous adenocarcinoma (delineating between 10%–15% of all cases) through an integrated analysis of genetic and epigenetic features (Wang et al., 2015). The authors of this study showed low RAD50 expression to be associated with poor prognosis in patients with MSS CRC, and postulated that increased RAD50 expression in MSS CRC could be a tumour suppressive cellular response to prevent further tumour progression (Wang et al., 2015). Based on these reports, further investigation of the molecular role of RAD50 in MSI and MSS CRCs could help inform patient responses to therapy.

In common with the other proteins investigated as part of this thesis, a larger sample size would

have been of benefit towards defining associations, as we will discuss. Furthermore, as this study also only involved two centres, larger, multi-center prospective studies are needed to validate the observations here reported.

7.2.5 Combinatorial panels – ATM/MRE11

The studies in this thesis provide conclusive evidence to support the further exploration of ATM, MRE11, NBS1 and RAD50 as biomarkers in rectal cancer. Still, despite the promise, no such biomarkers have been validated in randomised clinical trials, and their clinical value remains unclear.

Furthermore, the directions of the associations between DDR proteins and such outcome measures are highly variable (i.e. low or high expression of ATM has been associated with CRC survival outcomes; *please see previous sections*), and as discussed, this could be due to a number of reasons. Given this marked variation, and the clinical heterogeneity of cancer, combinatorial biomarker panels that aim to capture information about a greater number of disease-associated proteins in tandem, hold promise for helping to define the clinical picture in a greater proportion and number of patients; by analysing different members of the same pathway (i.e. the MRN complex) simultaneously, a more complete picture of the pathway's function within the tumour can be obtained (Rakha et al., 2010).

Based on the hypothesis that both ATM and MRE11 are integral to the detection of DNA damage and subsequent intracellular signalling following radiotherapy (and hence, that their deficiency would equate to increased radio-sensitivity), a two-marker panel of ATM and MRE11 expression was established by carrying out binary regression analysis of tumour samples and normal tissues.

When both ATM and MRE11 were combined into a single panel, subsequent analyses showed greater levels of association between biomarker expression and general clinicohistopathological parameters in our cohort, compared to either marker alone. The combined panel yielded a ROC-AUC value of 0.745, with high expression of the panel predicting poor histological tumour regression (i.e. a TRG score of 3) following radiotherapy. This value was superior to using ATM (0.618) or MRE11 (0.711) alone; demonstrating the increased power of the combinatorial approach. Furthermore, in both the neoadjuvant radiotherapy sub-cohort and the overall cohort, the combined ATM/MRE11 expression levels correlated with clinical survival outcomes; with high levels being associated with worse outcomes.

These results demonstrate that expression of these markers, when used together, are associated with both, early (histological tumour regression), and late (clinical survival) responses to neoadjuvant therapy in rectal cancers. Additionally, these findings support the hypothesis (that increased protein expression leads to increased cancer cell survival

after therapy), and are also consistent with the fact that inhibitors of ATM and the MRN complex have shown potential as radio-sensitising agents (Permuth-Wey et al., 2009).

On a final note, the role of lymph node status in predicting survival outcomes with regards to the 2-marker panel is intriguing, and further hints at the complexity of using biomarkers to predict patient outcomes. Among patients with LN-negative tumours in our cohort, there was no significant difference in survival between patients with a high or low, two-protein combined panel score in the tumour centre. In contrast, in patients with LN-positive tumours, a high two-protein combined panel score in the tumour centre was found to be significantly associated with worse survival. This suggests that the combined expression of ATM and MRE11 may be associated with LN involvement in relation to patient survival, and may represent increased DDR protein involvement in advanced disease stages.

7.2.6 Combinatorial Panels – MRN Complex, three-marker panel

To further explore the utility of a MRN complex combinatorial biomarker, a panel comprising the three MRE11, RAD50, and NBS1 was established. Combined MRN protein expression was found to have a high sensitivity and specificity in samples taken from both the centre and periphery of rectal tumours. Importantly, the sensitivity, specificity, and overall accuracy for the MRN panel was found to be higher than for the combined MRE11/ATM panel, both in the tumour centre and periphery; again supporting the argument that increased data points in such a setting serve to capture more of the clinical picture and increase the power to detect associations (Creixell et al., 2015; Samyn et al., 2015).

Notably, high expression levels of the three MRN complex proteins in the tumour centre was significantly associated with disease-free and overall survival. Interestingly, none of the other clinicohistopathological variables were significantly associated with combined MRN expression. Therefore, this panel appears to be specifically prognostic of DFS and OS.

A lack of associations between the three-marker panel and the other disease features in the cohort may be due to a lack of statistical power, as discussed, or because by combining data from three markers, the power of each marker, individually, is reduced; which, in limited sample sizes, is what allows for associations to be drawn from each marker independently. Furthermore, when the expression of the three-marker panel in the subset of patients that received pre-operative radiotherapy was examined, the association between combined MRN expression and outcome remained significant. This suggests that the prognostic value of this panel may be related to tumour radio-sensitivity.

Interestingly, high MRN protein levels have been found to be associated with better outcomes in some other cancer types. In early breast cancer, for example, patients with high MRN

complex expression experienced the greatest reduction in recurrence from radiotherapy (Söderlund et al., 2007). Furthermore, in two different studies of bladder cancer, high MRE11 expression was associated with better cancer-specific survival times in patients who also underwent radiotherapy, rather than a cystectomy (Choudhury et al., 2010; Laurberg et al., 2012). Therefore, the MRN complex may play a very different role in cancers arising from different tissues, or during the course of the same disease (Punt et al., 2017). In addition, it is also possible that the prognostic value of MRN complex expression is dependent on certain combinations of chemotherapeutics, radiotherapy regimens, and surgery, which vary between the treatment modalities preferred for different cancers and patients.

High combined MRN complex protein expression levels with DFS and OS was observed in LN-positive patients, but not in LN-negative patients; again, suggesting that these proteins are expressed during more advanced disease, where the mutation burden is greatest. In a study of rectal cancer patients undergoing long-course neoadjuvant chemo-radiotherapy, combining LN involvement with tumour grade was found to be prognostic for patient survival after treatment (Lindebjerg et al., 2009). Given these results, and the considerations discussed, it is possible that some biomarkers may specifically predict outcomes in patients with LN involvement or those without; again highlighting the need for further work in the area. Similarly, Quintanal-Villalonga and colleagues found that a mutated version of the FGFR4 gene was associated with overall survival only in LN-involved patients (Quintanal-Villalonga et al., 2017). Thus, the prognostic value of the MRN expression panel may be related to the lymph node involvement of the patient.

One mechanism that could lead to altered expression of the MRN complex proteins is defective MMR, as we have discussed. Along these lines, Giannini and colleagues found that the *MRE11* gene was mutated in MMR-deficient tumours and cell lines, but not in those with normal MMR function (Giannini et al., 2002). However, all of the tumours tested in the studies of this thesis expressed the two MMR proteins most frequently mutated in MMR-deficient patients, MLH1 and MSH2. Furthermore, the absence of MSH6 or PMS2 protein expression was not significantly associated with combined MRN expression, although this analysis was limited by the very small number of cases lacking expression of either of these proteins and requires validation in a larger cohort. Therefore, the mechanisms underlying the prognostic change in MRN expression identified here seems to be independent of the MMR pathway, and is a subject for further study.

The primary limitation of this three-marker panel study was the inability to analyse the relationship of combined MRN expression with tumour regression response. Only 10.6% of patients were classified as responders to radiotherapy, represented by a TRG score of 0 to 2. As such a small portion of the patients had good responses to radiotherapy, it also raises the possibility that combined MRN expression predicts late responses (survival) to treatment (radiotherapy), only in patients with relatively poor early responses.

Finally, since increased MRN protein expression is associated with worse outcomes in rectal cancer patients, reducing MRN protein expression and/or activity could improve response to

radiotherapy. Indeed, the MRN complex inhibitors, mirin and telomelysin, have great radio-sensitising effects, as has been shown in several pre-clinical studies (Dupré et al., 2008; Garner et al., 2009; Kuroda et al., 2012). As testament to their potential utility, telomelysin is currently undergoing Phase I and II trials for use in patients with melanoma (https://clinicaltrials.gov; NCT03190824), oesophageal cancer (NCT03213054), and hepatocellular carcinoma (NCT02293850).

Additionally, in patients with higher MRN expression (who are thus expected to have worse outcomes), additional radio-sensitising treatments could be used in combination with neoadjuvant radiotherapy to improve survival times. Heat treatment, for example, shows good radio-sensitising effects in cells *in vitro*, and is being explored in cancer patients (Dewey, 2009). Dynlacht and colleagues (2011) also found that heat radio-sensitisation was dependent on a functioning MRE11 protein further suggesting the utility of this treatment.

7.3 Study implications

By analysing the expression of four highly-conserved DNA damage response proteins with important roles in nucleic acid lesion repair and cell-cycle regulation, the studies in this thesis are able to identify associations between their expression, and outcomes in rectal cancer. These studies, therefore, implicated the DDR in rectal cancer pathophysiology and progression.

Importantly, many of these associations are novel, and many results arising from this work have already been published in peer-reviewed publications. Therefore, these studies provide important starting points for further refining the role of DDR proteins in rectal cancer.

In the case of ATM, the protein was shown to be more highly expressed in cancer tissue. Furthermore, its expression was highest in the periphery of tumours with a high histological grade, especially in older patients. Rectal tumours displaying low/absent ATM expression in peripheral cells, were also found to display a better response to radiotherapy, which correlated with improved DFS and OS in multivariate analyses.

In contrast, MRE11, was less informative, with its expression only having predictive value in patients with concurrent perineural invasion, metastasis, or high-grade disease. Given the reduced predictive power of MRE11 compared to ATM, it was reassuring to discover that the combined panel formed from both proteins showed improved predictive power to either marker alone, and allowed for DFS and OS to be analysed more accurately.

Still less informative individually proved to be NBS1, although a possible association between high expression and overall survival was reported.

Lastly, low levels of RAD50 were associated with worse disease-free and overall survival. Specifically, worse DFS outcomes were associated not only with low RAD50 expression in the tumour centre, but also with perineural invasion, highlighting the use of RAD50 expression

in combination with clinicohistopathological variables to discern novel disease features.

Finally, combination of the different proteins into a two- or three-marker biomarker panel was found to increase sensitivity and specificity, with regards to outcome measures, and the combinatorial panels proved to be a more powerful tool for describing associations with survival times in the cohort.

These results arising are particularly important and timely, as they define important predictive markers of radiotherapy success and disease progression, and because much work remains to be done to better delineate rectal cancers from colorectal tumours - aetiologically, mechanistically and prognostically; not to mention the fact that data on predictive markers of radiotherapy response are still lacking from the literature and clinic.

As developed economies face increasing non-communicable disease burdens and ageing populations (Boutayeb and Boutayeb, 2005; Christensen et al., 2009), determining which patients are likely to benefit from certain treatments to common human pathologies, will lead to large-scale cost savings and dramatic improvements in patient welfare (United Nations, 2015).

Together, given the different directions of associations in different regions of the tumours (i.e. high/low expression in the tumour centre/periphery) of different types of patients (i.e. those presenting with PNI), large studies incorporating longitudinal measurements (i.e. measure in the same patient over time), will help refine the mechanistic basis for these associations and further understand rectal cancer progression.

7.4 Study Limitations

Although this thesis has been successful in uncovering several novel associations between DDR protein expression and rectal cancer, much work remains to be done to address the discordancy present within the available literature and determine the mechanisms leading to disease and affecting disease progression/outcome. Several examples of associations between DDR protein expression and cancer survival outcomes running in different directions between studies exist; as has been discussed and shown for ATM and MRE11.

These differences are likely to be due to a number of different factors, which can include: the choice of antibody used to detect the protein of interest (Baker, 2015); variation in TRG scores between laboratories (as discussed); differences in patient cohort characteristics between studies (i.e. ethnicity, age, sex, and disease stage) (Payne et al., 2012); concurrent patient treatments (Glimelius et al., 1997); and small sample sizes (Sprent, 2003).

In the last decade, the number of irreproducible and erroneous biological and clinical medicine papers in the peer-reviewed literature due to small sample sizes has reached unprecedented levels (Baker, 2016; Emanuel et al., 2000), leading to considerable concern in the academic science community. Over the course of time, many results cannot be repeated (Halsey et al.,

2015), added confusion to researchers in the field and slowing scientific progress. This has helped the development of more transparent editorial/peer-review processes, such as those started by the *eLife* journal (https://elifesciences.org, funded by the major academic science funders, the Wellcome Trust, Howard Hughes Medical Institute and Max Planck Institutes), which adds the names of reviewers, and their comments, alongside the manuscript, as well as the development of peer comment sites (like https://pubpeer.com) and open access depositories, such as https://www.biorxiv.org and https://www.ebi.ac.uk/arrayexpress/, where data and results can be shared and validated by others in a more timely manner. While many journals do not accept submission that have been previously published, many prestigious journals such as *Science* and *Nature* do, with the journals standing to benefit from publishing science that has already been exposed to the tests of the community. In this respect, the tide is turning, and studies are being placed under increased scrutiny to ensure only the best work gets disseminated.

Part of the reason for this irreproducibility, many scientists suggest, is a lack of appropriate statistical testing of data (i.e. not applying Bonferroni corrections to multiple t-tests on the same dataset), leading researchers to publish false-positive and false-negative associations, which in many instances match their hypotheses (Halsey et al., 2015; Sprent, 2003). The pressures of *publish or perish,* are also believed to contribute to the dissemination of incomplete, rushed or fraudulent studies (Neill, 2008). Given the clinical and genetic heterogeneity of cancer patients, small sample sizes, and inappropriate statistical testing, can lead to spurious associations (as only a few extreme data points are sufficient to delineate mathematical differences between groups), and more likely, false-negatives. Thus, having a larger sample size not only provides more statistical power to detect differences, but also helps avoid not finding them (Devine and Smith, 1998; Faraggi and Reiser, 2002; Faul et al., 2009).

In this thesis, although all analyses were carried out in a blinded fashion and appropriate statistical frameworks were used to test for associations, the studies still suffered from a relatively small number of samples, especially when considering patients who underwent pre-operative radiotherapy (n = 55). It is important to remember that outbred human volunteers, show more phenotypic variability (due to different genes and environmental exposures) that common laboratory animal strains (which are often inbred for many generations), leading to greater standard deviations within outcome readings in humans; not to mention the fact that the biological pathways underpinning disease are different between species (Davis, 2008; Editorial, 2013).

Ideally, prior to commencing any such study, power calculations should be carried out to determine the number of samples required to statistically detect differences of a given magnitude (Suresh and Chandrashekara, 2012). Thus, if the experiment is powered (i.e. has sufficient samples) to detect 5% differences between two groups (at an alpha of 0.05), and none are detected, it is more likely that the associations do not exist; compared to when the experiment is underpowered, the result obscured, and the associations missed. However, although power calculations *a priori* are best in terms of experimental design, often the number of samples is limited by factors beyond the control of the researchers, not least in retrospective

studies, such as this one, where patient samples were collected historically, and a degree of missing data was present.

Given the immense power of banked patient samples to inform upon disease pathogenesis and the problems associated with recruiting large numbers of patients at any one time, it is encouraging to see increasing numbers of biobanks (where diverse patient samples are stored) opening up in academic and private centres around the world (De Souza and Greenspan, 2013). Good examples of these include the Cambridge BioResource at the University of Cambridge (https://www.cambridgebioresource.group.cam.ac.uk), The Human Immune Monitoring Centre at Stanford University (http://iti.stanford.edu/himc.html), and the Biospecimen Repository at Yale University (https://medicine.yale.edu/obgyn/drs/facilities/index.aspx), all of which aim to collect fresh human tissue samples, alongside medical history and genetic information to facilitate research.

Such resources will be essential for testing hypotheses and determining molecular mechanisms in future, as new knowledge comes to the fore. Furthermore, as public healthcare education continues to increase with the help of the internet and efforts aimed to place patients at the centre of care decisions and management (Barratt, 2008), more volunteers will be encouraged to donate the diseased and healthy tissue required for such research to take place. Of course, such a future, in which vast amounts of personal data are stored and accessible to different researchers, will require extremely secure data networks and robust ethical frameworks to safeguard volunteers (Shoenbill et al., 2014; Switula, 2000).

Another and important limitation concerned with this thesis is the fact that the patient cohort is composed predominantly of elderly male patients, making it difficult to extend the observations to females or younger patients; who are known to show profound differences in the molecular aspects of their physiology (Dopico et al., 2015; Karastergiou et al., 2012; Short et al., 2013). As this thesis represents a retrospective study, the make-up of the cohort cannot be controlled. Whilst the absolute number of colorectal cancers has been reported to be similar between males and females, male mortality and incidence rates are increasing at a faster rate than females, and males are more likely to have rectal tumours compared to females; who are more likely to have colonic cancers (Gao et al., 2008b; Murphy et al., 2011; Purim et al., 2013). Given these considerations, analysing the expression of these DDR proteins in additional cohorts seems essential to firmly establish the associations we here report with rectal cancer. Nevertheless, age, sex and other clinicohistopathological variables were added as confounding variables to multivariate analyses, allowing the values to be taken into account as much as possible.

Finally, it is worth considering the limitations associated with not knowing the tumour, germ-line, or somatic cell sequence of the patients in question – cancer is a genetic disease, after all (Balmain et al., 2003; Hayes and Kim, 2015). Given the rapidly falling cost of DNA sequencing (Church, 2006), at minimal it would have been an advantage to know the base-pair sequence of the proteins of interest, which could have been determined by capillary sequencing and PCR. For example, knowing a patient carries a lesion within ATM that reduces expression, would

allow the results from that patient to be considered separately to those without a mutation, allowing us to further refine how the DDR pathway works in disease.

Such genetic data would also allow determination of, for example, whether any patients were carrying lesions within genes of interest from birth, or whether any deleterious variants arose sporadically (both of which could influence their expression levels) in the cancer cell lineage – helping to define familial and non-familial cases (Chubb et al., 2016).

Furthermore, knowing which other variants are present genome-wide, could also be used to help refine oncogenic mechanism (i.e. which other variants are found to segregate with this cub-type of patients, and not this one), much in the same way that GWAS illustrated the loci predisposing to common, complex human diseases and traits (Sniekers et al., 2017). Indeed, a recent study of 120,000 cancer patients identified more than 80 loci contributing an increased risk to breast cancer (Michailidou et al., 2015). Similar studies in rectal cancer, and NGS of the index samples, would again allow for genetic risk to be better accounted for in such experiments.

Knowing the genetic background of the patients would also help stratify patients according to disease risk, and monitor outcomes with respect to the genetic code. Prognostic genetic signatures (including transcription from known common gene variants) have also been documented for a large number of diseases (Ferreira et al., 2014; McKinney et al., 2010), including cancer (van ’t Veer et al., 2002; Stratton et al., 2009), and could be combined with DDR protein expression with respect to outcomes.

However, having too much genetic data from a patient, and too few patients, will lead to erroneous results; highlighting that studies need to be powered to detect differences with respect to the technology employed. This is not surprising given that identifying causal mutations from a 3.3 billion sequence of letters again requires large sample sizes (to define health and an absence of risk) and high-fidelity sequencing to be able to detect single nucleotide polymorphisms that could drive disease.

7.5 Summary

Several novel associations are found in this thesis between rectal tumour expression of key DNA damage repair proteins and patient outcomes; which also considered the response to radiotherapy.

Highlights from the work in the thesis include the publication of two predictive combinatorial biomarkers for rectal cancer – where co-expression of these DNA damage response proteins were found associated with disease outcomes. These studies highlight the important role of these proteins in rectal cancer pathogenesis; adding to the numerous reports implicating them in more heterogeneous colorectal cancer cohorts. The fact that differences in their expression are predictive of disease outcomes emphasises their importance as key modulators of oncogenesis and drug- and gene-therapy targets, which may aim to boost complex activity or

repair genomic lesions using modern base-editing approaches.

These studies pave the way for numerous further studies, not only in rectal cancer, but also in other tumours and pathophysiological and developmental processes. A particularly interesting avenue of research, would be the exploration of these markers in rectal cancer patients in a non-invasive manner – such as by sampling cell-free DNA in blood samples (Diaz and Bardelli, 2014), which would also allow for longitudinal measures to be considered in the statistical models. Such approaches would allow for the parallel understanding of genetic sequence during the disease course, although how the peripheral circulation represents the nature of tumour *in situ* is a complex problem and remains an area of intensive investigation.

Future approaches will be discussed in the following chapter, and it suffices here to say that it is an incredibly exciting time in clinical medicine, where the heritage literature, research laboratory, and advances in statistics and technology, are increasingly clarifying the pathological basis of disease – welcome to the genomic age.

7.6 References

Abdel-Fatah, T.M.A., Arora, A., Moseley, P., Coveney, C., Perry, C., Johnson, K., Kent, C., Ball, G., Chan, S., and Madhusudan, S. (2014). ATM, ATR and DNA-PKcs expressions correlate to adverse clinical outcomes in epithelial ovarian cancers. BBA Clin. *2*, 10–17.

Aschele, C., Cionini, L., Lonardi, S., Pinto, C., Cordio, S., Rosati, G., Artale, S., Tagliagambe, A., Ambrosini, G., Rosetti, P., et al. (2011). Primary tumor response to preoperative chemoradiation with or without oxaliplatin in locally advanced rectal cancer: Pathologic results of the STAR-01 randomized phase III trial. J. Clin. Oncol. *29*, 2773–2780.

Austen, B., Powell, J.E., Alvi, A., Edwards, I., Hooper, L., Starczynski, J., Taylor, A.M.R., Fegan, C., Moss, P., and Stankovic, T. (2005). Mutations in the ATM gene lead to impaired overall and treatment-free survival that is independent of IGVH mutation status in patients with B-CLL. Blood *106*, 3175–3182.

Bai, A.H.C., Tong, J.H.M., To, K.F., Chan, M.W.Y., Man, E.P.S., Lo, K.W., Lee, J.F.Y., Sung, J.J.Y., and Leung, W.K. (2004). Promoter hypermethylation of tumor-related genes in the progression of colorectal neoplasia. Int. J. Cancer *112*, 846–853.

Baker, M. (2015). Blame it on the antibodies. Nature *521*, 274–276.

Baker, M. (2016). 1,500 scientists lift the lid on reproducibility. Nature *533*, 452–454.

Balmain, A., Gray, J., and Ponder, B. (2003). The genetics and genomics of cancer. Nat. Genet. *33*, 238–244.

Barratt, A. (2008). Evidence Based Medicine and Shared Decision Making: The challenge of getting both evidence and preferences into health care. Patient Educ. Couns. *73*, 407–412.

Bartkova, J., Hořejší, Z., Koed, K., Krämer, A., Tort, F., Zieger, K., Guldberg, P., Sehested, M., Nesland, J.M., Lukas, C., et al. (2005). DNA damage response as a candidate anti-cancer barrier in early human tumorigenesis. Nature *434*, 864–870.

Beggs, A., Domingo, E., McGregor, M., Presz, M., Johnstone, E., Midgley, R., et al. (2012). Loss of expression of the double strand break repair protein ATM is associated with worse prognosis in colorectal cancer and loss of Ku70 expression is associated with CIN. Oncotarget. *3*, 1348–1355.

Boutayeb, A., and Boutayeb, S. (2005). The burden of non communicable diseases in developing countries. Int. J. Equity Health *4*, 2.

Brandt, S., Samartzis, E.P., Zimmermann, A.K., Fink, D., Moch, H., et al. (2017). Lack of MRE11-RAD50-NBS1 (MRN) complex detection occurs frequently in low-grade epithelial ovarian cancer. BMC Cancer. *17*, 44.

Buccione, R., Caldieri, G., and Ayala, I. (2009). Invadopodia: Specialized tumor cell structures for the focal degradation of the extracellular matrix. Cancer Metastasis Rev. *28*, 137–149.

Cam, H., Easton, J.B., High, A., and Houghton, P.J. (2010). mTORC1 signaling under hypoxic conditions is controlled by atm-dependent phosphorylation of HIF-1α. Mol. Cell *40*, 509–520.

Camma, C., Giunta, M., Fiorica, F., Pagliaro, L., Craxi, A., and Cottone, M. (2000). Preoperative radiotherapy for resectable rectal cancer: A meta-analysis. JAMA *284*, 1008–1015.

Chang, L., Huang, J., Wang, K., Li, J., Yan, R., Zhu, L., Ye, J., and Wu, X. (2016). Targeting Rad50 sensitizes human nasopharyngeal carcinoma cells to radiotherapy. BMC Cancer. *16*, 190.

Chetty, R., Gill, P., Govender, D., Bateman, A., Chang, H.J., Deshpande, V., et al. (2012). International study group on rectal cancer regression grading: interobserver variability with

commonly used regression grading systems. Hum Pathol. *43*, 1917–23.
Choudhury, A., Nelson, L.D., Teo, M.T.W., Chilka, S., Bhattarai, S., Johnston, C.F., Elliott, F., Lowery, J., Taylor, C.F., Churchman, M., et al. (2010). MRE11 expression is predictive of cause-specific survival following radical radiotherapy for muscle-invasive bladder cancer. Cancer Res. *70*, 7017–7026.
Christensen, K., Doblhammer, G., Rau, R., and Vaupel, J.W. (2009). Ageing populations: the challenges ahead. Lancet *374*, 1196–1208.
Chubb, D., Broderick, P., Dobbins, S.E., Frampton, M., Kinnersley, B., Penegar, S., Price, A., Ma, Y.P., Sherborne, A.L., Palles, C., et al. (2016). Rare disruptive mutations and their contribution to the heritable risk of colorectal cancer. Nat. Commun. *7*, 11883.
Church, G. (2006). The Race for the $ 1000 Genome. Science. *311*, 1544–1546.
Collis, S.J., Swartz, M.J., Nelson, W.G., and DeWeese, T.L. (2003). Enhanced radiation and chemotherapy-mediated cell killing of human cancer cells by small inhibitory RNA silencing of DNA repair factors. Cancer Res. *63*, 1550–1554.
Colorectal Cancer Collaborative Group. (2001). Adjuvant radiotherapy for rectal cancer: a systematic overview of 8,507 patients from 22 randomised trials. Lancet *358*, 1291–1304.
Compton, C., Fenoglio-Preiser, C.M., Pettigrew, N., and Fielding, L.P. (2000). American Joint Committee on Cancer Prognostic Factors Consensus Conference: Colorectal Working Group. Cancer *88*, 1739–1757.
Cortes-Ciriano, I., Lee, S., Park, W.Y., Kim, T.M., and Park, P.J. (2017). A molecular portrait of microsatellite instability across multiple cancers. Nat. Commun. *8*, 15180.
Cox, K.L., Devanarayan, V., Kriauciunas, A., et al. (2012). Immunoassay Methods. [Updated 2014 Dec 24]. In: Sittampalam GS, Coussens NP, Brimacombe K, et al., editors. *Assay Guidance Manual [Internet]*. Bethesda (MD): Eli Lilly & Company and the National Center for Advancing Translational Sciences; 2004-. Available from: https://www.ncbi.nlm.nih.gov/books/NBK92434/
Creixell, P., Reimand, J., Haider, S., Wu, G., Shibata, T., Vazquez, M., Mustonen, V., Gonzalez-Perez, A., Pearson, J., Sander, C., et al. (2015). Pathway and network analysis of cancer genomes. Nat. Methods *12*, 615–621.
Davis, M.M. (2008). A Prescription for Human Immunology. Immunity *29*, 835–838.
De Souza, Y.G., and Greenspan, J.S. (2013). Biobanking past, present and future: responsibilities and benefits. AIDS *27*, 303–312.
Devine, O.J., and Smith, J.M. (1998). Estimating sample size for epidemiologic studies: The impact of ignoring exposure measurement uncertainty. Stat. Med. *17*, 1375–1389.
Dewey, W.C. (2009). Arrhenius relationships from the molecule and cell to the clinic. Int. J. Hyperth. *25*, 3–20.
Diaz, L.A., and Bardelli, A. (2014). Liquid biopsies: Genotyping circulating tumor DNA. J. Clin. Oncol. *32*, 579–586.
Difilippantonio, S., Celeste, A., Fernandez-Capetillo, O., Chen, H.T., San Martin, B.R., Van Laethem, F., Yang, Y.P., Petukhova, G. V., Eckhaus, M., Feigenbaum, L., et al. (2005). Role of Nbs1 in the activation of the Atm kinase revealed in humanized mouse models. Nat. Cell Biol. *7*, 675–685.
di Masi, A., and Antoccia, A. (2008). NBS1 Heterozygosity and Cancer Risk. Curr. Genomics *9*, 275–281.
Dopico, X.C., Evangelou, M., Ferreira, R.C., Guo, H., Pekalski, M.L., Smyth, D.J., Cooper, N., Burren, O.S., Fulford, A.J., Hennig, B.J., et al. (2015). Widespread seasonal gene expression reveals annual differences in human immunity and physiology. Nat. Commun. *6*, 7000.
Dupré, A., Boyer-Chatenet, L., Sattler, R.M., Modi, A.P., Lee, J.H., Nicolette, M.L., Kopelovich, L., Jasin, M., Baer, R., Paull, T.T., et al. (2008). A forward chemical genetic

screen reveals an inhibitor of the Mre11-Rad50-Nbs1 complex. Nat. Chem. Biol. *4*, 119–125.
Dynlacht, J.R., Batuello, C.N., Lopez, J.T., Kim, K.K., and Turchi, J.J. (2011). Identification of Mre11 as a target for heat radiosensitization. Radiat. Res. *176*, 323–332.
Editorial (2013). Of men, not mice. Nat. Med. *19*, 379.
Emanuel, E.J., Wendler, D., and Grady, C. (2000). What makes clinical research ethical? JAMA *283*, 2701–2711.
Faraggi, D., and Reiser, B. (2002). Estimation of the area under the ROC curve. Stat. Med. *21*, 3093–3106.
Faul, F., Erdfelder, E., Buchner, A., and Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. Behav. Res. Methods *41*, 1149–1160.
Feig, C., Jones, J.O., Kraman, M., Wells, R.J.B., Deonarine, A., Chan, D.S., Connell, C.M., Roberts, E.W., Zhao, Q., Caballero, O.L., et al. (2013). Targeting CXCL12 from FAP-expressing carcinoma- associated fi broblasts synergizes with anti – PD-L1 immunotherapy in pancreatic cancer. Proc Natl Acad Sci U S A *110*, 20212–20217.
Feng, X., Li, H., Dean, M., Wilson, H.E., Kornaga, E., Enwere, E.K., Tang, P., Paterson, A., Lees-Miller, S.P., Magliocco, A.M., et al. (2015). Low ATM protein expression in malignant tumor as well as cancer-associated stroma are independent prognostic factors in a retrospective study of early-stage hormone-negative breast cancer. Breast Cancer Res. *17*, 65.
Feng, X., Li, H., Kornaga, E.N., Dean, M., Lees-Miller, S.P., Riabowol, K., Magliocco, A.M., Morris, D., Watson, P.H., Enwere, E.K., et al. (2016). Low Ki67/high ATM protein expression in malignant tumors predicts favorable prognosis in a retrospective study of early stage hormone receptor positive breast cancer. Oncotarget *7*, 85798-85812.
Ferreira, R.C., Guo, H., Coulson, R.M.R., Smyth, D.J., Pekalski, M.L., Burren, O.S., Cutler, A.J., Doecke, J.D., Flint, S., McKinney, E.F., et al. (2014). A type I Interferon transcriptional signature precedes autoimmunity in children genetically at risk for type 1 diabetes. Diabetes *63*, 2538–2550.
Flint, T.R., Janowitz, T., Connell, C.M., Roberts, E.W., Denton, A.E., Coll, A.P., Jodrell, D.I., and Fearon, D.T. (2016). Tumor-Induced IL-6 Reprograms Host Metabolism to Suppress Anti-tumor Immunity. Cell Metab. *24*, 672–684.
Forker, L.J., Choudhury, A., and Kiltie, A.E. (2015). Biomarkers of Tumour Radiosensitivity and Predicting Benefit from Radiotherapy. Clin. Oncol. *27*, 561–569.
Freitas, A.A., and De Magalhães, J.P. (2011). A review and appraisal of the DNA damage theory of ageing. Mutat. Res. - Rev. Mutat. Res. *728*, 12–22.
Gao, J., Zhang, H., Arbman, G., and Sun, X.-F. (2008a). The different roles of hRAD50 in microsatellite stable and unstable colorectal cancers. Dis. Markers *24*, 127–134.
Gao, R.-N., Neutel, C.I., and Wai, E. (2008b). Gender differences in colorectal cancer incidence, mortality, hospitalizations and surgical procedures in Canada. J. Public Health (Oxf). *30*, 194–201.
Garcia-Campos, M.A., Espinal-Enriquez, J., and Hernandez-Lemus, E. (2015). Pathway Analysis: State of the Art. Front Physiol *6*, 383.
Garner, K.M., Pletnev, A.A., and Eastman, A. (2009). Corrected structure of mirin, a small-molecule inhibitor of the Mre11-Rad50-Nbs1 complex. Nat. Chem. Biol. *5*, 129–130.
Gatenby, R.A., Smallbone, K., Maini, P.K., Rose, F., Averill, J., Nagle, R.B., Worrall, L., and Gillies, R.J. (2007). Cellular adaptations to hypoxia and acidosis during somatic evolution of breast cancer. Br. J. Cancer *97*, 646–653.
Giannini, G., Ristori, E., Cerignoli, F., Rinaldi, C., Zani, M., Viel, A., Ottini, L., Crescenzi, M., Martinotti, S., Bignami, M., et al. (2002). Human MRE11 is inactivated in mismatch repair-deficient cancers. EMBO Rep. *3*, 248–254.
Glimelius, B., Ekström, K., Hoffman, K., Graf, W., Sjödén, P.O., Haglund, U., Svensson, C.,

Enander, L.K., Linné, T., Sellström, H., et al. (1997). Randomized comparison between chemotherapy plus best supportive care with best supportive care in advanced gastric cancer. Ann. Oncol. *8*, 163–168.
Grabsch, H., Dattani, M., Barker, L., Maughan, N., Maude, K., Hansen, O., Gabbert, H.E., Quirke, P., and Mueller, W. (2006). Expression of DNA double-strand break repair proteins ATM and BRCA1 predicts survival in colorectal cancer. Clin. Cancer Res. *12*, 1494–1500.
Guarch, R., Cortés, J.M., Lawrie, C.H., and López, J.I. (2016). Multi-site tumor sampling (MSTS) improves the performance of histological detection of intratumor heterogeneity in clear cell renal cell carcinoma (CCRCC). F1000Research *5*, 2020.
Gupta, G.P., Vanness, K., Barlas, A., Manova-Todorova, K.O., Wen, Y.H., and Petrini, J.H.J. (2013). The Mre11 Complex Suppresses Oncogene-Driven Breast Tumorigenesis and Metastasis. Mol. Cell *52*, 353–365.
Halsey, L.G., Curran-Everett, D., Vowler, S.L., and Drummond, G.B. (2015). The fickle P value generates irreproducible results. Nat. Methods *12*, 179–185.
Ham, M.F., Takakuwa, T., Luo, W.J., Liu, A., Horii, A., and Aozasa, K. (2006). Impairment of double-strand breaks repair and aberrant splicing of ATM and MRE11 in leukemia-lymphoma cell lines with microsatellite instability. Cancer Sci. *97*, 226–234.
Han, S., Gonzalo, D.H., Feely, M., Delitto, D., Behrns, K.E., Beveridge, M., Zhang, D., Thomas, R., Trevino, J.G., Schmittgen, T.D., et al. (2017). The pancreatic tumor microenvironment drives changes in miRNA expression that promote cytokine production and inhibit migration by the tumor associated stroma. Oncotarget *8*, 54054-54067.
Hayes, D.N., and Kim, W.Y. (2015). The next steps in next-gen sequencing of cancer genomes. J. Clin. Invest. *125*, 462–468.
Ho, V., Chung, L., Singh, A., Lea, V., Abubakar, A., Lim, SH., Ng, W., Lee, M., de Souza, P., Shin, J and Lee, CS. (2018). Overexpression of the MRE11-RAD50-NBS1 (MRN) complex in rectal cancer correlates with poor response to neoadjuvant radiotherapy and prognosis. BMC Cancer. *18*, 869.
Ioannidis, J.P.A.. (2005). Why most published research findings are false. PLoS Med. *2*, e124.
Jackson, S.P. (2002). Sensing and repairing DNA double-strand breaks. Carcinogenesis *23*, 687–696.
Kang, B., Guo, R.F., Tan, X.H., Zhao, M., Tang, Z. Bin, and Lu, Y.Y. (2008). Expression status of ataxia-telangiectasia-mutated gene correlated with prognosis in advanced gastric cancer. Mutat. Res. - Fundam. Mol. Mech. Mutagen. *638*, 17–25.
Karastergiou, K., Smith, S.R., Greenberg, A.S., and Fried, S.K. (2012). Sex differences in human adipose tissues – the biology of pear shape. Biol. Sex Differ. *3*, 13.
Kim, W., Vo, Q.N., Shrivastav, M., Lataxes, T.A., and Brown, K.D. (2002). Aberrant methylation of the ATM promoter correlates with increased radiosensitivity in a human colorectal tumor cell line. Oncogene *21*, 3864–3871.
Kraman, M., Bambrough, P.J., Arnold, J.N., Roberts, E.W., Magiera, L., Jones, J.O., Gopinathan, A., Tuveson, D.A., and Fearon, D.T. (2010). Suppression of antitumor immunity by stromal cells expressing fibroblast activation protein-alpha. Science *330*, 827–830.
Kuroda, S., Urata, Y., and Fujiwara, T. (2012). Ataxia-telangiectasia mutated and the Mre11-Rad50-NBS1 complex: Promising targets for radiosensitization. Acta Med. Okayama *66*, 83–92.
Laurberg, J.R., Brems-Eskildsen, A.S., Nordentoft, I., Fristrup, N., Schepeler, T., Ulhøi, B.P., Agerbæk, M., Hartmann, A., Bertz, S., Wittlinger, M., et al. (2012). Expression of TIP60 (tat-interactive protein) and MRE11 (meiotic recombination 11 homolog) predict treatment-specific outcome of localised invasive bladder cancer. BJU Int. *110*, E1228-36.
Lee, J.-H., and Paull, T.T. (2007). Activation and regulation of ATM kinase activity in response

to DNA double-strand breaks. Oncogene *26*, 7741–7748.
Li, W., Jian, W., Xiaoping, X., Yingfeng, L., Tao, X., and Xiaoyan, X. (2006). Enhanced radiation-mediated cell killing of human cervical cancer cells by small interference RNA silencing of ataxia telangiectasia-mutated protein. Int. J. Gynecol. Cancer. *16*, 1620–1630.
Lin, C., Yu, Y., Zhao, H.G., Yang, A., Yan, H., and Cui, Y. (2012). Combination of quercetin with radiotherapy enhances tumor radiosensitivity in vitro and in vivo. Radiother. Oncol. *104*, 395–400.
Lindebjerg, J., Spindler, K.L.G., Ploen, J., and Jakobsen, A. (2009). The prognostic value of lymph node metastases and tumour regression grade in rectal cancer patients treated with long-course preoperative chemoradiotherapy. Color. Dis. *11*, 264–269.
Little, R.J., D'Agostino, R., Cohen, M.L., Dickersin, K., Emerson, S.S., Farrar, J.T., Frangakis, C., Hogan, J.W., Molenberghs, G., Murphy, S.A., et al. (2012). The Prevention and Treatment of Missing Data in Clinical Trials. N. Engl. J. Med. *367*, 1355–1360.
McKinney, E.F., Lyons, P. a, Carr, E.J., Hollis, J.L., Jayne, D.R.W., Willcocks, L.C., Koukoulaki, M., Brazma, A., Jovanovic, V., Kemeny, D.M., et al. (2010). A CD8+ T cell transcription signature predicts prognosis in autoimmune disease. Nat. Med. *16*, 586–591.
McPherson, L.A., Shen, Y., and Ford, J.M. (2014). Poly (ADP-ribose) polymerase inhibitor LT-626: Sensitivity correlates with MRE11 mutations and synergizes with platinums and irinotecan in colorectal cancer cells. Cancer Lett. *343*, 217–223.
Michailidou, K., Beesley, J., Lindstrom, S., Canisius, S., Dennis, J., Lush, M.J., Maranian, M.J., Bolla, M.K., Wang, Q., Shah, M., et al. (2015). Genome-wide association analysis of more than 120,000 individuals identifies 15 new susceptibility loci for breast cancer. Nat. Genet. *47*, 373–380.
Murphy, G., Devesa, S.S., Cross, A.J., Inskip, P.D., McGlynn, K.A., and Cook, M.B. (2011). Sex disparities in colorectal cancer incidence by anatomic subsite, race and age. Int. J. Cancer *128*, 1668–1675.
United Nations. (2015). World Population Prospects: The 2015 Revision. United Nations Econ. Soc. Aff. *XXXIII*, 1–66.
Neill, U.S. (2008). Publish of perish, but at what cost? J. Clin. Invest. *118*, 2368.
Payne, R.A., Abel, G.A., and Simpson, C.R. (2012). A retrospective cohort study assessing patient characteristics and the incidence of cardiovascular disease using linked routine primary and secondary care data. BMJ Open *2*, e000723.
Pelosof, L.C., and Gerber, D.E. (2010). Paraneoplastic syndromes: an approach to diagnosis and treatment. Mayo Clin Proc *85*, 838–854.
Permuth-Wey, J., Boulware, D., Valkov, N., Livingston, S., Nicosia, S., Lee, J.H., Sutphen, R., Schildkraut, J., Narod, S., Parker, A., et al. (2009). Sampling strategies for tissue microarrays to evaluate biomarkers in ovarian cancer. Cancer Epidemiol. Biomarkers Prev. *18*, 28–34.
Punt, C.J.A., Koopman, M., and Vermeulen, L. (2017). From tumour heterogeneity to advances in precision treatment of colorectal cancer. Nat. Rev. Clin. Oncol. *14*, 235–246.
Purim, O., Gordon, N., and Brenner, B. (2013). Cancer of the colon and rectum: potential effects of sex-age interactions on incidence and outcome. Med. Sci. Monit. *19*, 203–209.
Quail, D.F., and Joyce, J.A. (2013). Microenvironmental regulation of tumor progression and metastasis. Nat Med *19*, 1423–1437.
Quintanal-Villalonga, Á., Carranza-Carranza, A., Meléndez, R., Ferrer, I., Molina-Pinelo, S., and Paz-Ares, L. (2017). Prognostic Role of the FGFR4-388Arg Variant in Lung Squamous-Cell Carcinoma Patients With Lymph Node Involvement. Clin. Lung Cancer *18*, 667–674.e1.
Rakha, E.A., Reis-Filho, J.S., and Ellis, I.O. (2010). Combinatorial biomarker expression in breast cancer. Breast Cancer Res. Treat. *120*, 293–308.
Reddy, J.P., Peddibhotla, S., Bu, W., Zhao, J., Haricharan, S., Du, Y.-C.N., Podsypanina, K., Rosen, J.M., Donehower, L.A., and Li, Y. (2010). Defining the ATM-mediated barrier to

tumorigenesis in somatic mammary cells following ErbB2 activation. Proc. Natl. Acad. Sci. *107*, 3728–3733.
Rodriguez, R., Hansen, L.T., Phear, G., Scorah, J., Spang-Thomsen, M., Cox, A., Helleday, T., and Meuth, M. (2008). Thymidine selectively enhances growth suppressive effects of camptothecin/irinotecan in MSI+ cells and tumors containing a mutation of MRE11. Clin. Cancer Res. *14*, 5476–5483.
Roos, W.P., and Kaina, B. (2013). DNA damage-induced cell death: From specific DNA lesions to the DNA damage response and apoptosis. Cancer Lett. *332*, 237–248.
Roossink, F., Wieringa, H.W., Noordhuis, M.G., Ten Hoor, K.A., Kok, M., Slagter-Menkema, L., Hollema, H., De Bock, G.H., Pras, E., De Vries, E.G.E., et al. (2012). The role of ATM and 53BP1 as predictive markers in cervical cancer. Int. J. Cancer *131*, 2056–2066.
Russell, R., Perkhofer, L., Liebau, S., Lin, Q., Lechel, A., Feld, F.M., Hessmann, E., Gaedcke, J., Güthle, M., Zenke, M., et al. (2015). Loss of ATM accelerates pancreatic cancer formation and epithelial-mesenchymal transition. Nat. Commun. *6*, 7677.
Samyn, B., Samyn, B., Sergeant, K., Sergeant, K., Castanheira, P., Castanheira, P., Faro, C., Faro, C., Beeumen, J. Van, and Beeumen, J. Van (2015). Pathway and network analysis of cancer genomes. Nat. Methods *2*, 1–6.
Santos, M.D., Silva, C., Rocha, A., Matos, E., Nogueira, C., and Lopes, C. (2013). Tumor regression grades: Can they influence rectal cancer therapy decision tree? Int. J. Surg. Oncol. *2013*, 572149.
Sheridan, J., Tosetto, M., Gorman, J., O'Donoghue, D., Sheahan, K., Hyland, J., Mulcahy, H., Gibbons, D., and O'Sullivan, J. (2013). Effects of radiation on levels of DNA damage in normal non-adjacent mucosa from colorectal cancer cases. J. Gastrointest. Cancer *44*, 41–45.
Shiloh, Y., and Ziv, Y. (2013). The ATM protein kinase: Regulating the cellular response to genotoxic stress, and more. Nat. Rev. Mol. Cell Biol. *14*, 197–210.
Shoenbill, K., Fost, N., Tachinardi, U., and Mendonca, E.A. (2014). Genetic data and electronic health records: a discussion of ethical, logistical and technological considerations. J. Am. Med. Inform. Assoc. *21*, 171–180.
Short, S.E., Yang, Y.C., and Jenkins, T.M. (2013). Sex, gender, genetics, and health. Am. J. Public Health *103*, Suppl 1, S93-101.
Sniekers, S., Stringer, S., Watanabe, K., Jansen, P.R., I Coleman, J.R., Krapohl, E., Taskesen, E., Hammerschlag, A.R., Okbay, A., Zabaneh, D., et al. (2017). Genome-wide association meta-analysis of 78,308 individuals identifies new loci and genes influencing human intelligence. Nat. Publ. Gr. *49*, 1107-1112.
Söderlund, K., Stål, O., Skoog, L., Rutqvist, L.E., Nordenskjöld, B., and Askmalm, M.S. (2007). Intact Mre11/Rad50/Nbs1 Complex Predicts Good Response to Radiotherapy in Early Breast Cancer. Int. J. Radiat. Oncol. Biol. Phys. *68*, 50–58.
Sprent, P. (2003). Statistics in medical research. Swiss Med. Wkly. *133*, 522–529.
Stratton, M.R., Campbell, P.J., and Andrew F, P. (2009). The cancer genome. Nature *458*, 719–724.
Suresh, K., and Chandrashekara, S. (2012). Sample size estimation and power analysis for clinical research studies. J. Hum. Reprod. Sci. *5*, 7.
Süt, N. (2014). Study designs in medicine. Balkan Med. J. *31*, 273–277.
Switula, D. (2000). Principles of good clinical practice (GCP) in clinical research. Sci. Eng. Ethics *6*, 71–77.
Thies, S., and Langer, R. (2013). Tumor regression grading of gastrointestinal carcinomas after neoadjuvant treatment. Front Oncol. *3*, 262.
Uhlen, M., Zhang, C., Lee, S., Sjöstedt, E., Fagerberg, L., Bidkhori, G., Benfeitas, R., Arif, M., Liu, Z., Edfors, F., et al. (2017). A pathology atlas of the human cancer transcriptome. Science. *357*, eaan2507.

Uhrhammer, N., Bay, J., Gosse-Brun, S., Kwiatkowski, F., Rio, P., Daver, A., and Bignon, Y. (2000). Allelic imbalance at NBS1 is frequent in both proximal and distal colorectal carcinoma. Oncol Rep *7*, 427–431.
van 't Veer, L.J., Dai, H., van de Vijver, M.J., He, Y.D., Hart, A.A.M., Mao, M., Peterse, H.L., van der Kooy, K., Marton, M.J., Witteveen, A.T., et al. (2002). Gene expression profiling predicts clinical outcome of breast cancer. Nature *415*, 530–536.
Varon, R., Seemanova, E., and Chrzanowska, K., et al. (2000). Clinical ascertainment of Nijmegen breakage syndrome (NBS) and prevalence of the major mutation, 657del5, in three Slav populations. Eur J Hum Genet. *8*, 900–902.
Vilar, E., Scaltriti, M., Balmaña, J., Saura, C., Guzman, M., Arribas, J., Baselga, J., and Tabernero, J. (2008). Microsatellite instability due to hMLH1 deficiency is associated with increased cytotoxicity to irinotecan in human colorectal cancer cell lines. Br. J. Cancer *99*, 1607–1612.
Vilar, E., Bartnik, C.M., Stenzel, S.L., Raskin, L., Ahn, J., Moreno, V., Mukherjee, B., Iniesta, M.D., Morgan, M.A., Rennert, G., et al. (2011). MRE11 deficiency increases sensitivity to poly(ADP-ribose) polymerase inhibition in microsatellite unstable colorectal cancers. Cancer Res. *71*, 2632–2642.
Wang, M.-J., Ping, J., Li, Y., Holmqvist, A., Adell, G., Arbman, G., Zhang, H., Zhou, Z.-G., and Sun, X.-F. (2015). Prognostic Significance and Molecular Features of Colorectal Mucinous Adenocarcinomas: A Strobe-Compliant Study. Medicine (Baltimore). *94*, e2350.
Warcoin, M., Lespinasse, J., Despouy, G., D'Enghien, C.D., Laugé, A., Portnoï, M.F., Christin-Maitre, S., Stoppa-Lyonnet, D., and Stern, M.H. (2009). Fertility defects revealing germline biallelic nonsense NBN mutations. Hum. Mutat. *30*, 424–430.
Xing, J., Wu, X., Vaporciyan, A.A., Spitz, M.R., and Gu, J. (2008). Prognostic significance of ataxia-telangiectasia mutated, DNA-dependent protein kinase catalytic subunit, and Ku heterodimeric regulatory complex 86-kD subunit expression in patients with nonsmall cell lung cancer. Cancer *112*, 2756–2764.
Yamashita-Kashima, Y., Shu, S., Yorozu, K., Hashizume, K., Moriya, Y., Fujimoto-Ouchi, K., and Harada, N. (2014). Importance of formalin fixing conditions for HER2 testing in gastric cancer: immunohistochemical staining and fluorescence in situ hybridization. Gastric Cancer *17*, 638–647.
Zhang, T., Penicud, K., Bruhn, C., Loizou, J.I., Kanu, N., Wang, Z.Q., and Behrens, A. (2012). Competition between NBS1 and ATMIN Controls ATM Signaling Pathway Choice. Cell Rep. *2*, 1498–1504.

CHAPTER 8
Future Directions

8.1 Future directions

The collection of accurate patient data in a timely manner represents a major hurdle to treating complex human diseases. The missing patient data is due to a number of reasons, including poor public understanding of disease risk, a lack of timely physician consultation, and a fear of death leading to half-truths (Vrinten et al., 2016), as well as poor professional healthcare protocols (for example, where notes are not detailed and useful to other professionals (Stefanacci and Riddle, 2016)), and poor patient follow-up after diagnosis (Dugdale et al., 1999). In an increasingly global world, where patients regularly switch doctors and domiciles, maintaining an accurate and searchable medical history is of paramount importance for understanding disease and the effectiveness of treatments.

An incomplete clinical patient picture leaves one guessing as to the factors associated with disease progression, and future computerised, online medical records, where patient history can be easily referenced and related to relevant entries (such as an individual's genome sequence), will help stratify patients for post- and ad-hoc research analyses (Bowman, 2013; Walport and Brest, 2011). Furthermore, the patient is not out with his/her environment (take tobacco and lung cancer, for example), and as is seen with regards to how seasons (Dopico et al., 2015), shift-work (Wang et al., 2011), jet-lag (Filipski et al., 2004), the microbiome (Cho and Blaser, 2012), diet (Carrera-Bastos et al., 2011), and artificial light (Chepesiuk, 2009), influence human physiology and cancer risk, detailed metrics of patient exposures to different environmental parameters will become increasingly important for building the complete clinical picture.

Aside from building a more complete clinical history for each patient, more patient samples are needed to understand disease. For example, colorectal cancer and rectal cancer should not be considered the same disease any longer, given the evidence discussed, and therefore, the need to continue collecting as many patient samples as possible to further refine disease aetiology, mechanism, treatment and outcome cannot be overlooked (Braun et al., 2014). Only thorough the collection of geographically- and ethnically-unique populations will diseases like cancer be better understood. It may seem counterintuitive to think that as we are increasingly able to profile individual health, that large, diverse cohorts are needed, but it is through such windows where commonality can be seen (many data points require many samples for robust associations to be discovered; multiple testing (Sprent, 2003)), human physiology better understood, and individuals be placed into a more-refined clinical context. This will be further helped by the collection of longitudinal samples from the same individual (perhaps using modern laparoscopic techniques), allowing disease development and progression to be monitored in human subjects (Trastulli et al., 2012), not only animal models; which display key differences to humans in cancer physiology (Shanks et al., 2009).

Such collections of tumours and healthy tissues from different patient groups and healthy volunteers will serve as valuable resources for genetics and environment to be associated with biological, molecular features of the tumour. For example, expanding cancer transcriptome and genome sequencing, proteomics, and metabolomics will allow for the finer molecular details

of tumours to be revealed (Uhlen et al., 2017; Weinstein et al., 2013). Indeed, the recent combination of high-throughput proteomics and gene expression microarray profiling from colorectal cancers led to the identification of putative cancer biomarkers (S100 calcium-binding protein A9, annexin A3, nicotinamide phosphoribosyltransferase, carboxylesterase 2 and calcium activated chloride channel A1) detectable *in situ* and the serum (Yu et al., 2016).

Such molecular interrogations will be helped by advances in ribosome profiling (where only mRNA transcripts being actively translated on the ribosome are sequence; reducing background noise considerably) (Ingolia et al., 2009), mass cytometry (which combines flow cytometry with mass spectrometry, will allow blood cell tumours, especially, to be analysed in a multi-parametric manner at the single-cell level, helping define cell-surface phenotypes, for example which can be used as prognostic markers, or sorted by FACS for detailed, downstream molecular analyses) (Bendall et al., 2011; Dempsey, 2017), and metabolomics based on the latest low-molecular weight mass spectrometry methods (which will allow small molecular metabolite signatures of tumours to be defined) (Beger, 2013).

All of these advances do not consider the advancement of pharmaceutical science, which is increasingly entering an era exploiting biological reagents (such as humanised monoclonal antibodies and autologous engineered T lymphocytes to destroy cancer cells), as well as human genome editing approaches to cure disease (Cartron et al., 2002; Gill and June, 2015; Jarboe et al., 2014; Paquet et al., 2016). A future in which deleterious variants can be replaced in your genome by a normal, wild-type allele may not be so distant, and already human trials involving CRISPR to replace highly-penetrant lesions in human embryos and patient cells have begun (Su et al., 2016; Tang et al., 2017); which add to older genomic methods able to correct lesions in specific cell types, and grow engineered autologous cells *in vitro* for later transplant back to the host (Hirsch et al., 2017).

The first job of a physician is to: identify those patients most at risk of disease; help such people mitigate those risks; identify those in need; medically treat patients as individuals; and prognostic for them in context of their disease. A future in which vast amounts of patient data and predictive algorithms facilitate our mission will be warmly welcomed.

8.2 DDR proteins

With regards to extending the MRN complex and ATM studies, future studies should involve a larger number of patients, and those with concurrent microsatellite instability. This may be sourced from additional collections of other Australian institutions.

Future studies should also be performed in the context of the tumour genome sequence, to further refine the associations, and attempt to expand upon the predictive value of the biomarkers here presented in rectal cancer.

As has been shown in this thesis, combinatorial marker panels can be highly informative to disease progression and treatment response, and of further interest will be detailing how

different members of the same pathway behave in relation to one another and in respect to their gene sequence, and this should be further explored. Along these lines, immune-fluorescence approaches, where multiple tumour molecules/antigens can be analysed simultaneously in the same sample (as recently shown when refining the phenotype of breast cancer stem cells (Balic et al., 2011)), represent an exciting possibility; allowing one to investigate the relative contribution of different MRN complex members to the on-going pathology in the lesion. For example, it the levels of ATM are decreased, how do the levels of RAD50 change, and so on.

The use of statistical models to describe patient cohorts and associations will be central future biomarker studies in rectal cancer where advanced mathematics will be required to tease apart biological interactions.

Thus, with these things in mind, future rectal cancer research will involve investigation into the mechanisms underpinning rectal cancer development and the definition of predictive markers of treatment and disease outcome.

8.3 References

Balic, M., Rapp, N., Stanzer, S., Lin, H., Strutz, J., Szkandera, J., Daidone, M.G., Samonigg, H., Cote, R.J., and Dandachi, N. (2011). Novel immunofluorescence protocol for multimarker assessment of putative disseminating breast cancer stem cells. Appl. Immunohistochem. Mol. Morphol. *19*, 33–40.

Beger, R. (2013). A Review of Applications of Metabolomics in Cancer. Metabolites *3*, 552–574.

Bendall, S.C., Simonds, E.F., Qiu, P., Amir, E.D., Krutzik, P.O., Finck, R., Bruggner, R. V, Melamed, R., Trejo, A., Ornatsky, O.I., et al. (2011). Single-cell mass cytometry of differential immune and drug responses across a human hematopoietic continuum. Science *332*, 687–696.

Bowman, S. (2013). Impact of electronic health record systems on information integrity: quality and safety implications. Perspect. Heal. Inf. Manag. *10*, 1c.

Braun, K.L., Tsark, J.U., Powers, A., Croom, K., Kim, R., Gachupin, F.C., and Morris, P. (2014). Cancer Patient Perceptions about Biobanking and Preferred Timing of Consent. Biopreserv. Biobank. *12*, 106–112.

Carrera-Bastos, P., Fontes, M., O'Keefe, J.H., Lindeberg, S., and Cordain. L. (2011). The western diet and lifestyle and diseases of civilization. Res. Reports Clin. Cardiol. *2*, 15-35.

Cartron, G., Dacheux, L., Salles, G., Solal-Celigny, P., Bardos, P., Colombat, P., and Watier, H. (2002). Therapeutic activity of humanized anti-CD20 monoclonal antibody and polymorphism in IgG Fc receptor FcγrIIIa gene. Blood *99*, 754–758.

Chepesiuk, R. (2009). Missing the dark: Health effects of light pollution. Environ. Health Perspect. *117*, A20-A27.

Cho, I., and Blaser, M.J. (2012). The human microbiome: at the interface of health and disease. Nat. Rev. Genet. *13*, 260–270.

Dempsey, L.A. (2017). CyTOF analysis of anti-tumor responses. Nat. Immunol. *18*, 254–254.

Dopico, X.C., Evangelou, M., Ferreira, R.C., Guo, H., Pekalski, M.L., Smyth, D.J., Cooper, N., Burren, O.S., Fulford, A.J., Hennig, B.J., et al. (2015). Widespread seasonal gene expression reveals annual differences in human immunity and physiology. Nat. Commun. *6*, 7000.

Dugdale, D.C., Epstein, R., and Pantilat, S.Z. (1999). Time and the patient-physician relationship. J. Gen. Intern. Med. *14 Suppl 1*, S34-40.

Filipski, E., Delaunay, F., King, V.M., Wu, M.W., Claustrat, B., Gréchez-Cassiau, A., Guettier, C., Hastings, M.H., and Francis, L. (2004). Effects of chronic jet lag on tumor progression in mice. Cancer Res. *64*, 7879–7885.

Gill, S., and June, C.H. (2015). Going viral: chimeric antigen receptor T-cell therapy for hematological malignancies. Immunol. Rev. *263*, 68–89.

Hirsch, T., Rothoeft, T., Teig, N., Bauer, J.W., Pellegrini, G., De Rosa, L., Scaglione, D., Reichelt, J., Klausegger, A., Kneisz, D., et al. (2017). Regeneration of the entire human epidermis using transgenic stem cells. Nature *551*, 327–332.

Ingolia, N.T., Ghaemmaghami, S., Newman, J.R.S., and Weissman, J.S. (2009). Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. Science *324*, 218–223.

Jarboe, J., Gupta, A., and Saif, W. (2014). Therapeutic human monoclonal antibodies against

cancer. Methods Mol. Biol. *1060*, 61–77.
Paquet, D., Kwart, D., Chen, A., Sproul, A., Jacob, S., Teo, S., Olsen, K.M., Gregg, A., Noggle, S., and Tessier-Lavigne, M. (2016). Efficient introduction of specific homozygous and heterozygous mutations using CRISPR/Cas9. Nature *533*, 1–18.
Shanks, N., Greek, R., and Greek, J. (2009). Are animal models predictive for humans? Philos. Ethics, Humanit. Med. *4*, 2.
Sprent, P. (2003). Statistics in medical research. Swiss Med. Wkly. *133*, 522–529.
Stefanacci, R.G., and Riddle, A. (2016). Preventing medication errors. Geriatr. Nurs. (Minneap). *37*, 307–310.
Su, S., Hu, B., Shao, J., Shen, B., Du, J., Du, Y., Zhou, J., Yu, L., Zhang, L., Chen, F., et al. (2016). CRISPR-Cas9 mediated efficient PD-1 disruption on human primary T cells from cancer patients. Sci. Rep. *6*, 20070.
Tang, L., Zeng, Y., Du, H., Gong, M., Peng, J., Zhang, B., Lei, M., Zhao, F., Wang, W., Li, X., et al. (2017). CRISPR/Cas9-mediated gene editing in human zygotes using Cas9 protein. Mol. Genet. Genomics *292*, 525–533.
Trastulli, S., Cirocchi, R., Listorti, C., Cavaliere, D., Avenia, N., Gullà, N., Giustozzi, G., Sciannameo, F., Noya, G., and Boselli, C. (2012). Laparoscopic vs open resection for rectal cancer: A meta-analysis of randomized clinical trials. Color. Dis. *14*, e277-296.
Uhlen, M., Zhang, C., Lee, S., Sjöstedt, E., Fagerberg, L., Bidkhori, G., Benfeitas, R., Arif, M., Liu, Z., Edfors, F., et al. (2017). A pathology atlas of the human cancer transcriptome. Science (80-.). *357*, eaan2507.
Vrinten, C., Wardle, J., and Marlow, L.A. (2016). Cancer fear and fatalism among ethnic minority women in the United Kingdom. Br. J. Cancer *114*, 597–604.
Walport, M., and Brest, P. (2011). Sharing research data to improve public health. Lancet *377*, 537–539.
Wang, X.-S., Armstrong, M.E.G., Cairns, B.J., Key, T.J., Travis, R.C., and Nicholson, P.J. (2011). Shift work and chronic disease: the epidemiological evidence. Occup. Med. (Chic. Ill). *61*, 443–4; author reply 444.
Weinstein, J.N., Collisson, E.A., Mills, G.B., Shaw, K.R.M., Ozenberger, B.A., Ellrott, K., Shmulevich, I., Sander, C., and Stuart, J.M. (2013). The Cancer Genome Atlas Pan-Cancer analysis project. Nat. Genet. *45*, 1113–1120.
Yu, J., Li, X., Zhong, C., Li, D., Zhai, X., and Hu, W. (2016). High-throughput proteomics integrated with gene microarray for discovery of colorectal cancer potential biomarkers. Oncotarget *7*, 75279–75292.