



The role of SPRTN protease in colorectal cancer chemotherapy

DPhil Thesis

Supervisors:

Prof Kristijan Ramadan

Prof Tim Maughan

Dr Cristiano Peron

St Peter's College

Department of Oncology

University of Oxford

Trinity Term 2023

Contents

Acknowledgements 6

Declaration 7

Abstract 8

List of Figures 9

List of Tables 11

List of Supplementary Tables 12

List of Abbreviations 13

Chapter 1: Introduction 17

The clinical stages of CRC (TNM staging) and their management 20

Genomic instability in CRC 26

Chromosomal Instability Pathway 26

Microsatellite instability (MSI) 37

CpG Island Methylator Phenotype (CIMP) 39

The molecular classification of CRC (CMS, CRIS) 41

Consensus Molecular Subtype (CMS) 42

Colorectal Intrinsic Subtypes (CRIS) 44

Mechanism of action of cytotoxic chemotherapy in CRC (5-Fluorouracil, Irinotecan, Oxaliplatin) 46

5-Fluorouracil 47

Oxaliplatin 49

Irinotecan 52

DNA repair mechanisms of chemotherapy-induced DNA damage in CRC (5-Fluorouracil, Irinotecan, Oxaliplatin) 54

5-Fluorouracil 54

Oxaliplatin 55

Irinotecan 57
DNA-protein crosslinks repair and the role of SPRTN protease 59
DNA-protein crosslinks: molecular structure and repair pathways 59
SPRTN, a DNA-dependent metalloprotease involved in replication-coupled DPC proteolysis 65
Alternative DPC proteases 67
Main hypothesis and aims 69
Chapter 2: Material and methods 70
Cell culture 70
siRNA 71
Colony forming assays 72
Western blots 74
RT-qPCR 76
RNA extraction 76
RNA conversion to cDNA 77
RT-qPCR 77
Auxin-inducible degron 80
Genome editing (CRISPR-Cas9) 82
Fluorescence-activated cell sorting (FACS) and clonal growth 83
Western blot screening 83
PCR screening of selected clones 83
Confirmation of SPRTN degradation by western blot 86
Cell cycle analysis 87
Staining and preparations of samples 87
Flow cytometric analysis 88

Quantification of the results 88
Immunofluorescence 91
Processing of adherent cells on coverslips 91
Imaging 93
RNA-seq 94
RNA extraction and quality assessment 96
RNA-seq 96
Immunohistochemistry and staining assessment 99
Immunohistochemical staining 99
Digital pathology (HALO) 100
Survival analyses 102
Chapter 3: Results 104
3.1 SPRTN role in mediating the sensitivity to cytotoxic chemotherapy in CRC 104
Summary 122
3.2 Development and characterization of a SPRTN-AID system in HCT116 cells 123
Summary 137
3.3 Whole genome transcriptomic analysis of SPRTN-depleted cells (RNA-seq) 138
Summary 149
3.4: SPRTN role as a biomarker in CRC patients 150
Exploratory analysis 156
Oxford Cohort 159
Final analysis (Oxford cohort + Leeds cohort) 166
Summary 172
Chapter 4: Discussion 174
The role of SPRTN in mediating the survival of CRC cells treated with chemotherapy 177

The consequences of SPRTN depletion on the transcriptome of CRC cells 181

The role SPRTN as predictive and prognostic biomarker in CRC patients 188

SPRTN as a prognostic biomarker for FOLFIRI chemotherapy 188

SPRTN as a predictive biomarker in FOLFIRI vs FOLFOX 189

Chapter 5: Future work 191

Supplementary tables 197

References 231

Acknowledgements

Firstly, I would like to thank both my DPhil supervisors, Professor Kristijan Ramadan and Professor Tim Maughan, for their unwavering support and for the many meetings we had together during the last four years. During each one of them, I have appreciated your valuable inputs and your mentorship, which allowed me to become a better scientist and a better doctor. Kristijan, your unmatched passion for science has taught me how to approach every scientific question: with energy, hard-work and excitement. Tim, your calm and steady guidance allowed me to never lose sight of what is really important: our patients, and the attempt to improve their lives. Without you both, I am not sure I would have been able to reach the end of this long journey. I will always be grateful for the way you have encouraged me.

Of course, I also would like to thank all the lab members I had the pleasure to work with. This includes Annamaria, Wei, Abhay, Ignacio, Sam, Shudong, John, Pauline, Gwennie, Ftoon, Susan, Enric and Sylvana. In particular, I feel the need to mention Pauline and Ftoon, who started their DPhils along with me. I have enjoyed the opportunity to share this experience with you, and I am grateful for the many memories we have created together.

Working in the Department of Oncology at the University of Oxford has been a pleasure, especially because of all the people I had the opportunity to meet in this place. Between all the principal investigators and senior fellows, this particularly includes Mark, Eric, Vincenzo, Ester and Simon. Between all the students and postdocs, I enjoyed every moment I have shared with Hannah, Cagla, Hala, Shahd, Claudia and Kay. And, of course, my favorite extroverted weirdo: Tiffany.

Finally, I want to thank my family. Thank you for supporting me all these years, none of this would have been possible without you.

Declaration

This dissertation is submitted for the degree of Doctor of Philosophy in Oncology. I hereby declare that the work presented in this thesis is my own except where explicitly otherwise stated. I declare that this work has not been submitted for any other degree at this or any other university.

Abstract

Colorectal cancer (CRC) currently represents one of the most commonly diagnosed tumors, accounting for 10% of all cancer diagnoses and cancer-related deaths. The first-line treatment of metastatic colorectal cancer (mCRC) is frequently based on combination chemotherapeutic regimens, where 5-fluorouracil is either associated with irinotecan (FOLFIRI) or oxaliplatin (FOLFOX). Both these regiments present similar level of efficacy, and both irinotecan and oxaliplatin have been reported as drugs with the potential to induce relevant amounts of DNA-protein crosslinks (DPCs), a heterogeneous form of DNA lesion in which proteins end up being covalently bound to the DNA. At the moment we are currently lacking reliable methods to predict which patients are more likely to respond to FOLFIRI or FOLFOX, which would allow us to choose the most appropriate therapy for each patient. In order to address this need, the role of SPRTN in colorectal cancer chemotherapy is investigated for the first time in this thesis.

SPRTN was recently identified as a metalloprotease involved in the repair of DPCs. To address SPRTN potential role in the repair of DPCs induced by irinotecan and oxaliplatin, two main approaches have been attempted in this thesis. First, a preclinical study with three different colorectal cancer cell lines was performed to address if SPRTN depletion could make cells more sensitive to the aforementioned drugs. However, despite presenting an increased amount of DNA damage, CRC cells appeared to survive the deleterious effects of SPRTN depletion, thanks to the expression of alternative DPC proteases and the upregulation of anti-apoptotic genes. Secondly, clinical mCRC samples obtained from the MRC FOCUS trial were analyzed to determine if SPRTN overexpression measured by immunohistochemistry could represent a useful predictive biomarker to determine the efficacy of FOLFIRI vs FOLFOX chemotherapy. In line with the preclinical data, however, SPRTN was not confirmed to represent a valuable biomarker in CRC chemotherapy.

List of Figures

Figure 1. Colorectal cancer carcinogenesis. 19
Figure 2. Application of the TNM system to colorectal cancer. 21
Figure 3. Clinical stages of colorectal cancer based on the TNM classification. 21
Figure 4. Clinical management of CRC patients [34, 35]. 24
Figure 5. Systemic treatment of CRC patients. 25
Figure 6. The chromosomal instability (CIN) pathway. 27
Figure 7. The WNT pathway. 30
Figure 8. The RAS pathway. 33
Figure 9. The p53 signaling pathway. 35
Figure 10. The microsatellite instability (MSI) pathway. 38
Figure 11. GpG island methylator phenotype (CIMP) 40
Figure 12. CRC subtypes according to CMS classification. 43
Figure 13. CRC subtypes according to CRIS classification 45
Figure 14. 5-Fluorouracil mechanisms of action. 48
Figure 15. Oxaliplatin mechanisms of action. 51
Figure 16. Irinotecan mechanisms of action. 53
Figure 17. Main mechanisms of DPC repair. 64
Figure 18. SPRTN and its role in DPC repair. 66
Figure 19. Colony forming assays. 73
Figure 20. SPRTN C-terminal insertion in HCT116 osTIR1 cells. 81
Figure 21. Auxin-inducible degron insertion workflow. 81
Figure 22. sgRNA plasmid and donor DNA for HCT116 SPRTN-AID cells. 82
Figure 23. DNA fragments obtained from PCR 84
Figure 24. Click-reaction for S-phase cells. 89
Figure 25. Cell cycle analysis. 90
Figure 26. RNA-seq workflow. 95
Figure 27. RNA-seq correlation heatmaps for HCT116 and HeLa cells. 97
Figure 28. PCA analysis across replicates. 98
Figure 29. SPRTN role and alterations in cancer. 107

Figure 30. SPRTN role in mediating the efficacy of 5-Fluorouracil. 110
Figure 31. SPRTN role in mediating the efficacy of Irinotecan. 113
Figure 32. SPRTN role in mediating the efficacy of Camptothecin. 115
Figure 33. SPRTN role in mediating the efficacy of Oxaliplatin. 119
Figure 34. p53 role in SPRTN depleted cells treated with Oxaliplatin. 121
Figure 35. Generation of HCT116 SPRTN-AID cells. 126
Figure 36. SPRTN-depleted cells survival to DPC-inducing drugs. 130
Figure 37. 53BP1 and yH2AX foci in SPRTN-depleted cells. 133
Figure 38. Cell cycle of SPRTN-depleted cells. 136
Figure 39. Gene expression in SPRTN-depleted cells. 143
Figure 40. Volcano plots of DEGs in SPRTN-depleted cells. 145
Figure 41. Kegg pathway analysis of SPRTN-depleted cells. 148
Figure 42. MRC focus trial and SPRTN IHC staining. 153
Figure 43. SPRTN antibodies and their validation for IHC. 155
Figure 44. SPRTN as a prognostic biomarker in FOLFIRI-treated patients (45 patients). 158
Figure 45. Survival analyses of CRC patients based on treatment (Oxford cohort – 400 patients) 163
Figure 46. SPRTN as a predictive biomarker in CRC patients (Oxford cohort – 400 patients). 165
Figure 47. SPRTN as a predictive biomarker in CRC patients (Oxford and Leeds cohorts – 914 patients). 169
Figure 48. Consequences of SPRTN-depletion in different cell lines. 180
Figure 49. SPRTN-depletion and its consequences on the cell cycle. 185
Figure 50. Regulation of apoptosis in SPRTN-depleted cells. 187
Figure 51. Volcano plots of ribosomal DEGs in SPRTN-depleted cells. 196

List of Tables

Table 1. Oligonucleotides used for siRNA depletions. 71

Table 2. Primary antibodies used for western blot. 75

Table 3. Secondary antibodies used for western blot. 75

Table 4. SPRTN primers used for qPCR. 79

Table 5. GAPDH primers used for qPCR. 79

Table 6. Thermocycler settings for qPCR 79

Table 7. Primers used for first PCR. 85

Table 8. Primers used for second PCR. 85

Table 9. PCR reaction components. 85

Table 10. Thermocycler conditions for PCR. 85

Table 11. Click-reaction components. 89

Table 12. Flow cytometry gating strategy. 90

Table 13. Extraction buffer for immunofluorescence. 92

Table 14. Primary antibodies for immunofluorescence. 92

Table 15. Secondary antibodies for immunofluorescence. 92

Table 16. Experimental conditions and replicates used for RNA-seq. 97

Table 17. Primary antibodies for IHC staining. 100

Table 18. Clinical and molecular variables of CRC patients. 103

Table 19. SPRTN-AID sequencing 127

Table 20. Main differences between Oxford and Leeds cohorts. 171

List of Supplementary Tables and Figures

Supplementary table 1. Main characteristic of cell lines. 198

Supplementary table 2. Genome-editing plasmids used for HCT116 SPRTN-AID. 200

Supplementary table 3. List of DEGs in HCT116 DMSO vs HCT116 Auxin. 206

Supplementary table 4. List of DEGs in HeLa WT vs HeLa SPRTN $^{-/-/+}$ (Top 100 over 4575).. 209

Supplementary table 5. List of DEGs present in both HCT116 Auxin and Hela SPRTN $^{-/-/+}$ 212

Supplementary table 6. List of DEGS present in HCT116 Auxin but not in HeLa SPRTN $^{-/-/+}$ .. 216

Supplementary table 7. Clinical and IHC data of 45 CRC patients treated with FOLFIRI (Exploratory analysis). 218

Supplementary table 8. Clinical and IHC data of 400 CRC patients (Oxford Cohort). 228

Supplementary figure 1. SPRTN possible role as a prognostic biomarker in CRC patients (Oxford cohort – 400 patients)……..…230

List of Abbreviations

Abbreviation	Full name
53BP1	P53 Binding Protein 1
5-FU	5-Fluorouracil
ADGRF1	Adhesion G Protein-Coupled Receptor F1
AHRR	Aryl Hydrocarbon Receptor Repressor
AJCC	American Joint Committee on Cancer
ALDH1A3	Aldehyde Dehydrogenase 1 Family Member A3
APC	Adenomatous Polyposis Coli Protein
APC/C	Anaphase promoting complex / cyclosome
APE1	Apurinic/Apyrimidinic Endodeoxyribonuclease 1
ARK	Advance Recovery of K-SDS precipitates
ASCO	American Society of Clinical Oncology
ATCC	American Type Culture Collection
ATM	Ataxia-telangiectasia mutated kinase
ATR	ATM-Rad3-related kinase
AUX/IAA	Auxin / Indole-3-Acetic Acid
Axin	Axis Inhibition Protein
BCL2L2	BCL2 Like 2 gene
BCL-W	BCL2 Like 2 protein
BER	Base Excision Repair
BRAF	V-Raf Murine Sarcoma Viral Oncogene Homolog B
BSA	Bovine serum albumin
Cas9	CRISPR associated protein 9
CASP3	Caspase 3
CASP8	Caspase 8
CASP9	Caspase 9
CDC25	Cell division cycle 25 protein
CDC25B	Cell division cycle 25B protein
CDK1	Cyclin Dependent Kinase 1
CDK1/2	Cyclin Dependent Kinase 1/2
CDK6	Cyclin Dependent Kinase 6
CDKN2A	Cyclin Dependent Kinase Inhibitor 2A
CGIs	CpG islands
CHK1/2	Checkpoint kinase 1/2
CIMP	CpG Island Methylator Phenotype
CIN	Chromosomal Instability
CK1	Casein Kinase 1
CMG Helicase	CDC45-MCM2/7-GINS Helicase
CMS	Consensus Molecular Subtypes
CpG	Cytosine-Guanine dinucleotides
CPT-11	Irinotecan
CR	Complete Response
CRC	Colorectal Cancer
CRCSC	CRC Subtyping Consortium
CRIS	Colorectal Intrinsic Subtypes
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
CYP1A1	Cytochrome P450 Family 1 Subfamily A Member 1
DC	Destruction Complex
DCC	Deleted in Colorectal Cancer protein
DDI1	DNA Damage Inducible 1 Homolog 1
DDR	DNA-damage response
DMEM	Dulbecco's Modified Eagle's Medium
dMMR	Mismatch repair deficient
DMSO	Dimethylsulfoxide
DNA	Deoxyribonucleic acid

DNMT1	DNA Methyltransferase 1
DNMT3A/B	DNA Methyltransferase 3 Alpha / Beta
DPC4	Deleted in pacreatic cancer 4 protein
DPCs	DNA-protein crosslinks
DBSs	Double-strand breaks
dTMP	Deoxythymidine monophosphate
dTTP	Deoxythymidine triphosphate
dUMP	Deoxyuridine monophosphate
DVL	Dishevelled Protein
E2F	Retinoblastoma-Binding Protein 3 (E2F-Like)
ECACC	European Collection of Authenticated Cell Cultures
EdU	5-Ethynyl-2'-deoxyuridine
EGFR	Epidermal growth factor receptor
EMT	Epithelial-mesenchymal transition
ERCC1	ERCC Excision Repair 1 protein
ERK1/2	Extracellular signal-regulated kinase 1/2
FA	Fanconi Anemia
FACS	Fluorescence-activated cell sorting
FAM111A/B	FAM111 Trypsin Like Peptidase A / B
FBS	Fetal Bovine Serum
FdUMP	Fluorodeoxyuridine monophosphate
FdUTP	Fluorodeoxyuridine triphosphate
FFPE	Formalin-fixed, paraffin-embedded
FOLFIRI	5-Fluorouracil + Irinotecan
FOLFOX	5-Fluorouracil + Oxaliplatin
FOLFOXIRI	5-Fluorouracil + Oxaliplatin + Irinotecan
FUTP	Fluorouridine triphosphate
FZD	Frizzled receptor
GCN5 (KAT2A)	Lysine acetyltransferase 2A
GCNA	Germ cell nuclear acidic protein
GRB2	Grow Factor Receptor-bound protein 2
GSK3β	Glycogen Synthase Kinase 3 Beta
HR	Homologous Recombination Repair
HR23B	RAD23 Homolog B
ICLs	Interstrand crosslinks
IGF2	Insulin Like Growth Factor 2
IHC	Immunohistochemistry
KEGG	Kyoto Encyclopedia of Genes and Genomes
KRAS	Kirsten rat sarcoma virus
LOH	Loss of heterozygosity
LRP5/6	LDL Receptor Related Protein 5 / 6
MAD	Median Absolute Deviation
MAPKs	Mitogen-activated protein kinases
mCRC	Metastatic colorectal cancer
MDM2	MDM2 Proto-Oncogene
MEK	MAP kinase kinase
MGMT	Methylguanine-DNA Methyltransferase
MLH	MutL Homolog
MMR	Mismatch repair
MRE11	Meiotic Recombination 11 Homolog 1
MRN complex	Mre11-Rad50-Nbs1 complex
mRNA	Messenger RNA
MSH	MutS Homolog
MSI	Microsatellite instability
MSI-H	Microsatellite instability - High
MSS	Microsatellite stable
MYC	MYC Proto-Oncogene

NBS1 (NBN)	Nijmegen Breakage Syndrome Protein 1
NER	Nucleotide Excision Repair
NHEJ	Non-Homologous End Joining repair
OS	Overall Survival
osTIR1	Oryza sativa Transport Inhibitor Response 1
p53	Tumor protein p53
p97 (VCP)	Valosin-containing Protein
PBS	Phosphate Buffered Saline
PCA	Principal component analysis
PCAF (KAT2B)	Lysine Acetyltransferase 2B
PCNA	Proliferating cell nuclear antigen
PD	Progressive Disease
PD-1	Programmed Cell Death Protein 1
PFS	Progression Free Survival
PI	Propidium iodide
PMS2	PMS1 Homolog 2
PNKP	Polynucleotide kinase 3' phosphatase
Pol β	DNA polymerase beta
Pol δ	DNA polymerase delta
PR	Partial Response
PROTAC	Proteolysis targeting chimera
RADAR	Rapid Approach to DNA-Adduct Recovery
RAS	Rat Sarcoma Virus
RB	Retinoblastoma protein
RJALS	Ruijs-Aalfs syndrome
RNAi	RNA interference
RNA-seq	RNA sequencing
RPA	Replication protein A
RPL11	Ribosomal Protein L11
RPL17	Ribosomal Protein L17
RPL24	Ribosomal Protein L24
RPS16	Ribosomal Protein S16
RR	Response Rate
rRNA	Ribosomal RNA
RT-qPCR	Quantitative reverse transcription polymerase chain reaction
SCLC	Small cell lung cancer
SCNA	Somatic copy number alteration
SD	Stable disease
siRNA	Small interfering RNA
SIVA1	SIVA1 Apoptosis Inducing Factor protein
SN-38	7-ethyl-10-hydroxycamptothecin
SOS	Son of Sevenless protein
SOX4	SRY-Box Transcription Factor 4
SPRTN	SprT-Like N-Terminal Domain protein
SSBs	Single-strand breaks
TBS-T	Tris Buffered Saline with Tween
TCF	T-cell Factor
TCGA	The Cancer Genome Atlas
TDP1	Tyrosyl-DNA phosphodiesterase 1
TEX264	Testis-Expressed Protein 264
TFIIH	Transcription factor II H
TGF-β	Transforming growth factor-beta
THL	Translational Histopathology Laboratory
TIPARP	TCDD-inducible Poly(ADP-ribose) Polymerase
TLS	Translesion synthesis
TNM	Tumor-node-metastasis
TOP1	Topoisomerase 1 protein

TOP1cc	TOP1 cleavage complex
tRNA	Transfer RNA
TS	Thymidylate synthase
UNG	Uracil-DNA glycosylase
VCPIP1	Valosin Containing Protein Interacting protein 1
WSS1	Weak suppressor of SMT3 protein 1
WT	Wild type
XPA	Xeroderma Pigmentosum Group A-complementing protein
XPC	Xeroderma Pigmentosum Group C-complementing protein
XPF (ERCC4)	ERCC Excision Repair 4 protein
XPG (ERCC5)	ERCC Excision Repair 5 protein
β-TrCP	Beta-transducin repeat containing E3 ubiquitin protein ligase
γH2AX	Phosphorilated H2A Histone Family Member X

Chapter 1: Introduction

Almost a quarter of century has passed since Hanahan and Weinberg proposed the idea that cells, as they progress from their normal state to more aggressive and cancerous phenotypes, are characterized by the progressive acquisition of functional capabilities which are common between different cancer types [1]. These originally six shared commonalities, known as the hallmarks of cancer, were then updated in 2011 and 2022 to take into account some of our most recent understanding of cancer development and progression [2, 3]. The hallmarks of cancers are now twelve (Figure 1A), and they provide a conceptual scaffold to recapitulate the complex phenotypes which can be observed in cancer.

Cancers currently represents one of the major burdens on human health and, in 2018, around 18.1 million new cancer cases were registered worldwide, with over 9.6 million cancer deaths [4]. The process of carcinogenesis is largely based on the progressive acquisition of genetic changes in proliferating cells, which influence their ability to keep their own replication under control [5, 6]. These changes can either involve the aberrant activation of proto-oncogenes or the inactivation of tumor suppressor genes, and they can be further divided in somatic or germline mutations [7-9]. This process has been particularly well studied in colorectal cancer (CRC), the focus of this thesis, with the identification of the most common mutations, chromosomal changes and translocations responsible for the development and progression of CRC from normal colorectal epithelium (Figure 1B). This interest is justified by the fact that CRC alone accounts for almost 10% of all new cancer diagnoses and cancer-related death, and it is currently one of the most common type of cancers [10]. Furthermore, despite being generally considered a single clinical entity, it is important to differentiate CRC in colon or rectal cancers, according to the anatomical site from which they originally develop. This is because rectal cancers, which represent one third of all diagnosed CRC, are characterized by higher risk of local invasion due to the presence of the mesorectal fascia,

a lax layer of connective tissue surrounding the rectum. As a result, rectal cancers generally require neoadjuvant radiotherapy or chemotherapy before surgical treatment [11]. The main differences between colon and rectal cancer are summarized in Figure 1C.

Colorectal cancer is the second most frequent cancer-related cause of death in developed countries, and around 50% of patients already have metastases at the time of diagnosis or will subsequently develop them during the clinical follow up of the disease [12, 13]. While the median overall survival of patients with metastatic colorectal cancer (mCRC) has improved in the past 20 years, and it is now around 30 months, the metastatic disease remains far from curable [14].

As for many other cancers, CRC can be classified either according to its molecular or clinical features [15-17]. Both classifications will be described since they represent and important reference to understand the challenges characterizing the treatment of CRC patients, and the questions that this thesis desire to address.

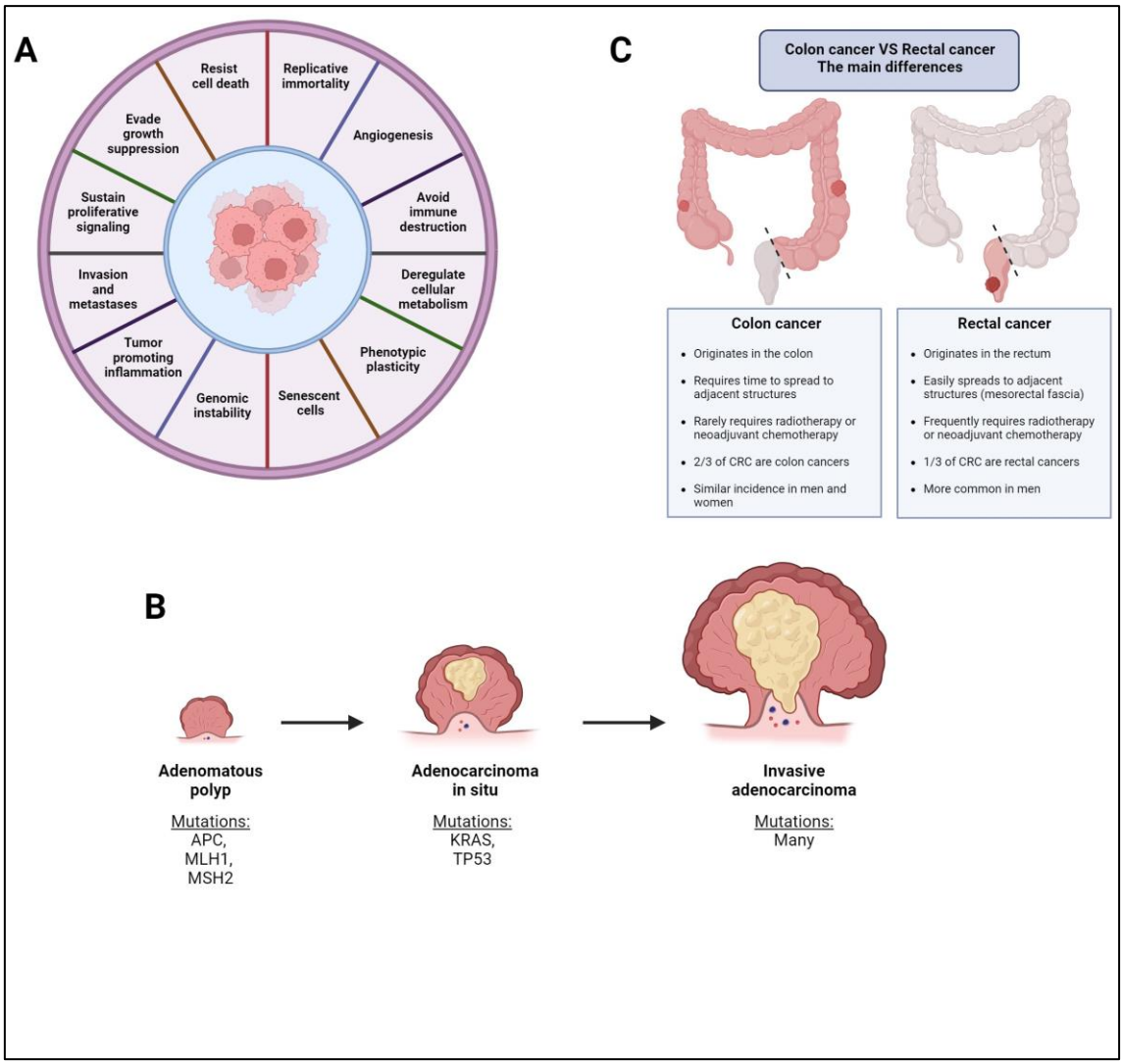


Figure 1. Colorectal cancer carcinogenesis.

(A) Hallmarks of cancer. (B) Carcinogenesis of CRC, from adenomatous polyps to invasive adenocarcinomas. (C) The main differences between colon and rectal cancers.

The clinical stages of CRC (TNM staging) and their management

The clinical classification of cancers is generally based on what is known as the tumor-node-metastasis system (TNM staging). This system is used even before the initiation of therapy, and it is particularly useful to facilitate the exchange of information between clinicians and the establishment of appropriate therapeutic decisions. It is based on the evaluation, through physical examinations, laboratory finding and imaging studies, of three main parameters:

- Tumor size and local growth (T)
- Extent of lymph nodes involvement (N)
- Presence of distant metastases (M)

The TNM system can be further divided in clinical (cTNM) or pathological (pTNM) classification, according if this has been determined before or after surgical exploration and the examination of histological samples collected during the surgery [18, 19]. Thanks to the TNM staging every parameter (T, N, M) can be properly classified, and they can be put together to determine the clinical stage of the cancer, which generally represents the most important parameter in determining the clinical management of the disease. The latest update regarding the application of the TNM system to the staging of colorectal cancer has been published in 2017, in the 8th edition of the American Joint Committee on Cancer (AJCC) staging manual [19, 20]. This classification has been summarized in Figure 2 and Figure 3.

TNM staging of colorectal cancer

Tumor	Characteristics
Tis	Carcinoma in situ
T1	Tumor invades the submucosa
T2	Tumor invades muscularis propria
T3	Tumor invades subserosa
T4a	Tumor invades the nearby peritoneum
T4b	Tumor invades nearby organs
Nodes	**Characteristics**
N0	No nodal metastases
N1a	Metastasis in 1 pericolic or perirectal lymph nodes
N1b	Metastases in 2-3 pericolic or perirectal lymph nodes
N1c	No nodal metastasis, but cancer cells can be found in nearby tissue
N2a	Metastases in 4-6 pericolic or perirectal lymph nodes
N2b	Metastases in more than 7 pericolic or perirectal lymph nodes
Metastases	**Characteristics**
M0	No distat metastases
M1a	Metastasis in 1 distant organ
M1b	Metastases in 2 or more distant organs
M1c	Metastases can be found in the peritoneum

Figure 2. Application of the TNM system to colorectal cancer.

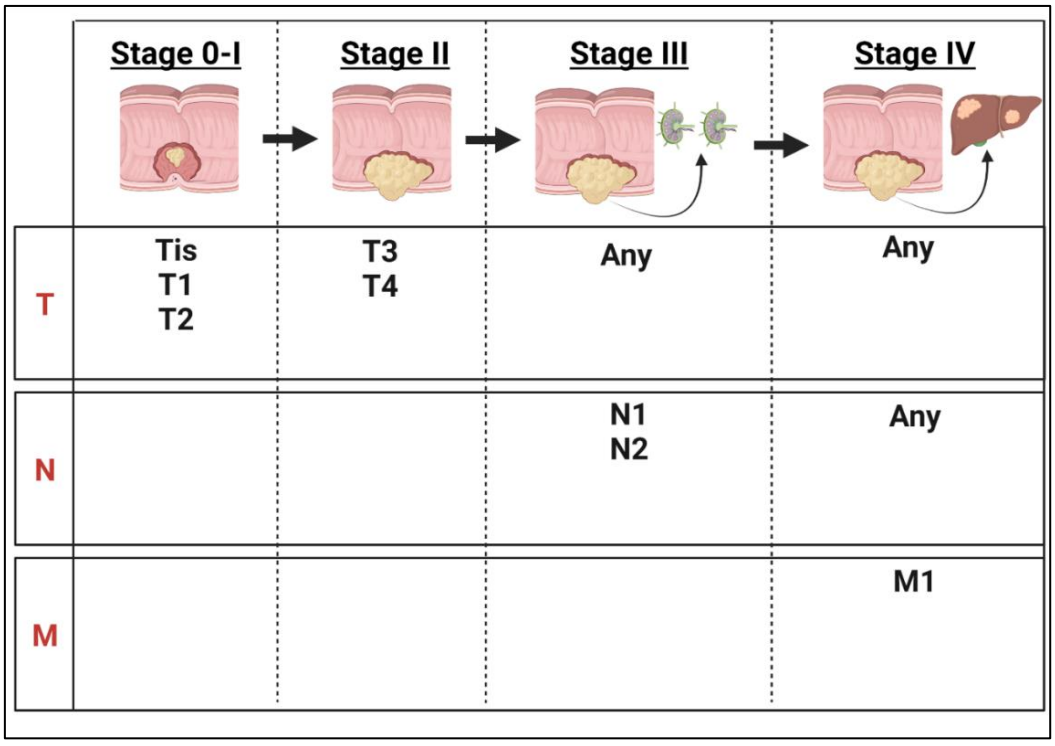


Figure 3. Clinical stages of colorectal cancer based on the TNM classification.

Understanding how the clinical stage of colorectal cancer is determined is particularly important, as the stage represents the strongest predictor of survival for these patients and it affects most therapeutic decisions [21]. A summary of the clinical management of patients affected by CRC, according to their clinical stage, is reported in Figure 4.

Surgery remains the primary treatment of colorectal cancer, as it represents the only one which can be performed with curative intent, but only 75% of newly diagnosed patients can benefit for it. This is because surgery does not generally represent a viable option in the metastatic setting (Stage IV). Moreover, around 30% of Stage II and 55% of patients affected by stage III CRC will develop a recurrence within 5 years after the initial surgery [22]. Understanding the risk of recurrence after surgery also allow us to understand why, in the end, around half of patients affected by CRC will reach a stage IV and will need to be treated with chemotherapy during the course of their disease [12]. Accordingly, a brief description of the main systemic treatments available for metastatic CRC will now be provided. Their molecular mechanisms of action, instead, will be reported later in this chapter.

The most common chemotherapeutic drugs used to treat mCRC patients are 5-fluorouracil (5-FU, generally administered along with folinic acid), irinotecan and oxaliplatin. In particular, combination chemotherapy regimens are now the gold standard and 5-FU can be associated with irinotecan (FOLFIRI) or oxaliplatin (FOLFOX) [23, 24]. Since their introduction in the clinical practice over 20 years ago, numerous studies have demonstrated the equivalence of FOLFOX and FOLFIRI regimens in terms of efficacy [25, 26], and both this regiments can be considered as the first line treatment. Moreover, the combination of 5-FU with Irinotecan plus Oxaliplatin (FOLFOXIRI) has also been evaluated and a meta-analysis of different phase III trials was able to demonstrate an improved progression free survival and overall survival, at the cost of an higher toxicity profile [27].

In recent years, our ability to treat patients with mCRC has also been enriched by the introduction of three different monoclonal antibodies: bevacizumab (anti-VEGF), cetuximab (anti-EGFR) and panitumumab (anti-EGFR) [28-30]. While the addition of monoclonal antibodies to the chemotherapeutic regiments previously described has been able to slightly increase the efficacy of chemotherapy, our project will focus on 5-FU, irinotecan and oxaliplatin as these agents are the ones manly involved in the damage of DNA. Furthermore, it is important to report the results of the recent KEYNOTE-177 trial, published in 2022, which investigated the administration of immunotherapy (pembrolizumab) in patients diagnosed with microsatellite instability-high or mismatch repair-deficient (MSI-H / dMMR) metastatic colorectal cancer [31]. Pembrolizumab determined a longer progression free survival when compared to standard chemotherapy, along with less severe adverse effects, and became a valid first-line treatment in this subset of patients. While these results are certainly encouraging, it is also important to underline that only 4-5% of metastatic CRC patients present MSI-H / dMMR tumors [32].

The possible role of immunotherapy has also been recently investigated in locally advanced dMMR rectal cancer. In an exciting Phase II clinical trial published in 2022, 12 patients with locally advanced dMMR rectal cancer were treated with neoadjuvant dostarlimab, an anti-Programmed Cell Death Protein 1 (anti-PD-1) monoclonal antibody, before being considered for chemoradiotherapy or surgery. Surprisingly, all 12 patients showed a complete response without requiring chemoradiotherapy or surgery, which was sustained for at least 6 months of follow up [33].

A summary of the current guidelines from the systemic treatment of CRC, as recommended by the American Society of Clinical Oncology (ASCO), is reported in Figure 5 [34, 35].

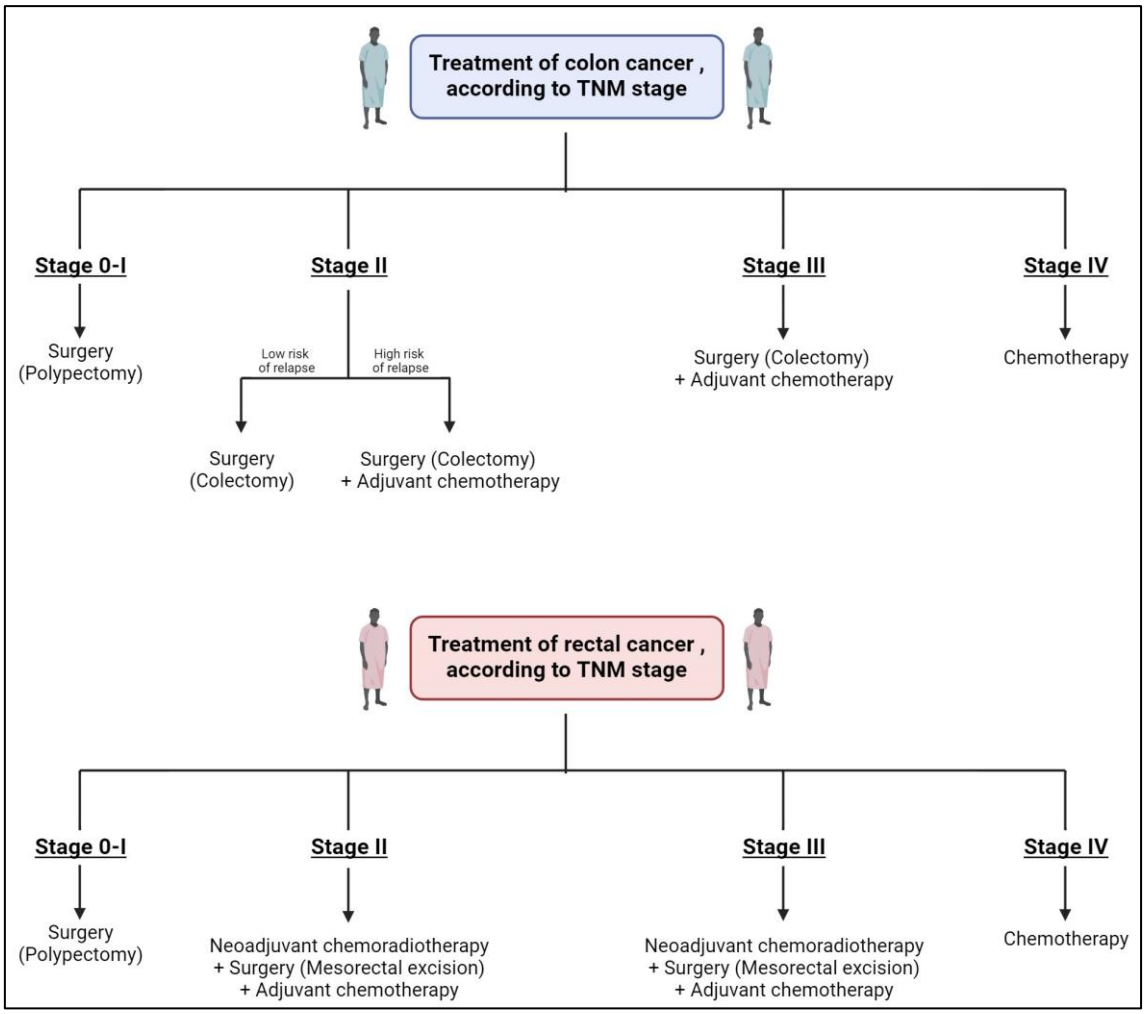


Figure 4. Clinical management of CRC patients [34, 35].

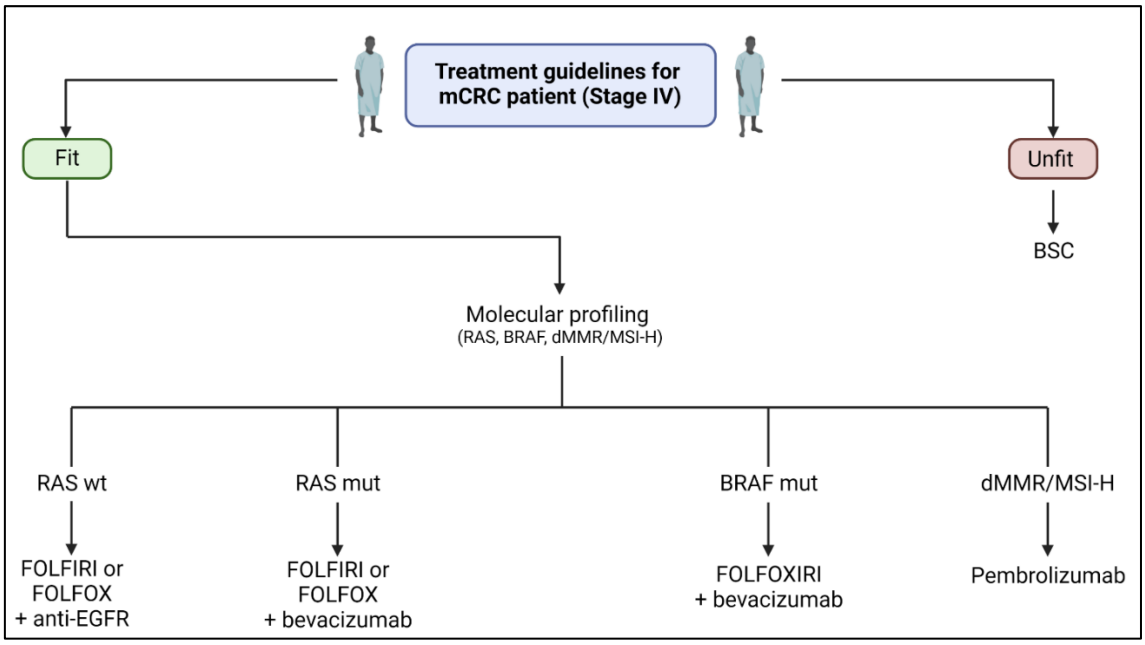


Figure 5. Systemic treatment of CRC patients.

The systemic treatment of CRC patients depend on the molecular characteristics of the tumour [34, 35]. Best supportive care is indicated as BSC.

Genomic instability in CRC

Before being able to address the molecular classification of CRC, it is important to recognize which genetic alterations are generally responsible for the development of CRC and their molecular implications. Three main pathways are generally considered to be responsible for CRC development and progression [36]:

- Chromosomal Instability (CIN)
- Microsatellite Instability (MSI)
- CpG Island Methylator Phenotype (CIMP)

Chromosomal Instability Pathway

The Chromosomal Instability (CIN) Pathway, also known as the adenoma-carcinoma sequence, is the most common pathway involved in CRC pathogenesis, and it is responsible for 65-70% of sporadic CRC [36]. While the term broadly refers to an accelerated rate of gains or losses of portions of chromosomes, resulting in karyotypic variability, these alterations generally end up to progressively introduce a characteristic set of mutations of specific oncogenes and tumor suppressor genes. These mutations generally involve the activation of the rat sarcoma virus (RAS) and/or Wingless-related integration site (WNT) pathways, the loss of cellular tumor antigen p53 (p53) and the loss of heterozygosity for the long arm of chromosome 18 (18q LOH). A summary of the main processes which can be involved in the production of these characteristic mutations is reported in Figure 6. Despite this, it is still unclear whenever the chromosomal instability is responsible for the accumulation of these mutations, or if the mutations itself are responsible for inducing this characteristic chromosomal instability [36].

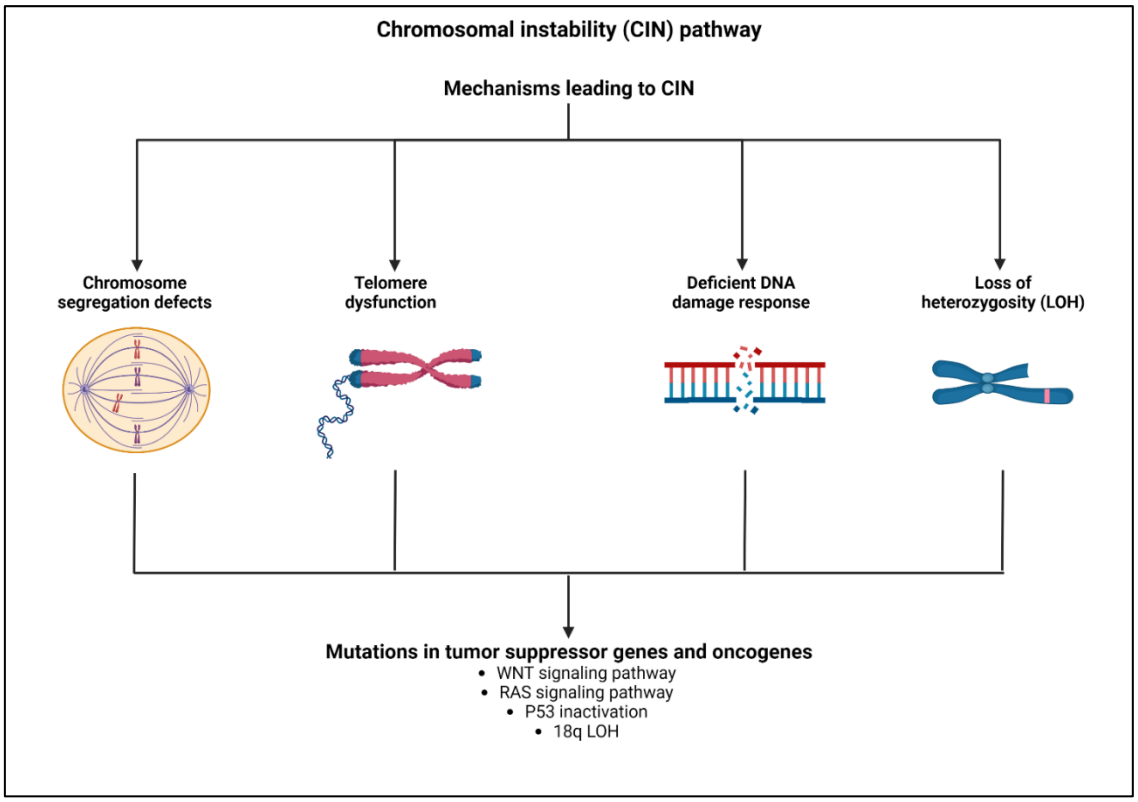


Figure 6. The chromosomal instability (CIN) pathway.

Activation of WNT pathway

The Wingless-related integration site (WNT) pathway represents one of most important pathways driving CRC pathogenesis, since over 92% of sporadic CRCs contains at least one alteration in one of the genes involved in this pathway [37].

The WNT signaling provides a distinctive proliferation stimulus, acting by regulating stemness along with cell migration and organogenesis during embryonic development. These morphogenic effects are mediated by the ability of WNT signaling to broadly affect gene expression, with subsequent effects on the mitotic spindle and the cytoskeleton [38, 39]. The WNT family consist of 19 secreted glycoprotein ligands which can be recognized by a complex of two molecules (FZD and LRP5/6), activating two main possible downstream signaling branches: the canonical (Wnt/β-catenin dependent) pathway and the non-canonical (β-catenin independent) pathway [40]. While much less its known about the non-canonical pathway in the context of CRC, the canonical WNT signaling pathway is a recognized driver of CRC development and it will be the one addressed here [41].

The binding of the Wnt proteins induces the dimerization of the two receptors, leading to the phosphorylation of the cytoplasmic portion of low-density lipoprotein receptor-related protein (LRP) and the recruitment of dishevelled (DVL) and axis inhibition (Axin) proteins, which acts as scaffold proteins. This induces the sequestration of the destruction complex (DC), a complex formed by the tumor suppressor APC and two constitutively active kinases (CK1 and GSK3β), towards the cell membrane, avoiding its interaction with β-catenin (as showed in Figure 7 – Right side). The missing interaction between the DC and β-catenin allows the progressive accumulation of β-catenin and its translocation to the nucleus, where it binds to TCF to upregulate target genes [42]. While the FZD/LRP receptors are not dimerized, instead, the destruction complex remains in the cytoplasm and it is responsible for the phosphorylation of β-catenin. The phosphorylated β-catenin

can then be recognized by the E3 ubiquitin ligase β-TrCP, which mediates its ubiquitination and proteasomal degradation (as showed in Figure 7 – Left side) [43].

Generally, all alteration of the WNT signaling in colorectal cancer are responsible for inducing, in the end, the progressive accumulation of β-catenin, which can be observed as a result of different mutations. The most common of these mutations involves the loss-of-function of the tumor suppressor gene APC, which has been showed to be altered in 83% of colorectal cancer cell lines [44]. These loss-of-function APC mutations are generally due to truncating mutations causing the loss of the β-catenin-binding domain of APC [45].

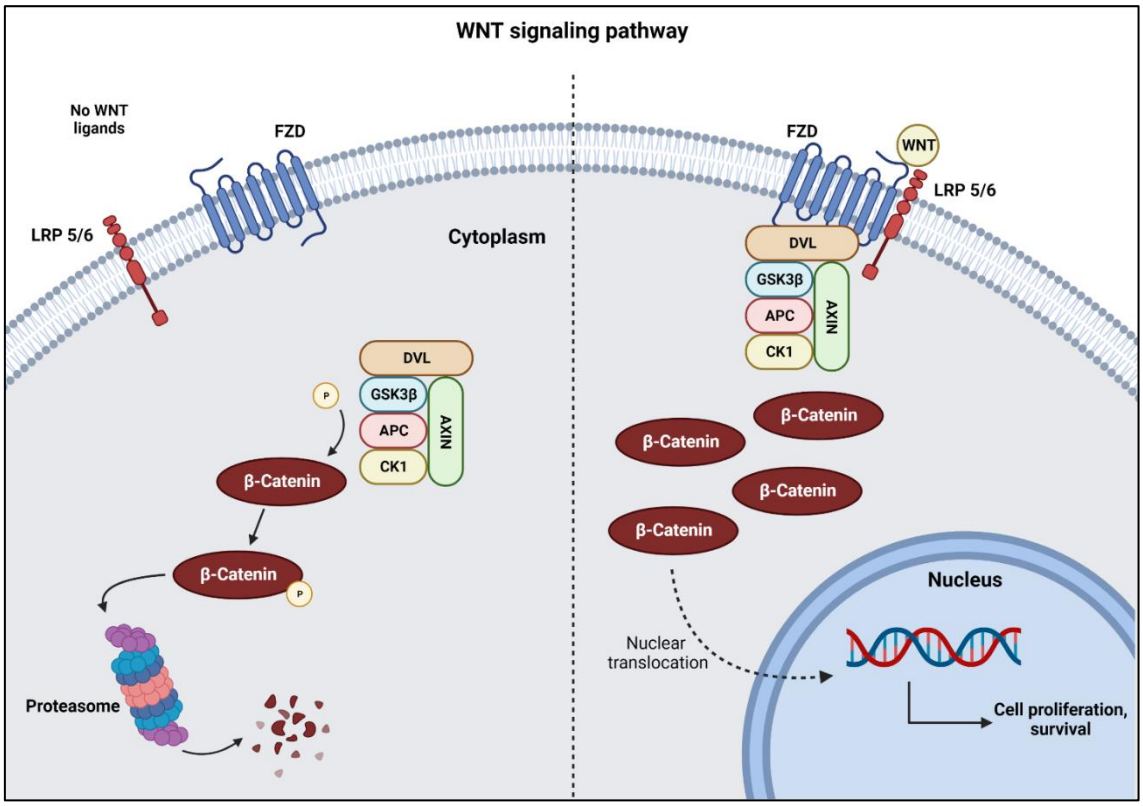


Figure 7. The WNT pathway.

The dimerization of the WNT receptors FZD and LRP 5/6 allows for the sequestration of the destruction complex and the progressive accumulation of β-catenin.

Activation of RAS pathway

The rat sarcoma virus (RAS) pathway is involved in regulating numerous aspects of cell proliferation, and tumors frequently presents activating mutations of the RAS genes or of other components involved in this pathway. This is even more frequent in CRC, where over 40% of CRC presents an activating mutation of the oncogene KRAS and 10% present a $BRAF^{V600E}$ mutation [37, 46].

The epidermal grow factor receptor (EGFR) is a membrane receptor located upstream of the RAS proteins and, in normal condition, its kinase activity is suppressed. After binding to any of its seven extracellular ligands, EGFR dimerizes and this activate its intracellular kinase domain. This process allows for the recruitment of growth factor receptor-bound protein 2 (GRB2) and the activation of guanine exchange factors, like Son of Sevenless (SOS) [47].

Three main RAS proteins have been associated with cancer development (K-RAS, N-RAS, H-RAS), and they normally cycle between one inactive GDP-bound state and one active GTP-bound state. Son of Sevenless (SOS) activation by EGFR causes GTP to be loaded on the inactive GDP-bound RAS proteins, turning them into their active conformations [48]. GTP-bound RAS is now able to bind and activate three closely related RAF proteins (BRAF, c-RAF1, ARAF), which can then phosphorylate and activate protein kinase kinases 1 and 2 (MEK1 and MEK2). Activated MEK proteins, in turn, phosphorylate and activate the mitogen-activated protein kinases (MAPKs) ERK1 and ERK2 [49]. ERK1 and ERK2, in turn, both phosphorylate numerous transcription factors involved in cellular proliferation, stimulating the cell cycle progression through the G1/S phase transition [50]. A schematic representation of the RAS pathway is offered in Figure 8.

Unsurprisingly, the common KRAS and BRAF mutations observed in CRC have important implications for the treatment of metastatic colorectal cancer with the anti-EGFR antibodies cetuximab and panitumumab. As these proteins are acting as

downstream effectors of EGFR, this subset of patients is generally resistant to the effect of anti-EGFR antibodies [51-53].

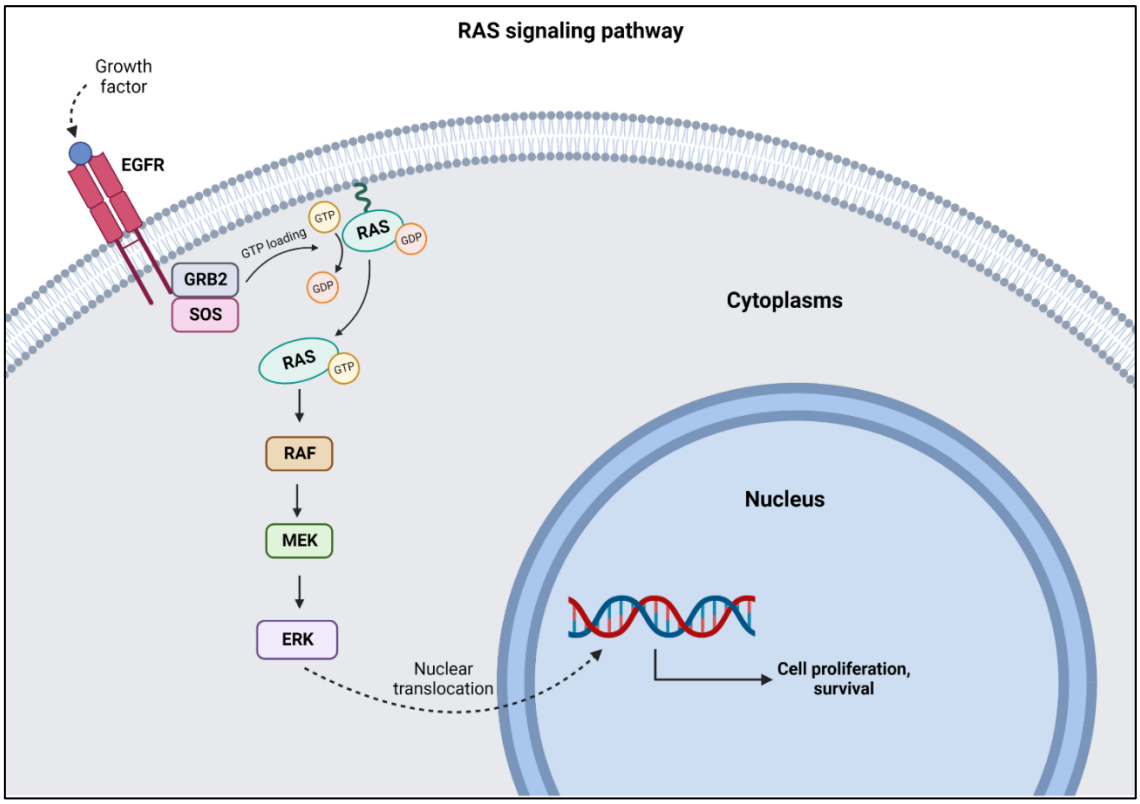


Figure 8. The RAS pathway.

EGFR dimerization causes the activation of SOS and the consequent GTP loading on RAS. GTP-bound RAS activates RAF, starting a cascade of signals which will result in the final nuclear translocation of ERK1/2 and the progression through the G1/S phase of the cell cycle.

P53 inactivation

TP53 represents one of the most studied tumour suppressor genes. It encodes for the phosphoprotein cellular tumour antigen p53 (p53) which, after DNA damage, can arrest the cell cycle and coordinate a series of responses involved in DNA repair or, if the damage is too severe, apoptosis [54]. The relevance of TP53 in regulating the response to DNA damage and cell survival is well represented by the fact that this gene is the most commonly mutated gene in human cancers, and over 60% of CRC cases present a loss-of-function mutation of TP53 [55, 56].

While a range of stimuli is involved in the regulation of p53 activity, the dominant role in the context of DNA damage is certainly due to the activation of the ataxia-telangiectasia mutated (ATM) and ATM-Rad3-Related (ATR) kinases [57]. While the ATM-Chk2 pathway is activated as a response to DNA double strand breaks (DSBs), the activation of ATR-Chk1 has been mostly associated to the response to DNA single strand breaks (SSBs) and replication stress instead [58]. Both these pathways converge in phosphorylating and activating p53, either because of the phosphorylation induced by Chk2, Chk1 or directly by ATR [59-61]. Phosphorylated p53 can then fulfil its role as a transcription factor and stimulate the expression of genes involved in cell cycle arrest, DNA repair and apoptosis (as showed in Figure 9) [62].

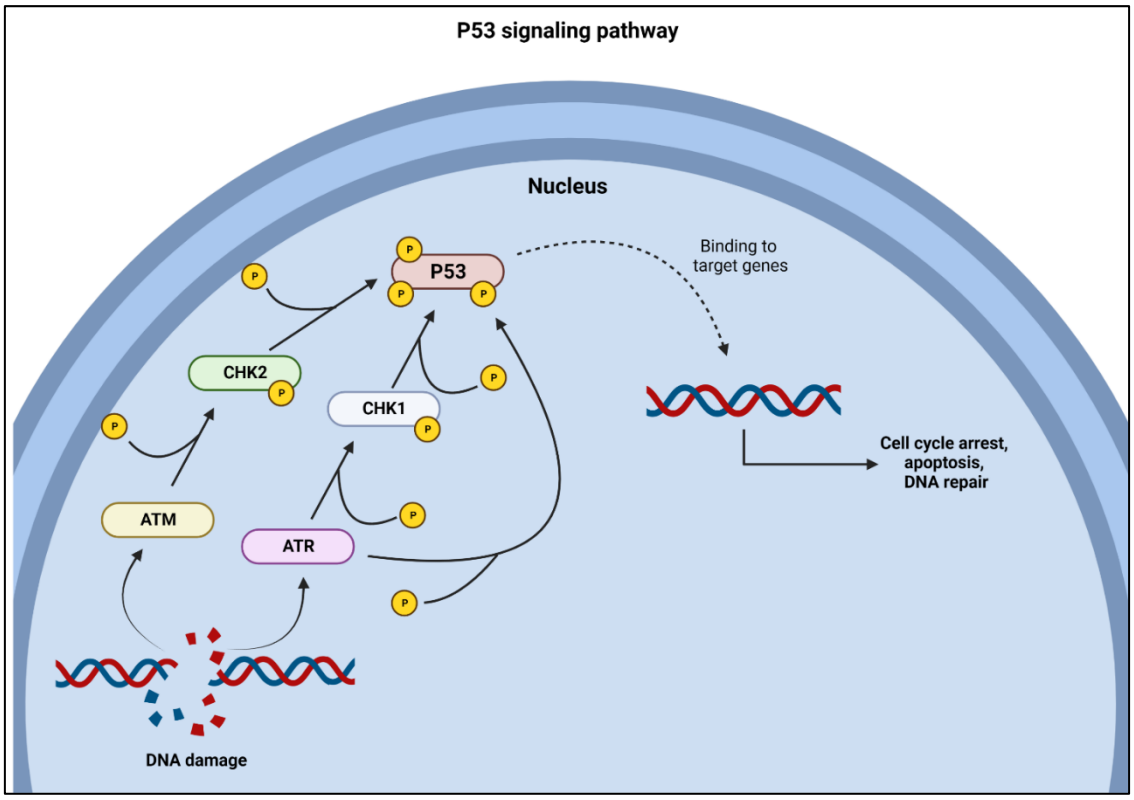


Figure 9. The p53 signaling pathway.

Activation of p53 after DNA damage is mediated by ATM-Chk2 and ATR-Chk1 pathways.

18q loss of heterozygosity

Loss of heterozygosity (LOH) is a common genetic abnormality in cancer, and it is defined as the loss of a portion of a chromosome. In diploid organisms, as two copies of the same chromosome are present, this phenomenon leads to the persistence of only one copy of each gene contained in the fragment of chromosome which has been lost [63].

LOH involving the long arm of chromosome 18 (18q) is particularly common in CRC cancer, and has been reported to be present in around 65% of CRC cases [64]. 18q LOH is generally associated to microsatellite stable (MSS) CRC cancers, but its prognostic significance has remained unclear [65]. The involvement of 18q LOH in CRC carcinogenesis is most likely due to the loss of tumor suppressor genes which are located within the long arm of chromosome 18, like SMAD Family Member 2 (SMAD2) and SMAD Family Member 4 (SMAD4) genes [65-67].

Microsatellite instability (MSI)

Microsatellites are generally defined as brief DNA sequences where a short motif of 1 to 6 bases is repeated multiple times, and they represent DNA hotspots characterized by a higher than normal mutation rate [68, 69]. This rate is further increased in microsatellite instable (MSI) cancers, which present a deficient mismatch repair (dMMR). DNA mismatch repair (MMR) is a well-studied DNA repair pathway involved in assuring DNA replication fidelity, genome stability and mutation avoidance [70]. Base-base mismatches in eukaryotic cells are recognized by one of the MutS heterodimers (MSH2/MSH6 or MSH2/MSH3), allowing for the recruitment of the MutLα heterodimer MLH1/PMS2 and the excision of the mismatched base. Then, the excision gap can be re-synthesized by DNA polymerase δ (Polδ), allowing for the right base to be inserted (Figure 10) [71, 72].

MSI can be observed in 15% of colorectal cancer, and it is generally due to the epigenetic silencing of MLH1 or a germline mutation of the main genes involved in mismatch repair (MLH1, MSH2, MSH6 or PMS2) [73]. Interestingly, the hypermutability observed in this subset of CRC cases has also been associated with a better response to immunotherapy, probably induced by the higher expression of neoantigens [74]. This has led to the recent approval of pembrolizumab as the first line treatment for metastatic MSI-H/dMMR [75].

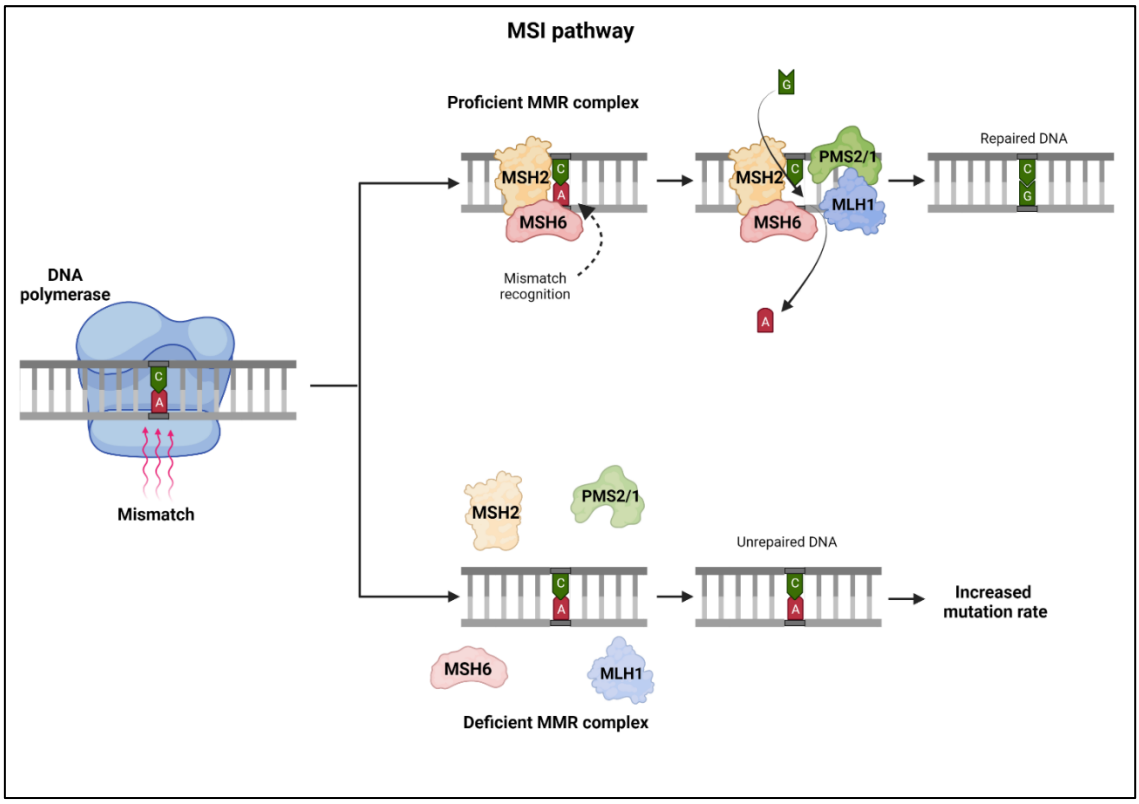


Figure 10. The microsatellite instability (MSI) pathway.

Microsatellite instability is induced by a deficient MMR and leads to an increased DNA mutation rate.

CpG Island Methylator Phenotype (CIMP)

The CpG Island Methylator Phenotype (CIMP) represents a condition characterized by a global increase in genome methylation, resulting in a more compacted chromatin structure and inaccessible promoters, with the potential of blocking the transcription of several tumor suppressor genes (like CDKN2A) and genes involved in DNA repair (like MGMT and MLH1) [76, 77]. CIMP has been determined to have a role in the pathogenesis of around 15% of all CRC cases, and it has generally been associated to older patients and the presence of $BRAF^{v600E}$ mutations [78-80].

In healthy human tissues, over 95% of CpG islands (CGIs) located in promoter regions are unmethylated, allowing the relaxed chromatin to be accessed by transcription factors involved in regulating genes expression [81]. The hypermethylation observed in cancer tissues is generally consider to be due to the increased activity of the de novo DNA methyltransferases (DNMT3A and DNMT3B) or the maintenance DNA methyltransferase (DNMT1) [82, 83]. The role of hypermethylation as a possible regulator of gene expression in colorectal cancer is summarized in Figure 11.

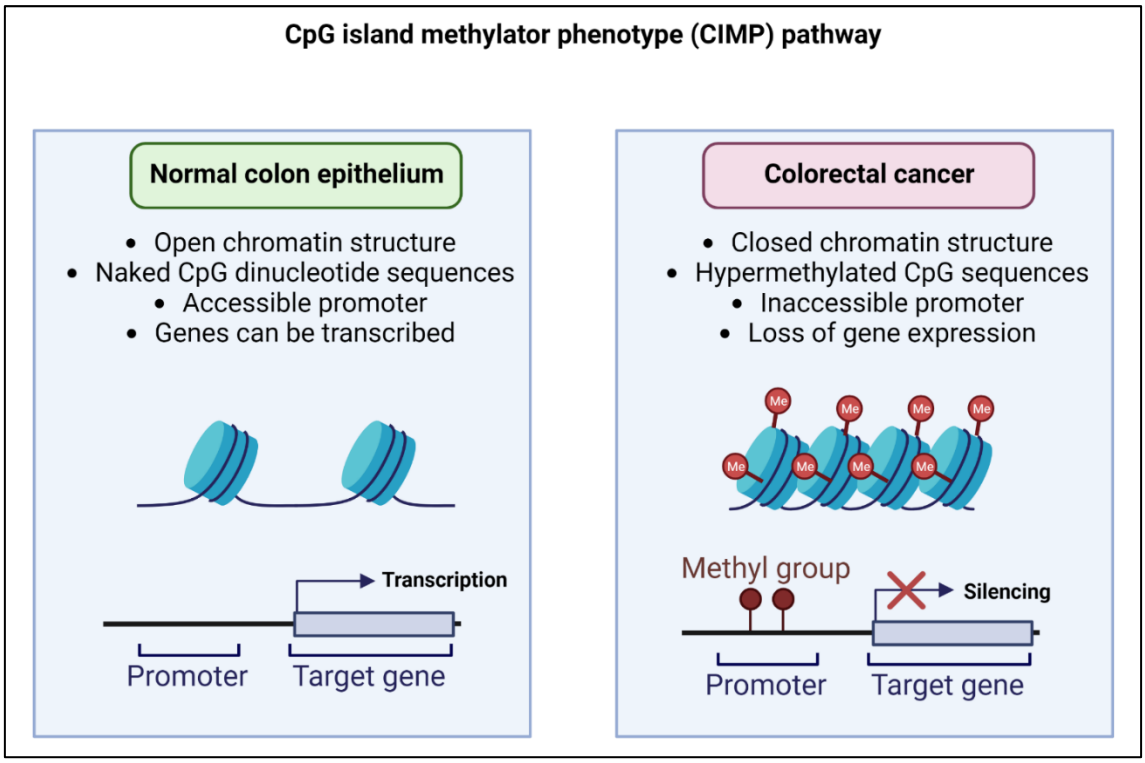


Figure 11. GpG island methylator phenotype (CIMP)

The increase methylation of promoter sequences compacts the chromatin, making those regions inaccessible to transcriptions factors. This consequently causes the silencing of target tumour suppressor genes.

The molecular classification of CRC (CMS, CRIS)

Having reviewed the most common genomic alterations responsible for the pathogenesis of colorectal cancer, it is now possible to address how this knowledge has been applied in the development of an appropriate molecular classification of different CRC subtypes. In particular, two main types of molecular classifications are now accepted and widely used:

- Consensus Molecular Subtypes (CMS) [16]
- Colorectal Intrinsic Subtypes (CRIS) [17]

Where both these molecular classifications (CMS and CRIS) are based on the analysis of transcriptomic data obtained from tumor samples, and a significant overlap exists between them, it is important to understand also what can differentiate them from each other [17]. The CMS classification was the first one to be developed in 2015, and it is by far the most largely accepted. However, this classification is also potentially influenced by the transcriptional components originating from the stromal cells surrounding the colorectal cancer epithelium. To address this, CRIS classification was subsequently developed two years later with the aim to focus more on cancer-intrinsic transcriptional features [17].

Consensus Molecular Subtype (CMS)

The CMS classification is a gene expression-based molecular classification of CRC which was developed in 2015 by the CRC Subtyping Consortium (CRCSC), with the aim to identify the presence of core subtype patterns among six independent classification algorithms [84-89]. This work allowed the identification of four main consensus molecular subtypes (CMSs), with distinguishing characteristics [16]:

- CMS1 (14% of cases) → Microsatellite instability, hypermutations, strong immune activation.
- CMS2 (37% of cases) → Epithelial phenotype, relevant WNT and MYC signaling activation
- CMS3 (13% of cases) → Epithelial phenotype, with metabolic dysregulation
- CMS4 (23% of cases) → Mesenchymal phenotype, stroma invasion, angiogenesis, TGF-β activation
- Mixed features (13% of cases) → Transition phenotype and intratumoral heterogeneity

The main characteristics of the four CMSs are recapitulated in Figure 12. Interestingly, the clinical relevance of the CMS classification has been recently confirmed by the fact that different CMS subtypes stratify the overall survival of metastatic CRC patients, independently from first-line treatment choice [90].

	CMS1 (MSI Immune)	CMS2 (Canonical)	CMS3 (Metabolic)	CMS4 (Mesenchymal)
MOLECULAR FEATURES	MSI CIMP high Hypermutation BRAF mutations	SCNA high	Mixed MSI status SCNA low CIMP low KRAS mutations	SCNA high
PATHWAYS	Immune activation and infiltration	WNT and MYC activation	Metabolic deregulation	Stromal infiltration TGF-β activation Angiogenesis
OUTCOME	Worse survival after relapse			Worse relapse-free and overall survival
THERAPY	Pembrolizumab	FOLFIRI or FOLFOX + anti-EGFR	FOLFIRI or FOLFOX + bevacizumab	FOLFIRI or FOLFOX + anti-EGFR

Figure 12. CRC subtypes according to CMS classification.

List of abbreviations: MSI (Microsatellite instability); CIMP (GpG island methylator phenotype); SCNA (Somatic copy number alterations).

Colorectal Intrinsic Subtypes (CRIS)

The CRIS classification of CRC was developed two years after the CMS classification was originally published, with the aim to reduce the impact that the stromal content could have on the overall transcriptomic data [17]. This was achieved by human-specific expression profiling of patient-derived xenografts, where the stromal cells derive from mice instead. By doing so, the CRIS classification focuses on the analysis of cancer-cell intrinsic transcriptional features, with the identification of five possible CRC intrinsic subtypes (CRISs):

- CRIS-A → Microsatellite instability, KRAS mutations, glycolytic
- CRIS-B → Epithelial-mesenchymal transition, TGF-β activation, poor prognosis
- CRIS-C → High EGFR signaling, good response to anti-EGFR antibodies
- CRIS-D → High WNT signaling, IGF2 overexpression and amplification
- CRIS-E → Paneth-cell like phenotype, TP53 mutations

The main characteristics of the five CRISs are recapitulated in Figure 13. Despite the reasons for which the CRIS classification was originally developed, it is important to notice that a significant partial overlap exists between it and the CMS classification. In particular, CRIS-A closely associates with CMS1 while CRIS-C is generally associated to CMS2 [91].

	CRIS-A	CRIS-B	CRIS-C	CRIS-D	CRIS-E
MOLECULAR FEATURES	MSI CIMP high KRAS mutations BRAF mutations		CIN high TP53 mutations MYC amplification	CIN high IGF2 amplification	CIN high KRAS mutations TP53 mutations
PATHWAYS	Glycolysis Inflammation Mucinous	EMT TGFβ activation	EGFR activation ERBB2 activation	IGFR activation FGFR activation WNT activation	WNT activation Paneth cells-like
OUTCOME	+	-	+	++	+
THERAPY	No response to anti-EGFR		Good response to anti-EGFR	No response to anti-EGFR	No response to anti-EGFR

Figure 13. CRC subtypes according to CRIS classification

List of abbreviations: MSI (Microsatellite instability); CIMP (GpG island methylator phenotype); CIN (Chromosomal instability); EMT (Epithelial-mesenchymal transition).

Mechanism of action of cytotoxic chemotherapy in CRC (5-Fluorouracil, Irinotecan, Oxaliplatin)

As previously described, the most common chemotherapeutic regimens generally used to treat mCRC patients are FOLFOX and FOLFIRI, which are considered to be equivalent in term of efficacy and are characterized by a median progression free survival (PFS) of around 8 months [23-26]. Their equivalence represents an interesting clinical challenge, because it brings along the following questions:

- Which patients are more likely to respond to FOLFOX, or FOLFIRI?
- Can we develop new biomarkers to stratify patients treated with FOLFIRI or FOLFOX?

While aiming to address these questions, it is important to remember that FOLFOX and FOLFIRI are regimens based on three different chemotherapeutic drugs: 5-fluorouracil (both), oxaliplatin (only FOLFOX) and irinotecan (only FOLFIRI). Understanding the mechanism of action of these three cytotoxic drugs becomes then a necessary starting point to address their efficacy when administered to mCRC patients.

5-Fluorouracil

5-FU was originally introduced to the clinic five decades ago and, today, more than two million patients are still treated with this drug every year [92]. The mechanism of cytotoxicity related to this compound is complex and due to the fact that, after entering the cell, 5-FU is converted in three different active metabolites (fluorodeoxyuridine monophosphate [FdUMP], fluorodeoxyuridine triphosphate [FdUTP] and fluorouridine triphosphate [FUTP]) [93]. Each one of these different metabolites has been proved to be involved on a different mechanism of action (Figure 14):

- FdUMP has been mostly connected to the inhibition of thymidylate synthase (TS) and the consequently inability of the cells to produce sufficient deoxythymidine monophosphate (dTMP), resulting in deoxynucleotide pool imbalances [93].
- FdUTP is manly incorporated into DNA and end up forming FU:A (mostly) or FU:G base pairs, which are then repaired by the base excision repair pathway (BER) with minor contribution from the mismatch repair pathway (MMR) [94].
- FUTP is incorporated into RNA, causing the disruption of normal RNA processing and functions [93, 94].

Despite these numerous mechanisms of action, recent studies seem to suggest that cells treated with 5-FU accumulated 3000 to 15000 more 5-FU in RNA than in DNA, and that the cytotoxicity of 5-FU is mainly RNA-mediated and consequently associated with the role of FUTP [94]. Accordingly, 5-FU is unlikely to be able to induce a relevant amount of DNA-protein crosslinks.

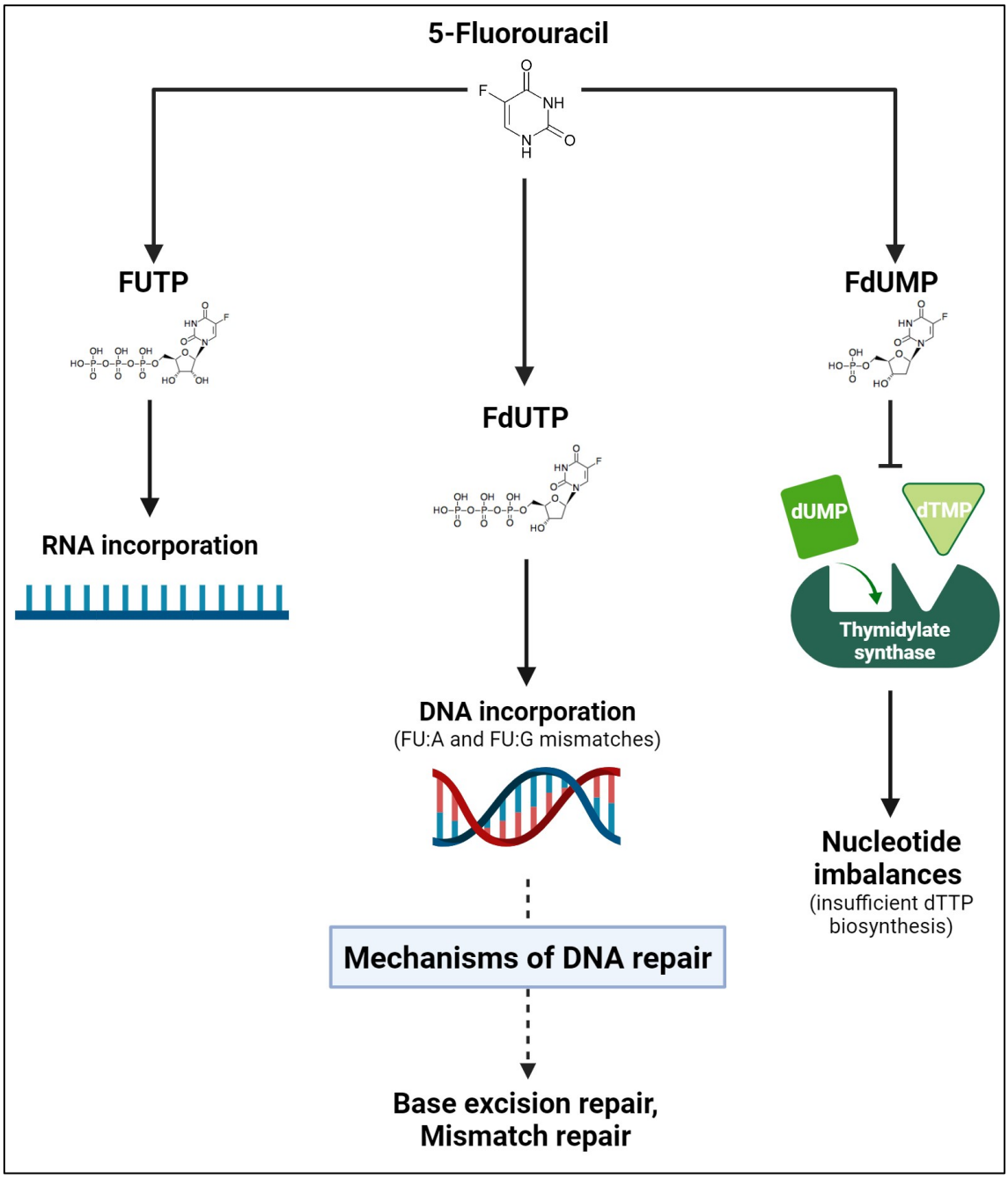


Figure 14. 5-Fluorouracil mechanisms of action.

Oxaliplatin

Platinum compounds like cisplatin, carboplatin and oxaliplatin have been largely used in the clinic since the discovery of cisplatin in 1970 [95]. Oxaliplatin, in particular, was developed in 1978 as a potential less toxic and more efficient alternative to cisplatin, but in vitro studies performed in the following twenty years identified oxaliplatin as an interesting possibility also for the treatment of cancers usually resistant to cisplatin, like colorectal cancer [96, 97]. Oxaliplatin is a platinum compound with a (1R,2R)-1,2-diaminocyclohexane (DACH) ligand and oxalate as a leaving group [98]. Due to its bulkier structure and higher lipophilicity, oxaliplatin-DNA adducts have been showed to be more efficient in the inhibition of DNA synthesis than cisplatin-DNA adducts [99]. This is true despite the fact that, when compared to cisplatin, oxaliplatin has reduced tendency of forming DNA-DNA crosslinks on guanine clusters by two to six-fold [100]. Another interesting difference with the previous compounds is the absence of cross-resistance between cisplatin and oxaliplatin [101].

The main cytotoxic lesion associated with platinum compounds is generally considered to be the formation of DNA-DNA crosslinks, which can be further divided in intrastrand (the most common) and interstrand crosslinks (Figure 15) [102]. DNA intrastrand crosslinks are typically repaired by the Nucleotide Excision Repair (NER) pathway, while DNA interstrand crosslinks are primarily repaired through the Fanconi Anemia (FA) pathway [103, 104].

Furthermore, some studies have started to assess oxaliplatin ability to induce secondary types of DNA lesions, including monoadducts and DNA-protein crosslinks (DPCs) [105, 106]. Specifically, regarding DPCs, the authors measured the amount of DPCs formed when isolated cell nuclei were exposed to equimolar amount of oxaliplatin and cisplatin. Similar to previously reported data regarding the formation of DNA-DNA crosslinks, cisplatin was two-fold more efficient in the formation of DPCs when compared to

oxaliplatin. However, despite these lower DNA and protein reactivities, oxaliplatin bulkier lesions seem to be more efficient in inhibiting cell growth [105].

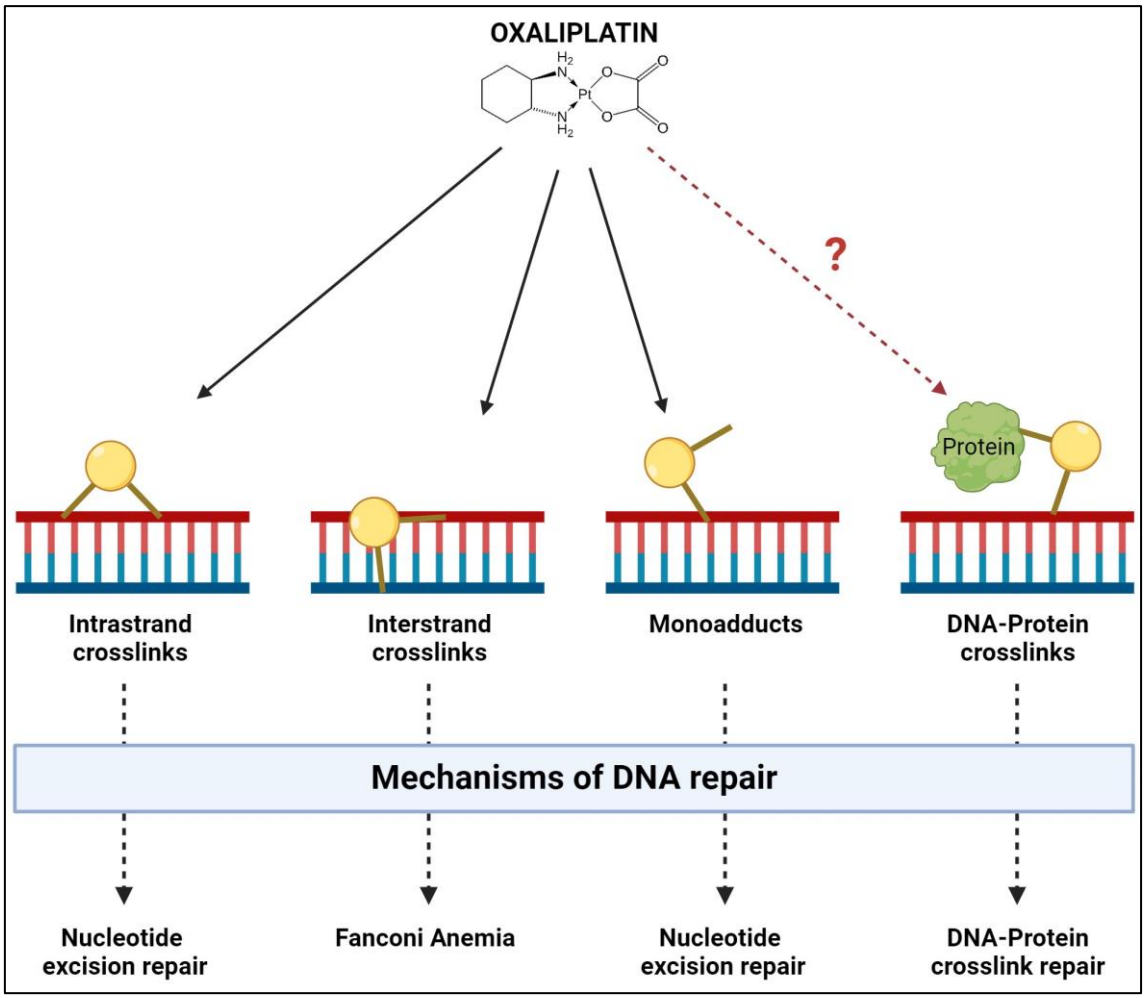


Figure 15. Oxaliplatin mechanisms of action.

Irinotecan

Irinotecan (CPT-11) is a prodrug derived from camptothecin and is metabolized by the carboxylesterase enzyme into its active metabolite, 7-ethyl-10-hydroxycamptothecin (SN-38) [107]. While most of the conversion of Irinotecan to SN-38 in patients is generally believed to happen within the liver tissue, a range of studies have demonstrated that this compound can be activated also by the native carboxylesterase expressed by tumor cells [108, 109]. Like other camptothecin derivates, irinotecan exert its antitumor activity by binding to topoisomerase I (TOP1) and preventing its release from the DNA, which leads to the persistence of single-strand breaks (SSBs) and double-strand breaks (DSBs) formation at replication forks (Figure 16) [110]. In normal conditions, TOP1 has been shown to be involved in allowing proper rotation and relaxation of supercoiled DNA. It does so by cleaving DNA to create a SSB to which it remains covalently bound to (this intermediate state, known as TOP1 cleavage complex [TOP1cc], is targeted and stabilized by camptothecin and its analogues), allowing the rotation of DNA before ligating it again and releasing itself [111].

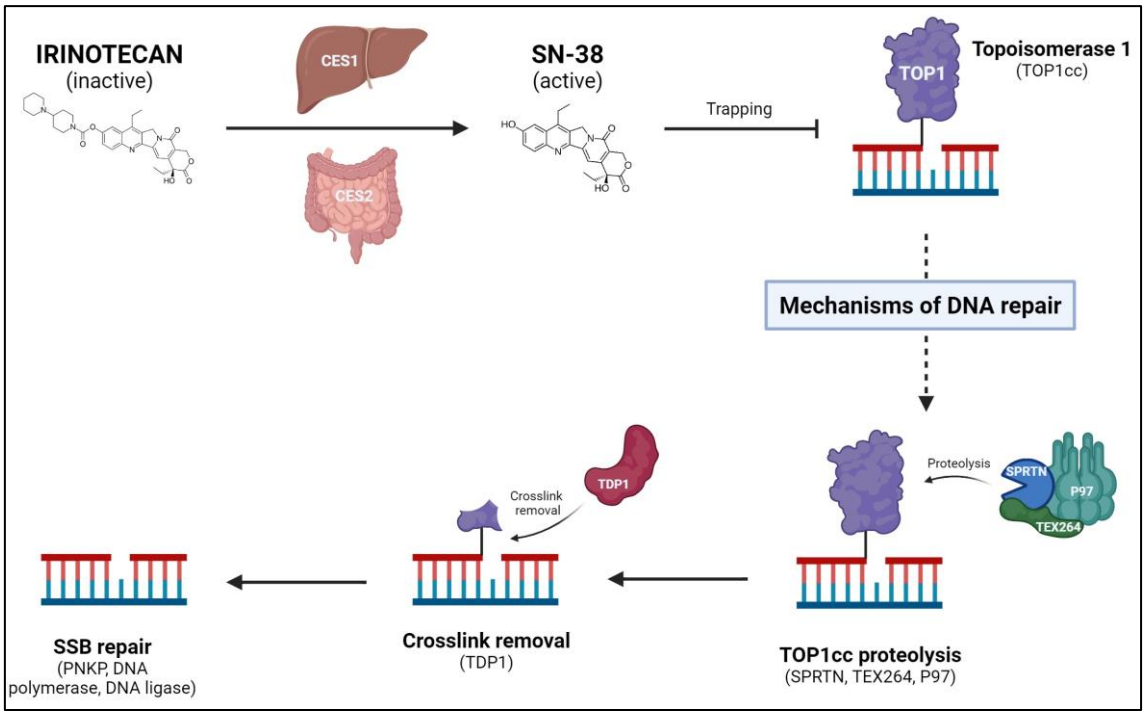


Figure 16. Irinotecan mechanisms of action.

DNA repair mechanisms of chemotherapy-induced DNA damage in CRC (5-Fluorouracil, Irinotecan, Oxaliplatin)

Understanding the main DNA repair mechanisms available to the cells for repairing the DNA damage induced by chemotherapy is important, as they have the potential to influence treatment outcomes [112]. Since the chemotherapeutic drugs most commonly used for treating CRC patient (5-Fluorouracil, Oxaliplatin, Irinotecan) are characterized by different mechanism of actions, however, these will now be reviewed according to the specific type of DNA damage they can cause and the type of DNA damage response (DDR) involved in its repair.

5-Fluorouracil

- *FdUMP inhibition of thymidylate synthase (TS):* TS is a folate-dependent enzyme, responsible of catalyzing the methylation of 2'-deoxyuridine-5'-monophosphate (dUMP) to 2'-deoxythymidine-5'-monophosphate (dTMP). As dTMP represents the sole intracellular source of 2'-deoxythymidine-5'-triphosphate (dTTP), a fundamental nucleotide necessary for the synthesis of DNA, the inhibition of TS by FdUMP leads to nucleotide pool imbalances [113].

  Interestingly, nucleotide pool imbalances have recently been associated to replication stress, with reduced replication fork speed and activation of ATR signaling [114]. ATR has been involved in preventing cell cycle progression through the regulation of intra-S-phase checkpoint, which is mediated by the phosphorylation of CHK1. pCHK1 in turn can phosphorylate CDC25A-C, leading to the inhibition of CDK1 and CDK2 [115].

- *FdUTP incorporation into DNA:* As previously discussed, FdUTP is manly incorporated into DNA to form FU:A and FU:G base pairs [94]. The main pathway in charge of repairing this form of DNA damage is base excision repair (BER). While BER can potentially be initiated by any of the 11 distinct mammalian DNA

glycosylases, in case of FU:A and FU:G base pairs this is generally mediated by the uracil-DNA glycosylase UNG [94, 116]. UNG exerts its activity by removing the damaged base, leaving an abasic site. The 5'-DNA strand located in this site is then cleaved by apurinic/apyrimidinic endodeoxyribonuclease 1 (APE1), allowing to the formation of 3'-OH and 5'-deoxyribose phosphate ends which can be recognized by the DNA polymerase beta (Pol β). At this point, Pol β can insert the correct nucleotide and complete the repair of these lesions [94, 116-118]

- *FUTP incorporation into RNA:* Despite the fact that FUTP incorporation into RNA has been reported to be the main mechanisms responsible for the cytotoxicity of 5-Fluorouracil, this is by definition not repaired by any DNA damage repair (DDR) process as it is happening on RNA instead [93, 94]. However, it is worth mentioning that the incorporation of FUTP into RNA has been reported to cause splicing defects affecting numerous intron-containing mRNAs along with rRNA and tRNA processing defects [119].

Oxaliplatin

- *Intrastrand crosslinks:* The repair of platinum-induced intrastrand crosslinks has been demonstrated to rely mostly on the nucleotide excision repair (NER) pathway [120]. A number of studies were able to indicate that XPC is able to recognize cisplatin-mediated intrastrand crosslinks [121] and that NER-deficient bacterial and mammalian cells are more sensitive to platinum drugs [122, 123]. NER is characterized by two different sub-pathways which differ in their mode of DNA-damage recognition, but converge into one pathway for downstream processing and repair events. The two sub-pathways for damage recognition are global genome NER (GG-NER), which relies on XPC, TFIIH and XPD, and transcription-coupled NER (TC-NER), which relies on the recruitment of CSA and CSB following the stalling of RNA polymerase 2. Once the presence of a lesion has been confirmed, the recruitment of structure-specific nucleases like XPF-

ERCC1 and XPG allows the incision of the damaged DNA strand and the filling of the resulting ssDNA gap by the coordinate action of DNA polymerases and DNA ligase I [124]**.**

- *Interstrand crosslinks (ICLs):* The repair of cisplatin-induced ICLs is mediated by both the NER and the homologous recombination (HR) pathways, depending on the specific phase of the cell cycle in which the ICL is induced [103]. Moreover, also the Fanconi Anemia (FA) pathway has been associated to platinum drugs sensitivity [125, 126].

  In quiescent G0/G1 cells, ICLs are recognized by the NER machinery and first incised by a complex formed by XPF-ERCC1 and XPG. After this, the ICL lesion is bypassed by a translesion synthesis (TLS) polymerase before a second incision can be made and the lesion can be completely removed [127, 128]. The recruitment of a more permissive TLS polymerase, like REV1, has been showed to be dependent on the ubiquitination of PCNA, which increases its affinity with TLS polymerases capable of performing DNA synthesis across the lesion, in exchange for lower replication fidelity [129].

  In proliferating S phase cells, ICLs are able to induce stalling of replication forks and the recruitment of the FA core complex, permitting the incision of the ICLs and the introduction of a DBS, which can then be repaired by homologous recombination (HR) [127, 130, 131]. More in details, following collision of replication forks due to the presence of the ICL during DNA replication, the lesion is recognized by UHRF1 which recruits the FANCD2-FANCI complex. This complex is then monoubiquitinated by FANCL, a E3 ubiquitin ligase part of the FA core complex, in order to allow the recruitment of structure-specific nucleases like XPF-ERCC1. By introducing a DBS at the lesion, the ICL is unhooked and the two double helix, one containing the residual ICL and one containing a DBS,

are separated. Following this step, the double helix containing the ICL is repaired through the previously described NER pathway, while the remaining double helix containing the DBS will be repaired through the HR pathway [132].

- *Monoadducts:* Monofunctional platinum-DNA adducts represents a significant obstacle for DNA transcription and, similarly to intrastrand crosslinks, their repair is mediated by the NER pathway [133].

- *DNA-protein crosslinks (DPCs):* The mechanisms of DPCs repair will be addressed in the section entitled “DNA-protein crosslinks repair and the role of SPRTN protease”. Despite this, it is important to underline how some papers have now finally started to study platinum-induced DPCs [134], and their relevance in mediating the cytotoxic effect of platinum compounds has finally begun to be addressed. In particular, SPRTN has been showed to have a role in mediating the efficacy of cisplatin both in cell lines and in more complex systems, like C. elegans [135, 136].

Irinotecan

- *TOP1 cleavage complex (TOP1cc) formation:* Until recently, the most critical enzyme for TOP1cc resolution was believed to be the tyrosyl-DNA phosphodiesterase 1 (TDP1), which is able to hydrolyze the bond that links a tyrosine in the active site of TOP1 to a 3’ phosphate group on the SSB DNA [137]. It has also been demonstrated that TDP1 depletion induces hypersensitivity to Irinotecan in a TOP1-dependent manner in RKO and SW480 colorectal cancer cells [138]. While the importance of TDP1 remains partially true, it is worth noticing that this enzyme can only process small peptide fragments and most of the TOP1 protein must be proteolytically degraded before TDP1 can intervene [139]. In the last few years, our laboratory has been able to address this important area and has identified the p97-SPRTN-TEX264 complex, which

specifically repair Top1-cc and acts upstream of TDP1. We have shown that TEX264 is able to recognize TOP1ccs and then to recruit the p97-SPRTN complex to proteolytically degrade those lesions [140]. Basing on the aforementioned results obtained evaluating the TOP1ccs induced by camptothecin, it can be assumed this could be true also for TOP1ccs induced by irinotecan. The role of SPRTN in irinotecan-induced DNA damage has never been previously evaluated, and this currently represent an interesting area which requires to be further addressed.

DNA-protein crosslinks repair and the role of SPRTN protease

DNA-protein crosslinks (DPCs) are a heterogeneous type of DNA lesion characterized by proteins covalently bound to the DNA. DPCs, due to their bulky structure, represent relevant obstacles to the progression of the replisome during DNA replication, causing stalling of replication forks and leading to DNA double-strand breaks [141].

DNA-protein crosslinks: molecular structure and repair pathways

DPCs can be caused by numerous sources, including a wide range of chemotherapeutic drugs, radiations or even by the endogenous cellular metabolism [142, 143]. They can also be categorized as enzymatic DPCs (like TOP1ccs) or non-enzymatic DPCs (like the one caused by histones), based on the molecular functions of the cross-linked protein [144].

The molecular structure of DPCs is characterized by three main components (the DNA, the crosslink and the crosslinked protein), all of which can potentially be targeted through different repair mechanisms (nuclease-dependent repair, direct crosslink hydrolysis, protease-dependent repair) as showed in Figure 17 [142]. The main characteristics of each of these repair pathways will now be briefly described.

Nuclease-dependent repair

The MRN complex (formed by MRE11, RAD50 and NBS1), with its nuclease activity, has proved to the able to cleave the DNA component of DPCs [145, 146]. This is due to the fact that DPCs have the potential to induce significant distortions of the DNA double helix and, especially during DNA replication, to end up producing double-strand breaks (DSBs) containing a DPC-blocked DNA end. The correlation between DPCs formation and DSBs offers an opportunity to the cell to repurpose nucleases normally involved in DSB repair in order to process the DPC-blocked DNA end and obtain a clean DBS instead [142, 147]. This is generally performed by meiotic recombination 11 (MRE11), an endonuclease part of the MRN complex, which coordinates the endonucleolytic cleavage of the DPC-blocked DNA end. Then, the resulting clean DSB has the potential to be repaired using canonical DBS repair mechanisms (HR or NHEJ), despite some data suggest that this repair is more likely to occur by NHEJ [145, 148].

Direct crosslink hydrolysis

Tyrosyl-DNA phosphodiesterases (TDP1 and TDP2) are responsible for the direct hydrolysis of specific subtypes of crosslink bonds, like the ones between the DNA and the trapped topoisomerases (TOP1 and TOP2, respectively) [149-153].

These enzymes have evolved to specifically remove TOP1 and TOP2 crosslinks, which are physiological intermediates formed between the Tyr residue in the active site of TOP1 or TOP2 and the 3' end of a single-strand break (SSB). These transient SSB are normally generated to induce the relaxation of supercoiled DNA in order to allow DNA replication and transcription, but they can be stabilized by topoisomerase poisons like camptothecin and etoposide [142, 154, 155]. TDP1 and TDP2 are able to hydrolyze the phosphotyrosyl bond between topoisomerases and DNA in order to produce 3'-phosphate ends, which will then be removed by the polynucleotide kinase 3′-phosphatase (PNKP) before the canonical SSB repair can proceed [151, 156]. While these general principles remain true, however, recent studies also suggest that TDP1 is not able to remove full-length TOP1 from DNA, and TOP1 should first be partially degraded through the proteasome or the p97-SPRTN-TEX264 complex [140, 157]. These findings suggest an interesting interaction between different DPC-repair pathways (protease-dependent repair and direct crosslink hydrolysis), hinting that these pathways are likely to collaborate in order to successfully repair TOP1 and TOP2ccs.

Protease-dependent repair

The crosslinked protein can be targeted directly by the proteasome or by DPC proteases, like SprT-Like N-Terminal Domain (SPRTN) in metazoans and weak suppressor of SMT3 protein 1 (WSS1) in lower eukaryotes, in a process known as replication-coupled DPC proteolysis [158-162]. This pathway and its relevance in colorectal cancer represents the main focus of this thesis.

The first suggestion regarding the existence of a protease-dependent DPC repair mechanism originated in 2014, from a synthetic lethality screen performed in *S. cerevisiae*. In this experiment, it was noted that TDP1-deficient cells are particularly reliant on the expression of WSS1 for their survival. ΔWSS1ΔTDP1 double mutant cells were characterized by cell cycle arrest, slow growth and DNA damage checkpoint activation [161]. A similar repair mechanism was also suggested in higher eukaryotes [163], but two more years of work were required before SPRTN could be characterized as the first DNA-dependent protease involved in DPC repair in higher eukaryotes [136, 158, 164]. Soon after that, SPRTN was recognized as an important protease contributing to the repair of a large range of DPCs, from formaldehyde-induced DPCs to TOP1/2ccs [159, 160]. Furthermore, germline mutations of SPRTN (Ruijs Aalfs syndrome) were also observed in three patients characterized by premature ageing, early onset hepatocellular carcinoma and genomic instability [165, 166].

Interestingly, proteases involved in DPCs repair are required to be characterized by a significant lack of substrate specificity, due to the numerous potential substrates capable of originating DPCs. In SPRTN, this is mostly achieved by the recognition of the DNA-double helix distortion and stalling of the replication fork, rather than the actual crosslink [163]. The proteolytic degradation mediated by SPRTN during DNA-replication can produce an almost completed degradation of the protein covalently bound to the DNA, leaving only a small peptide-DNA adduct which can then be bypassed by the replicative CDC45–MCM2/7–GINS (CMG) helicase [142, 167]. The DNA synthesis can then

proceed over the peptide-DNA adduct thanks to the intervention of a translesion synthesis polymerase like DNA polymerase ζ [163]. As an alternative, the remaining peptide-DNA adduct can also be removed through the NER pathway [168], which has already been described while discussing the repair of intrastrand crosslinks induced by oxaliplatin.

Beside the identification of proteases involved in DPC repair, the proteasome has also been investigated as a possible contributor to the repair of DNA-protein crosslinks [169]. The intervention of the proteasome implies that DPCs should first be ubiquitinated, and in fact numerous DPCs have been reported to be subject to post-translational modifications like ubiquitylation [170, 171]. Despite this, the role of proteasome in DPC repair remains controversial, as the ubiquitylation could also simply have a signaling role within this context [143, 172].

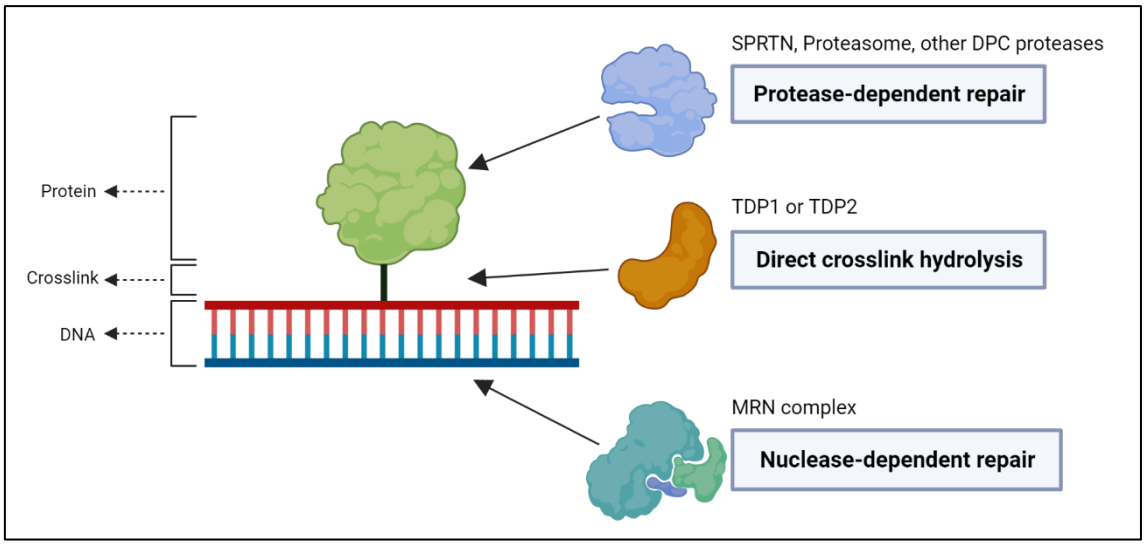


Figure 17. Main mechanisms of DPC repair.

SPRTN, a DNA-dependent metalloprotease involved in replication-coupled DPC proteolysis

Since the discovery of SprT-Like N-Terminal Domain (SPRTN) as an essential enzyme involved in the repair of DPCs in 2016 [136, 158], most of the work in the field has been focused on understanding how SPRTN activity is regulated. Thanks to this we are now aware that, in response to DPCs induction, ATM/ATR phosphorylates and activate the deubiquitinase VCPIP1, which in turn deubiquitinates SPRTN and promotes its relocalization within the chromatin. After its relocalization, the subsequent acetylation of SPRTN by PCAF or GCN5 allows SPRTN to interact with the specific site of DNA damage and exert its proteolytic activity, as summarized in Figure 18A-D [173]. Moreover, it has been showed that SPRTN preferentially cleaves DPCs in proximity of DNA overhang between single-stranded and double-stranded DNA structures. These structure are specifically recognized with two distinct interfaces on the N-terminal part of SPRTN [174]. SPRTN substrates are also not characterized by a common sequence motif, and can generally be cleaved in any positively charged regions rich in arginine, lysine or serine residues [175]. This characteristic is extremely relevant as it underlines how SPRTN-mediated proteolysis has the potential to involve a large number of different substrates.

The PIP domain, mediating the interaction between SPRTN and PCNA, is particularly important for assuring SPRTN correct function as a DPC protease (Figure 18A) [159]. PCNA trimeric complexes, an integral component of the human replisome, encircle the DNA and assure the correct interaction between DNA polymerases and DNA strands during the S-phase of the cell cycle [176, 177]. Thanks to its binding to PCNA, SPRTN becomes able to detect the DPC-mediated stalling of polymerases during DNA replication, mostly restricting its activity to a replication-dependent process [162].

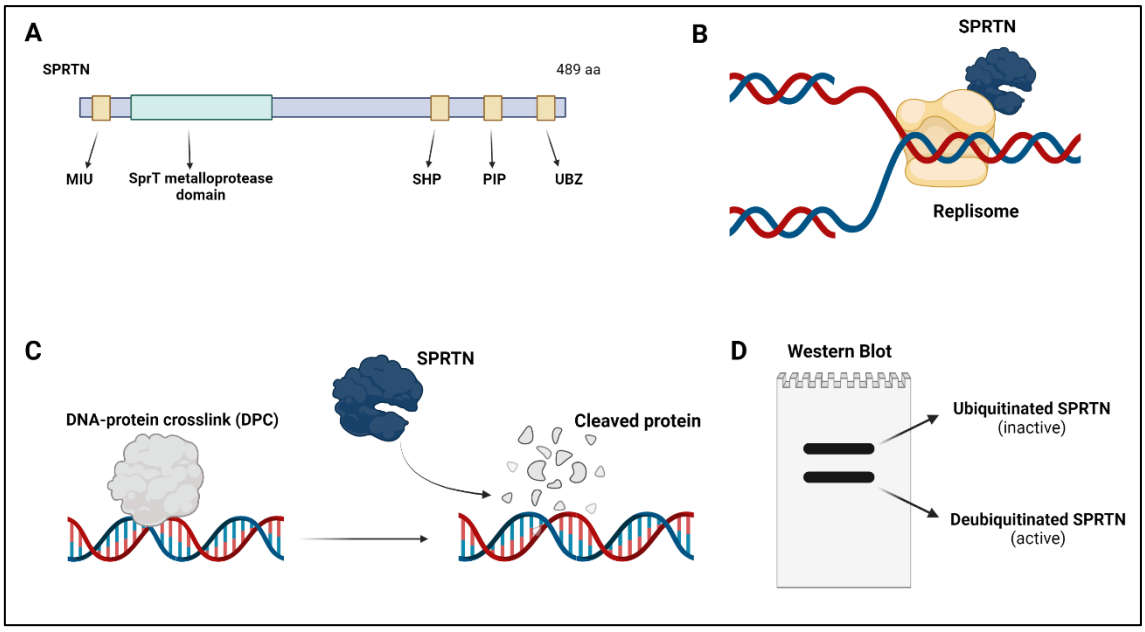


Figure 18. SPRTN and its role in DPC repair.

(A) SPRTN gene and its different functional domains: MIU (ubiquitin binding domain), SprT metalloprotease domain (zinc metalloprotease domain), SHP (p97 binding domain), PIP (PCNA binding domain), UBZ (ubiquitin binding zinc finger domain); (B) Through the interaction between SPRTN and PCNA, SPRTN binds to the replisome and removes DPCs in a replication-dependent proteolysis; (C) SPRTN recognizes DPCs and cleaves them, removing the bulky lesion from the DNA; (D) SPRTN present itself as a double band on western blots, due to its ubiquitinated and deubiquitinated forms.

Alternative DPC proteases

The study of DPC proteases still represents a fascinating area of scientific enquiry, as SPRTN and WSS1 are not the only DPC proteases which have been identified in recent years. GCNA, FAM111A/B and DDI1 have all been suggested as possible alternative DPC proteases [170, 178-180]. However, the relevance of each protease in the clearance of specific DPC subtypes has not yet been properly investigated, especially in the context of DPCs induced by different chemotherapeutic drugs.

GCNA

GCNA expression is generally limited to pluripotent cells and germ cells and, in this context, it has been confirmed to have a role in repairing the TOP2ccs induced by topoisomerase-2 poisons like etoposide [181]. The importance of GCNA in DPC repair during the early phases of vertebrate development has also recently been confirmed in a zebrafish model, were GCNA deficiency results in embryonic death due to DPC accumulation [182].

FAM111A

FAM111A has been identified in 2020 as a PCNA-interacting protease possibly involved in DPC repair. Through its trypsin-like protease domain, it has been suggested to play an important role in removing DPCs during DNA replication, in particular in case of PARP1-DNA complexes induced by PARP inhibitors like niraparib [178, 183]. Furthermore, FAM111A knock-out cells have been reported to be more sensitive to topoisomerase 1 poisons like camptothecin, but not to platinum compounds like cisplatin [178]. These results suggest that FAM111A could present a partially overlapping activity with SPRTN, but how the cells might choose between the two different DPC proteases has never been investigated.

FAM111B

FAM111B is a poorly characterized trypsin/cysteine-like protease and its gene represents a paralog of FAM111A, sharing over 45% of sequence homology [184]. Like

SPRTN, FAM111B levels have been reported to increase during the S-phase of the cell cycle and FAM111B knockout cells are characterized by a difficult progression through the S and G2/M phases [185, 186]. While the specific roles of FAM111B in DNA-repair have never been investigated, its similarity to FAM111A suggest that FAM111B might also represent an alternative DPC protease.

DDI1

DDI1 has been recently identified as a ubiquitin-dependent aspartic protease with a role in the S-phase dependent repair of TOP1ccs in Saccharomyces cerevisiae. Furthermore, loss of DDI1 or its protease domain has also been reported to hypersensitize cells to numerous DPC-inducing agents like camptothecin, etoposide and formaldehyde [180, 187]. Despite these encouraging results, the role of DDI1 in higher eukaryotes, including humans, remains to be properly investigated.

Main hypothesis and aims

To summarize, my hypothesis is the following:

- SPRTN is likely to have a role in repairing the DNA-protein crosslinks (DPCs) caused by the chemotherapeutic drugs currently used to treat colorectal cancer patients (5-fluorouracil, oxaliplatin and irinotecan). SPRTN could consequently modulate the efficacy of those drugs, with the potential to become an interesting biomarker and a possible therapeutic target.

To test my hypothesis, I aim to:

1. Investigate if common chemotherapeutic drugs used in colorectal cancer are able to cause DPCs.
2. Determine whether defects in DPCs repair affect the outcome of chemotherapy in cell cultures and in patients.
3. Understand the role of SPRTN in processing DPCs induced by chemotherapeutic drugs in colorectal cancer.

Chapter 2: Material and methods

Cell culture

All cell lines were grown in Dulbecco's Modified Eagle's Medium (DMEM) – High glucose [Sigma-Aldrich - D6429-500ML], supplemented with 10% Fetal Bovine Serum (FBS) [Sigma-Aldrich - F2442-500ML] and 1% Penicillin/Streptomycin [Sigma-Aldrich - P0781-100ML].

- HCT116 wild type (CCL-247™) was purchased from ATCC.
- HCT116 TP53 $^{-/-}$ was obtained from Ester Hammond's laboratory at the University of Oxford [188].
- HCT116 OsTIR1 was obtained from Ross Chapman's laboratory at the University of Oxford [189].
- HCT116 OsTIR1 SPRTN-AID was produced as part of this PhD project, starting from HCT116 OsTIR1. Further information is available on the Result (Development and characterization of a SPRTN-AID system in HCT116 cells) and Material and Methods (Auxin-inducible degron) sections of this thesis.
- HT55 wild type (ECACC 85061105) was purchased from ECACC.
- RKO wild type (CRL-2577™) was purchased from ATCC.
- HeLa wild type (CCL-2™) was purchased from ATCC.
- HeLa SPRTN $^{-/-/+}$ was previously produced in our laboratory by John Fielden [158].

An overview of HCT116, HT-55, RKO and HeLa different characteristics can be found in Supplementary Table 1. Cells were tested monthly with MycoAlert Mycoplasma Detection Kit [Lonza - LT07-318] to exclude possible mycoplasma contaminations [190].

siRNA

Briefly, cells were dissociated into single cell suspension by using a solution of 10% Trypsin-EDTA [Gibco – 15400-054] in Phosphate-Buffered Saline (PBS) before being transfected with the indicated siRNA (Table 1) and Lipofectamine RNAiMAX [Invitrogen – 13778-150] in their culture medium for 24 hours. siRNAs were previously diluted in 1000 µl Opti-MEM [Gibco – 31985-070] along with 10 µl Lipofectamine RNAiMAX. After 20 minutes of incubation at room temperature, the complexes were added to the cells to a final volume of 5 ml. An overview of how siRNAs can be used to silence the expression of different genes is provided by Birmingham A. et al. [191].

siRNA	Target	Working concentration	Company	Sequence
siLUC	Luciferase	20 nM	Microsynth	CGUACGCGGAAUACUUCGATT
siSPRTN 3'UTR	SPRTN	50 nM	Microsynth	GUCAGGAAGUUCUGGUUAA
siSPRTN Q1	SPRTN	100 nM	Microsynth	AGCCAAUAUAACGGUAUACCA
siSPRTN Q2	SPRTN	20 nM	Microsynth	CACGAUGAGGUGGAUGAGUAU

Table 1. Oligonucleotides used for siRNA depletions.

Colony forming assays

Cells were dissociated into single cell suspension by using a solution of 10% Trypsin-EDTA [Gibco – 15400-054] in PBS. Then, cells were seeded using DMEM – High glucose [Sigma-Aldrich - D6429-500ML] in 6-wells plates at a concentration between 1000 and 1400 cells/well according to the specific depletion that was performed and how this affected seeding efficiency. Cells were given 24 hours to attach to the bottom of the well before the media was removed and the treatment could begin, for the indicated amount of time. After this, cells were incubated until they were able to develop visible colonies (1 week for HCT116, RKO and HeLa; 2 weeks for HT-55). Colonies were fixed in methanol [Fisher Chemical - M/3950/17] for 10 minutes and stained with crystal violet solution [Merck - V5265-500ML] for 30 minutes. The colonies were then washed in H_2O before letting them dry overnight at room temperature. The following day, the number of colonies was counted using the automatic colony counter GelCount™. An overview of the colony forming assay has been published by Franken N et al. and can be consulted to obtain further details regarding this technique [192].

Figure 19A-C shows how the colony forming assays were performed according to different experimental conditions (siRNA, HCT116 SPRTN-AID, HeLa WT and HeLa SPRTN $^{-/-/+}$). Each data point in the colony forming assays presented in this thesis (Figures 30-34, 36) represent the mean result obtained from three independent biological replicates, each containing three technical replicates. To determine the statistical significance of the findings, a comparison by Student's t-test was performed for each data point corresponding to the same drug concentration [193].

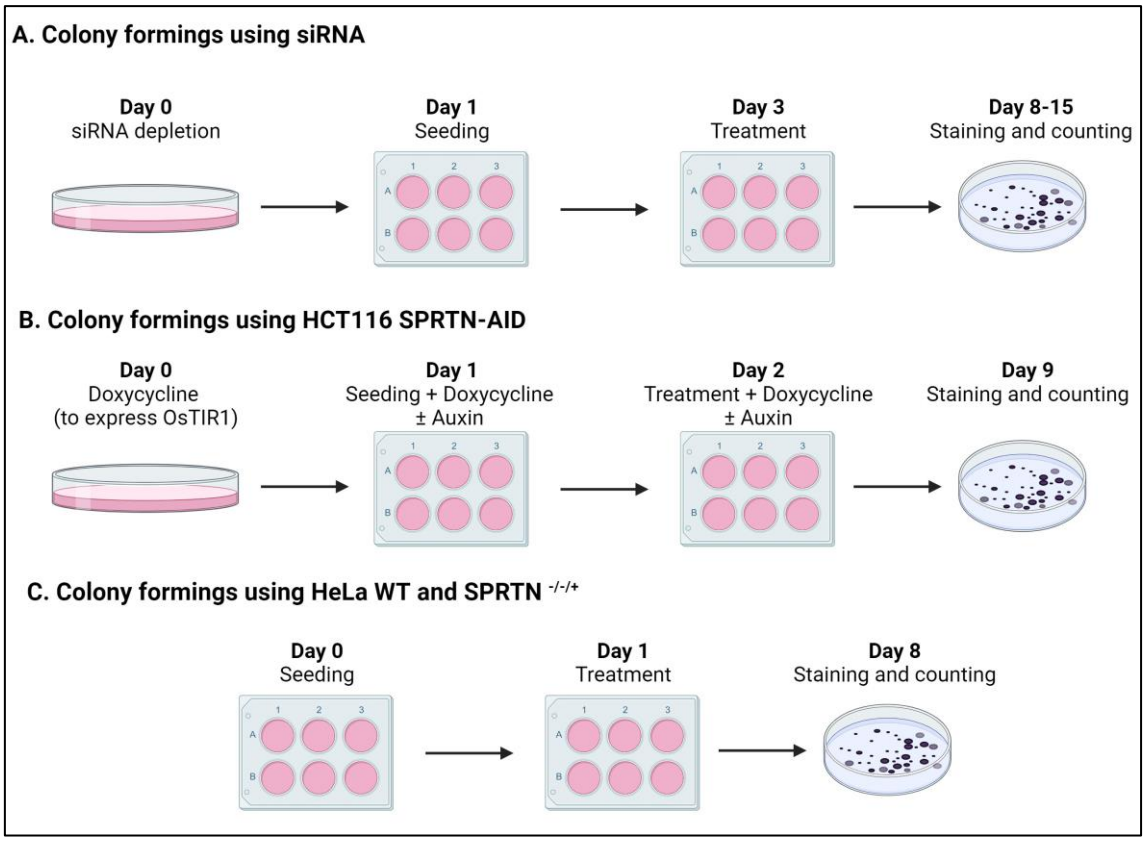


Figure 19. Colony forming assays.

(A) How colony forming assays were performed in HCT116, RKO and HT55 cells while depleting SPRTN by siRNA. (B) How colony forming assays were performed in HCT116 OsTIR1 SPRTN-AID. (C) How colony forming assays were performed in HeLa WT and SPRTN $^{-/-/+}$

Western blots

Cell were harvested and lysed with urea buffer (8 M urea, 50 mM Tris pH 6.8, 5% SDS and 10% glycerol in H_2O). The whole cell extract protein amounts were then measured using the bicinchoninic acid (BCA) assay [194], and 65 µg of proteins for every sample were run using gradient 5-18% gels. An brief overview of how western blots are performed and their theoretical background can be found in the work of Mahmood T et al. [195].

The membrane obtained at the end of the transfer was then blocked with 5% skim milk in Tris-Buffered Saline with Tween (TBS-T) for 1 hour, before being incubated with the desired primary antibody overnight at 4°C, in a solution of 3% BSA in TBS-T. The list of the primary antibodies used in this thesis is reported on Table 2.

The membrane was then washed and incubated for 1 hour in a solution of 3% BSA in TBS containing the appropriate secondary antibody (Table 3). The detection of the chemiluminescent signal was obtained by adding 1 ml of SuperSignal™ West Pico PLUS Chemiluminescent Substrate (Thermo Fisher Scientific – 34580). Images were obtained and saved using the ChemiDoc™ XRS+ Imaging System and the Image Lab™ Software from Bio-Rad.

Primary antibody	Working dilution	Company	Origin
β-Actin	1:2,000	Thermo Fisher Scientific (AM4302)	Mouse
H3	1:3,000	Abcam (AB1791)	Rabbit
osTIR1	1:1,000	MLB (PD048)	Rabbit
P53	1:1,000	Santa Cruz (SC-126)	Rabbit
SPRTN	1:500	Atlas (HPA025073)	Rabbit
V5	1:1,000	Novex (R96025)	Mouse
Vinculin	1:1,000	Santa Cruz (SC-73614)	Mouse

Table 2. Primary antibodies used for western blot.

Secondary antibody	Working dilution	Company	Origin
Anti-Mouse IgG	1:10,000	Thermo Fisher Scientific (31430)	Goat
Anti-Rabbit IgG	1:10,000	Thermo Fisher Scientific (65-6120)	Goat

Table 3. Secondary antibodies used for western blot.

RT-qPCR

Quantitative reverse transcription polymerase chain reactions (RT-qPCR) were performed in order to confirm the depletion of SPRTN using the siSPRTN 3'UTR siRNA. This technique provides a reliable method to quantify the amount of mRNA in a given sample and further details are available here in the work of Nolan T et al. [196].

The protocol was divided into three main parts:

1. RNA extraction
2. RNA conversion to cDNA
3. RT-qPCR

RNA extraction

The total RNA was extracted using GeneJET RNA Purification Kit (#K0731) provided by Thermo Fisher Scientific.

In briefly, samples were resuspended in 600 µl of Lysis Buffer supplemented with β-mercaptoethanol, before passing the lysate through a 20-gauge RNAase-free syringe several times. Then, 360 µl of ethanol were added to each sample and mixed by pipetting.

The lysate was then inserted into the GeneJET RNA Purification Column and centrifuged for 1 minute at >12.000 g. The column was washed with washing buffers for two times, and again centrifuged for 1 minute at >12.000 g, discarding the flow-through. The purification column was then transferred into a sterile 1.5 ml RNAase-free microcentrifuge tube, and 100 µl of nuclease-free water was added to the purification column.

The RNA was eluted centrifuging the sample again for 1 minute at >12.000 g. The total amount of RNA extracted by the procedure was quantified using the NanoDrop™

One/OneC Microvolume UV-Vis Spectrophotometer from Thermo Fisher Scientific (ND-ONE-W).

RNA conversion to cDNA

The total RNA was converted to cDNA using the SuperScript™ III First-Strand Synthesis SuperMix for qRT-PCR (Invitrogen – 11752-050).

Up to 1 µg of RNA from each sample was mixed with 10 µl of 2X RT Reaction Mix and 2 µl of RT Enzyme Mix. Then, a final volume of 20 µl was reached by the addition of DEPC-treated water. Samples were gently mixed and incubated at 25 °C for 10 minutes, before increasing the incubation temperature to 50 °C for 30 minutes.

The reaction was terminated by warming up the samples to 85 °C for 5 minutes, before being chilled on ice. The RNA was then degraded by the addition of 1 µl of E. coli RNase H and incubation at 37 °C for 20 minutes, before the remaining cDNA could be used for qPCR.

RT-qPCR

The qPCR was performed using the LightCycler 480 SYBR Green 1 Master kit (Roche – 04707516001). The primers used for qPCR are reported in Table 4 and Table 5. For each reaction, 2.5 µl of forward and reverse primers (stock concentration: 100 µM), 12.5 µl of 2x Mastermix, 5 µl of cDNA and 2.5 µl of ultra-pure water were used to a final volume of 25 µl and placed in a 96 well plate (Roche – 04729692001).

The qRT-PCR was then performed using the thermocycler StepOnePlus Real-Time PCR System from Applied Biosystems, according to the conditions showed in Table 6. The results obtained from the RT-qPCR presented in this thesis (Figures 30) represent the mean result obtained from three independent biological replicates. To determine the

statistical significance of the findings, a comparison by Student's t-test was performed between the two experimental conditions (siLUC and siSPRTN-3’UTR) [193].

SPRTN			
	Name	Working concentration	Sequence
Forward Primer	SPRTN_Exon5_Fw	10 µM	AGATTAATGAGCACTTGGACTGG
Reverse Primer	SPRTN_3'UTR_Rv	10 µM	ACCAAAACACCTGACAGGATTCCC
Purpose	To confirm SPRTN depletion by siRNA		

Table 4. SPRTN primers used for qPCR.

GAPDH			
	Name	Working concentration	Sequence
Forward Primer	NM_002046_Fw	10 µM	GTCTCCTCTGACTTCAACAGCG
Reverse Primer	NM_002046_Rv	10 µM	ACCACCCTGTTGCTGTAGCCAA
Purpose	GAPDH was used as a control. GAPDH Human qPCR Primer Pair were provided by OriGene (HP205798)		

Table 5. GAPDH primers used for qPCR.

Step	Temperature (°C)	Time (seconds)
Initial denaturation	95	30
40 cycles	95	5
	58	15
	72	10
Final extension	72	120
Hold	4	∞

Table 6. Thermocycler settings for qPCR

Auxin-inducible degron

The generation of a stable auxin-inducible degron knock-in cell line for specific protein degradation in mammalian cells has been extensively studied in recent years [197]. In our case, we aimed to generate a stable cell line with the insertion of an auxin-inducible degron (AID) in endogenous SPRTN gene at C-terminal position, along with several markers for subsequent protein identification. In order to be functional, the recipient cell line should also express the TIR1 gene. HCT116 OsTIR1, in which the TIR1 gene from *Oryza sativa* was already inserted under a doxycycline-inducible promoter, was kindly provided to us by Ross Chapman's laboratory at the University of Oxford [189].

The outcome cell line will express endogenously the tagged SPRTN protein showed in Figure 20.

The protocol was divided into five main parts, as summarized in Figure 21:

1. Genome editing
2. FACS and clonal growth
3. Western blot screening
4. PCR screening of selected clones
5. Confirmation of SPRTN degradation by western blot

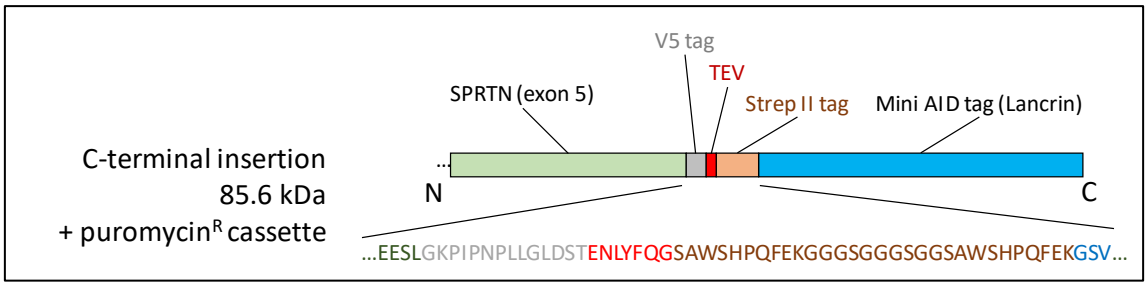


Figure 20. SPRTN C-terminal insertion in HCT116 osTIR1 cells.

Details of the C-terminal insertion performed on the endogenous SPRTN gene in HCT116 OsTIR1 cells. This included a V5 tag for western blot detection, a TEV tag for the removal of affinity purification tags and a Strep II tag for protein purification. Right after them, an AID tag was introduced to allow the auxin-mediated degradation of SPRTN.

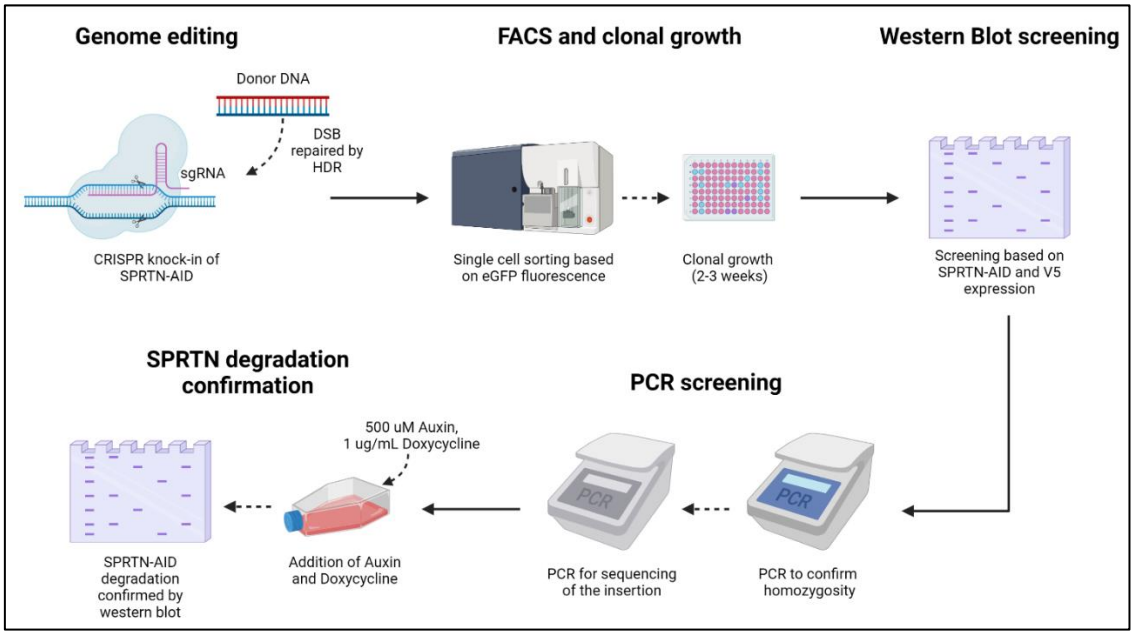


Figure 21. Auxin-inducible degron insertion workflow.

Genome editing (CRISPR-Cas9)

1.0 x10^6 HCT116 OsTIR1 cells were seeded in a well of a 6-well plate (Costar –3506) to a final volume of 2 ml. The following day, they were transfected with 5 µg of sgRNA plasmid and 10 µg of donor DNA in 200 µl of FBS-free OptiMEM (Gibco – 31985-047) and 45 µl of FuGene (Promega – E2311). After the mix was added to the cells, they were incubated for 2 days. The sgRNA-plasmid co-delivered scaffold RNA, Cas9 (S.p. wt) and a fluorescent tag for the FACS selection (eGFP), while the donor DNA contained the repair template DNA with the insertion cassettes (SPRTN Exon 5 + V5 + TEV + StrepII + AID + Stop + Puromycin cassette + 3'HOM).

The genome-editing DNAs were provided by the WIMM Genome Engineering Facility and are reported in Supplementary Table 2 and Figure 22. The sgRNA plasmid, in particular, was generated from the plasmid pSpCas9(BB)-2A-GFP (PX458) available on AddGene (https://www.addgene.org/browse/sequence/345552/), with the addition of the sgRNA sequence showed in Supplementary Table 2.

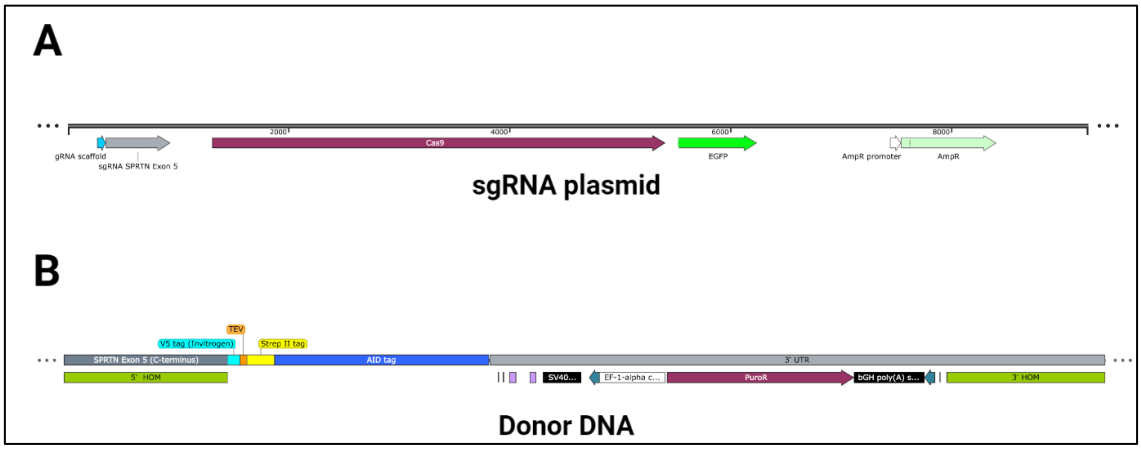


Figure 22. sgRNA plasmid and donor DNA for HCT116 SPRTN-AID cells.
(A) Structure of the sgRNA plasmid used for the creation of SPRTN-AID cell lines. (B) Structure of the donor DNA.

Fluorescence-activated cell sorting (FACS) and clonal growth

48 hours after transfection, cells were washed with PBS and trypsinized until a suspension of single cells was obtained. With the support of a technician (Jonathan Webber), cells were sorted into single cells by FACS according to their fluorescence provided by the eGFP signal contained in the sgRNA plasmid. This was performed on FacsAria III cell sorter using Diva 8.01, provided by the Kennedy Institute of Rheumatology at the University of Oxford. eGFP-positive cells were allowed to grow in 96-well plates for 2-3 of weeks, until each surviving clone could expand and provide enough material for further testing.

Western blot screening

Samples from each obtained clone were collected and tested according to the protocol described in the "Material and Methods – Western blot" section of this thesis. To summarize, clones were screened and selected according to their expression of SPRTN-AID (85 kDa, while SPRTN wild-type is only 55 kDa) and the presence of the V5 tag (85 kDa, a useful marker for protein identification inserted between SPRTN Exon 5 and the AID tag). Desired clones were selected according to their expression of both SPRTN-AID and V5, and the missing signal observed for SPRTN WT. Those clones were selected for further screening by PCR, to confirm their homozygosity and that the desired insertion was obtained.

PCR screening of selected clones

The clones which were considered more promising during the initial round of screening by western blot were further expanded and tested by PCR. This included two different PCRs, the first one to confirm the homozygosity of the selected clones (2511 bp) and

the second one, shorter, to be used for sequencing (502 bp). A description of the primer used for each PCR is offered in Table 7, Table 8 and in Figure 23.

To summarize, cells from a 10 cm plate were harvested and suspended in 300 µl of DirectPCR Lysis Reagent (Viagen Biotech – 301-C) and 0.4 mg/ml Proteinase K (Sigma – P6556) in a rotating hybridization oven at 55 °C for 6 hours to obtain a complete lysis. Crude lysates were then warmed at 85 °C for 45 minutes before they could be used for PCR reactions (1 µl of crude lysate every 50 µl PCR reaction is recommended).

The PCR reactions were set up using the primers previously described (Table 7, Table 8) and according to the Q5 High-Fidelity 2X Master Mix (Biolabs - NEB #M0515) protocol, summarized in Table 9 and Table 10. This was performed using the Veriti 96 Well Thermal Cycler from Applied Biosystem.

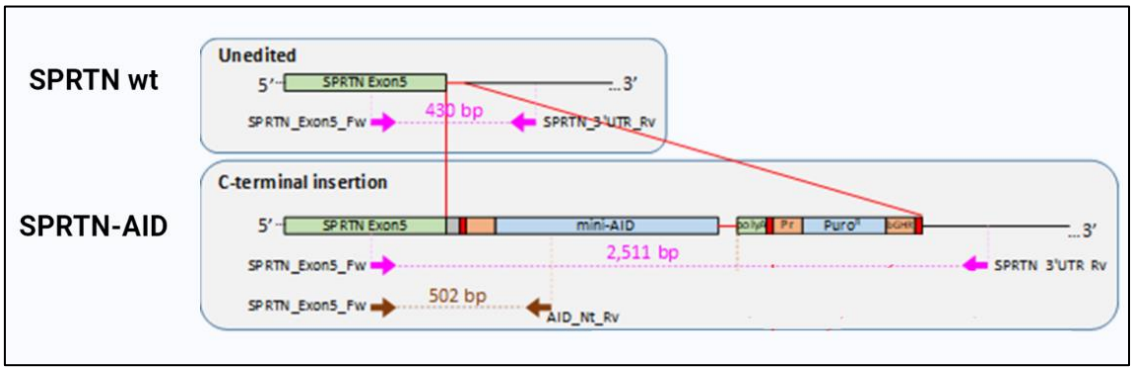


Figure 23. DNA fragments obtained from PCR

DNA fragments obtained from the PCR reactions used to screen SPRTN-AID clones.

First PCR			
	Name	Working concentration	Sequence
Forward Primer	SPRTN_Exon5_Fw	500 nM	AGATTAATGAGCACTTGGACTGG
Reverse Primer	SPRTN_3'UTR_Rv	500 nM	ACCAAAACACCTGACAGGATTCCC
Purpose	To identify homozygous clones for SPRTN-AID		
Comments	This PCR reaction can produce two different products: a smaller product (430 bp) for SPRTN wt and a longer product (2511 bp) for SPRTN-AID. Accordingly, homozygous clones will be characterized by the presence of a single band corresponding to 2511 bp.		

Table 7. Primers used for first PCR.

Second PCR			
	Name	Working concentration	Sequence
Forward Primer	SPRTN_Exon5_Fw	500 nM	AGATTAATGAGCACTTGGACTGG
Reverse Primer	AID_Nt_Rv	500 nM	ACGTTCTTTCTGTAGCTGCG
Purpose	Sequencing		
Comments	This PCR reaction was set up to produce a shorter fragment (502 bp) which could then be sequenced to confirm the correct C-terminal insertion.		

Table 8. Primers used for second PCR.

Component	50 µl reaction	Final concentration
Q5 High-Fidelity 2X Master Mix	25 µl	1x
10 µM Forward Primer	2.5 µl	500 nM
10 µM Reverse Primer	2.5 µl	500 nM
Template DNA	1 µl	Variable
Nuclease-free water	19 µl	-

Table 9. PCR reaction components.

Step	Temperature (ºC)	Time (seconds)
Initial denaturation	98	30
30 cycles	98	10
	65	30
	72	30
Final extension	72	120
Hold	4	∞

Table 10. Thermocycler conditions for PCR.

Confirmation of SPRTN degradation by western blot

Homozygous SPRTN-AID clones were further expanded and tested to determine their ability to correctly degrade the SPRTN-AID protein after the addition of 500 µM of auxin. Since this effect is dependent on the correct expression of OsTIR1, cells were pretreated with 1 µg/ml doxycycline for 24 hours to ensure a sufficient expression of OsTIR1.

The following day, cells were treated with 500 µM auxin and the depletion of SPRTN-AID was determined by western blot at different time points (30 minutes, 3 hours, 6 hours). Furthermore, the time required to restore SPRTN-AID expression after auxin washout was also tested at 12 hours, 24 hours, 48 hours. More detailed information is available in the “Material and Method – Western blot” section of this thesis.

Cell cycle analysis

The effect of SPRTN depletion on the cell cycle of HCT116 and HeLa cells was analyzed by flow cytometry by using a dual-labeling experiment, as previously described [198]. The DNA of proliferating cells was labeled with 5-Ethynyl-2'-deoxyuridine (EdU, a thymidine analog), detected with a Click-iT EdU assay and subsequently stained with propidium iodide (PI, a DNA dye) before being analyzed. The protocol can be divided into three main parts:

1. Staining and preparation of samples
2. Flow cytometric analysis
3. Quantification of the results

Staining and preparations of samples

Cells were seeded in 10 cm dishes and treated according to the required experimental conditions (none for HeLa; addition of 1 µg/ml doxycycline ± 500 uM auxin for HCT116 SPRTN-AID) for 48 hours. After this, they were incubated with 10 µM EdU for 1 hour to provide enough time for actively replicating cells to incorporate the EdU. before being washed with PBS for two times. Cells were washed two times with PBS, collected and resuspended in fixing buffer (ice cold methanol) for 30 minutes before being centrifuged for 5 minutes at 1000 g and resuspended in FACS buffer (0.25% Saponin, 1% BSA in PBS) for two times.

Cells were then treated according to the Click-iT™ EdU Alexa Fluor™ 488 Flow Cytometry Assay Kit (Termo Fisher Scientific - C10420) protocol. The Click-iT EdU assay is based on click chemistry, were a covalent reaction between an azide (coupled to Alexa Fluor™ 488) and an alkyne (found in the ethynyl moiety of EdU) is catalyzed in the presence of copper, as showed in Figure 24.

Briefly, cells were suspended in 250 µl of click reaction cocktail (Table 11) for 30 minutes in the dark. Then, they were centrifuge and washed two times with FACS buffer before

being resuspended in 250 µl of propidium solution (FACS buffer with the addition of 20 µg/ml of propidium iodide and 50 µg/ml ribonuclease A). Samples were once again incubated in the dark, for 20 minutes, to allow enough time for the fluorescent propidium iodide (PI) to intercalate between the DNA bases. While the click reaction enables the identification of replicative S-phase cells, the PI staining provide a measurement of the quantity of DNA in each cell (which doubles in G2/M phases when compared to G0/G1 phases). Cells were then washed with FACS buffer and resuspended in 300 µl of PBS before proceeding to flow cytometric analysis.

Flow cytometric analysis

For the analysis, the flow cytometer CytoFLEX (Beckman Coulter) and the program CytoExpert were used. At least 12,000 separate events were recorded for each sample, and the gating strategy which was used is reported in Figure 25 and Table 12.

Quantification of the results

The results obtained from the analysis of EdU and PI fluorescence were saved and analyzed using FlowJo, to determine the percentage of cells within each phase of the cell cycle. Samples were run in triplicate and the same gating strategy, previously described, was applied to all samples. The graphs were then saved and exported for publication. The comparison between the percentage of cells in each phase of the cell cycle (G0/G1, S, G2/M) according to the different experimental conditions (with or without SPRTN depletion) was performed by Student's t-test (Figure 38) [193].

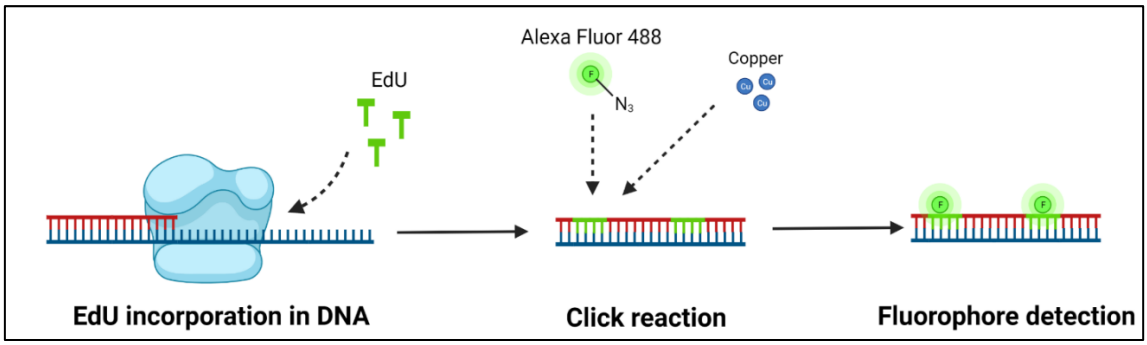


Figure 24. Click-reaction for S-phase cells.

Components of click reaction cocktail	Volume needed for preparing 1000 uL
PBS	875 µl
Component F (CuSO4)	20 µl
Alexa Fluor™ 488 azide	5 µl
Reaction buffer additive	100 µl

Table 11. Click-reaction components.

Component of the click-reaction cocktail were provided by Click-iT™ EdU Alexa Fluor™ 488 Flow Cytometry Assay Kit (Thermo Fisher Scientific - C10420).

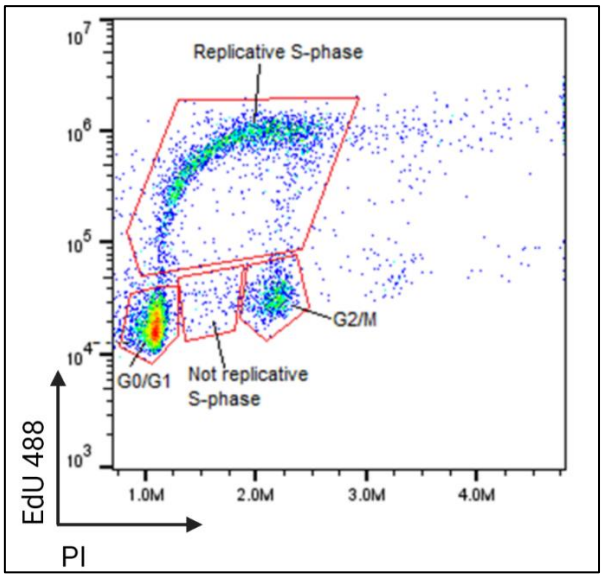


Figure 25. Cell cycle analysis.

Cell cycle phases recognized according to the fluorescent signals provided by EdU and PI.

Gate	Parameters	Purpose	Comments
P1	SSC-A / FSC-A	Exclude debris	Forward (FSC) and side (SSC) scatter give an estimation of the size and granularity of the cells. Debris tend to have lower FSC and SSC and they are found at the bottom left corner of the graph.
P2	FSC-H / FSC-A	Exclude non-single cells	The relationship between the forward scatter peak height (FSC-H) and the area under the peak (FSC-A) correlates as long as the cells are not clumped together. Accordingly, single cells appear on a diagonal while selected according to these parameters.
P3	EdU-488 / PI	Cell cycle analysis	After having excluded debris and non-single cells, it is now possible to analyze the signal obtained from EdU and PI. An example of the fluorescent signal and its interpretation is provided in Figure 25.

Table 12. Flow cytometry gating strategy.

Immunofluorescence

The immunofluorescence analysis of γH2AX and 53BP1 foci after DNA damage has been reported to be a reliable method to examine DNA double-strand breaks formation, as showed by Popp HD et al. [199]. This protocol is characterized by two main parts:

- Processing of adherent cells on coverslips
- Imaging

Processing of adherent cells on coverslips

Cells were grown in 12-wells plates (Costar - 3512) on sterile coverslips (Epredia – CB0009RA120MNZ0) until 70% confluent. Cells were treated according to the required experimental conditions, and then washed with PBS before being fixed with a 4% formaldehyde (Sigma-Aldrich - 252549) in PBS solution for 10 minutes at room temperature. Then, the coverslips were washed again 3 times with PBS before being permeabilized with extraction buffer (Table 13) for 5 minutes at room temperature and washed with PBS 3 times.

Samples were then blocked with a 5% BSA (Sigma-Aldrich - 9048-46-8) in PBS solution for 1 hour at 37 ^{o}C, before being incubated with the desired primary antibodies prepared in a 2.5% BSA in PBS solution. 20 µl of diluted primary antibody (Table 14) were placed on parafilm and the coverslips were positioned over them at room temperature for 1 hour, in the dark. The coverslips were then washed 3 times with PBS before being incubated, using the same procedure, with the desired secondary antibodies (Table 15) in a 2.5% BSA in PBS solution for 1 hour.

After the incubation with the secondary antibody was completed, nucleuses were stained with 300 nM DAPI (Thermo Fisher Scientific - D1306) in PBS for 5 minutes. The coverslips were then washed 3 times with PBS and 3 times in H_2O before being mounted on appropriate slides (Thermo Scientific – BS7011/2) for microscopy, with 7.5 µl of

Fluoromount-G mounting medium (Invitrogen - 00-4958-02). Samples were dried overnight before being analyzed.

Components of extraction buffer	Final concentration
HEPES	10 mM
NaCl	100 mM
Sucrose	300 mM
$MgCl_2$	3 mM
Triton X-100	0.25%
PBS	Up to the required volume

Table 13. Extraction buffer for immunofluorescence.

Primary antibody	Working dilution	Company	Origin
53BP1	1:250	Cell Signaling (4937)	Rabbit
ɣ-H2AX (Ser 139)	1:1000	Millipore (05-636)	Mouse

Table 14. Primary antibodies for immunofluorescence.

Secondary antibody	Working dilution	Company	Origin
Alexa Fluor 488 anti-rabbit IgG	1:500	Thermo Fisher Scientific (A-11034)	Goat
Alexa Fluor 555 anti-mouse IgG	1:500	Thermo Fisher Scientific (A-31570)	Donkey

Table 15. Secondary antibodies for immunofluorescence.

Imaging

Immunofluorescence images were taken using a Carl Zeiss LSM 7XX confocal microscope utilizing a 40x / 1.3 NA oil immersion objective with a pixel size of 0.208 µm in both X and Y planes. The blue fluorophore was excited using 5% power of the 405 nm laser and emission was collected 410 – 513 nm range, the gain for the system was set at 498. The green fluorophore was excited using 5% power of the 488 nm laser and emission was collected 595 – 612 nm range, the gain for the system was set at 615. The red fluorophore was excited using 10% power of the 561 nm laser and emission was collected 585 – 733 nm range, the gain for the system was set at 475.

The number of 53BP1 and γ-H2AX foci with and without SPRTN depletion at different timepoints were calculated, and the statistical significance between different experimental conditions was determined by One-way ANOVA (Figure 37) [200]. The experiments was independently repeated three times, providing comparable results.

RNA-seq

RNA sequencing (RNA-seq) is a well-known method, first developed in 2008, that relies on deep-sequencing technologies to study the transcriptome of different cells [201-203]. The possibility to analyze the whole-genome transcriptome of cells under different conditions has further expanded our possibilities to understand how cells can respond and adapt to different stimuli, and a range of companies now offer support to laboratories interested in performing a transcriptomic analysis.

Since our laboratory unfortunately did not have access to the required expertise and facilities to perform RNA-seq, we decided to rely on the services provided by an external company: BGI Genomics (https://www.bgi.com/global). This Chinese genomics company was originally founded in 1999 in order to support the global efforts of the Human Genome Project, and has since specialized in DNA and RNA sequencing. Accordingly, after treating the cells as required and extracting the RNA to be sequenced, samples were shipped to BGI Genomics to perform the initial transcriptomic analysis, as described in Figure 26.

The protocol was divided into two main parts:

- RNA extraction and quality assessment
- RNA-seq

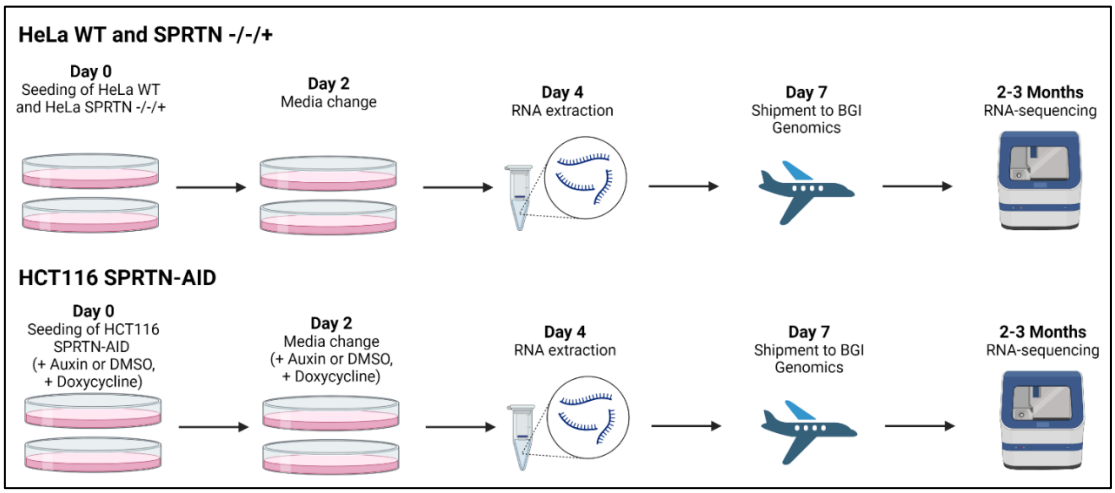


Figure 26. RNA-seq workflow.

Workflow of the RNA-seq experiment, performed in collaboration with BGI Genomics.

RNA extraction and quality assessment

The RNA extraction was performed using the GeneJET RNA Purification Kit (#K0731) provided by Thermo Fisher Scientific. Please refer to the description provided on the "Material and Methods – RT-qPCR" section of this thesis for further details.

RNA samples were then analyzed using NanoDrop™ One/OneC Microvolume UV-Vis Spectrophotometer from Thermo Fisher Scientific (ND-ONE-W), to determine if their quality was good enough to proceed for RNA-seq. All tested samples had a concentration >10ng/µl, an OD260/280 ratio > 2.0, an OD260/230 ratio >2.0, and they were deemed suitable to proceed. Samples were then stored at -80 ^{o}C for 48 hours until ready to be shipped to BGI Genomics.

RNA-seq

Samples were analyzed according to the indicated experimental conditions (two different cell lines, each with or without the expression of SPRTN), and each condition had 3 independent biological replicates (Table 16). Accordingly, a total of 12 samples were tested using the DNBSEQ platform, with an average yield of 4.49G data per sample.

In order to reflect the correlation of gene expression between biological replicates, the Pearson correlation coefficients of all gene expression between each two samples were calculated, and these coefficients were reflected in the form of a heatmap in Figure 27A-B. The correlation coefficients can reflect the similar situation of the overall gene expression between each sample. The higher the correlation coefficient is, the more similar the gene expression level is. Next, a principal component analysis (PCA) was conducted to reduce the large amount of gene expression information contained in each sample to a few unrelated principal components, for comparison between samples and identification of outliers. The PCA is showed in Figure 28 and it is a multivariate statistical analysis method that reduces multiple variables to a small number of independent variables (that is, principal components) by dimensionality reduction, while retaining as much of the original data information as possible.

Cell line	SPRTN status	Number of replicates
HeLa	Normal (SPRTN wt)	3
	Depleted (SPRTN -/-/+)	3
HCT116 SPRTN-AID	Normal (DMSO)	3
	Depleted (Auxin)	3

Table 16. Experimental conditions and replicates used for RNA-seq.

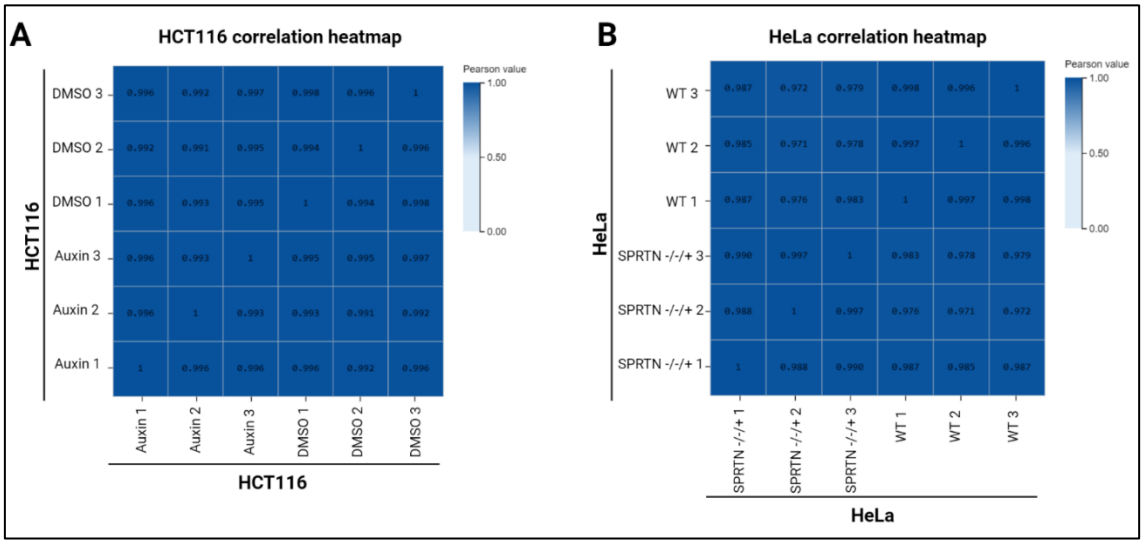


Figure 27. RNA-seq correlation heatmaps for HCT116 and HeLa cells.

(A) Correlation heatmap of HCT116 cells, with and without SPRTN depletion (B) Correlation heatmap of HeLa cells, with and without SPRTN depletion.

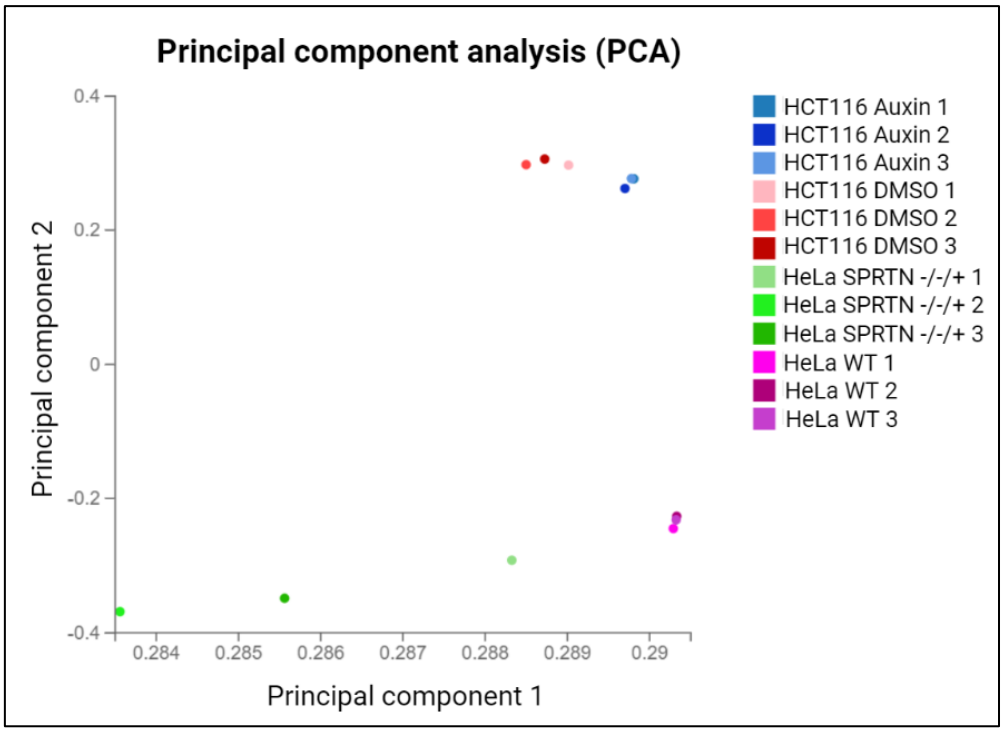


Figure 28. PCA analysis across replicates.

Immunohistochemistry and staining assessment

In order to perform the survival analyses of CRC patients based on SPRTN protein expression, we relied on tumor samples available from the MRC FOCUS trial [204]. This trial enrolled 2135 unpretreated CRC patients, which were randomized and allocated to three different possible treatments (5-fluorouracil alone, FOLFIRI or FOLFOX). Our analysis included an initial exploratory cohort containing tissue slides obtained from 45 patients treated with FOLFIRI, and an expanded cohort of tissue microarrays (TMAs) obtained from 914 patients treated with 5-fluorouracil alone, FOLFIRI or FOLFOX. Clinical samples were obtained through the support of the FOCUS Trial Investigators and, in particular, professor Timothy Maughan at the University of Oxford.

All the samples, independently of their cohort, were processed with the support and supervision of the Translational Histopathology Laboratory (THL) at the University of Oxford, according to the following steps:

- Immunohistochemical staining
- Digital pathology (HALO)

Immunohistochemical staining

The immunohistochemical staining of formalin-fixed, paraffin-embedded (FFPE) colorectal cancer samples was performed by Leticia Campo, a senior THL technician. We relied on chromogenic immunohistochemistry (IHC), a technique that take advantage of enzymes that convert soluble substrates into insoluble chromogenic products, to visualize the presence of antigens in tissue [205, 206].

In brief, automated staining was carried out with the Leica BOND-MAX autostainer (Leica, Microsystem) using the following conditions:

- Antigen retrieval at 100ºC for 20 min with Epitope Retrieval Solution 2 (AR9640, Leica Biosystems), equivalent to Tris-EDTA Buffer pH 9.0. This solution was designed to separate protein aggregates and unmask the antigen and epitopes present in FFPE tissues, to enhance the staining intensity of the primary antibodies.
- Primary antibody incubation with the indicated antibodies (Table 17) for 1 hour.
- Detection using the BOND™ Polymer Refine Detection System (DS9800, Leica Biosystems), an anti-rabbit polymer conjugated to horseradish peroxidase (HRP), as per manufacturer's instructions.

Primary antibody	Working dilution	Company	Origin	Comments
SPRTN	1:100 (Tissue) 1:125 (Cell lines)	Atlas (HPA025073)	Rabbit	Used for the exploratory cohort (45 patients)
SPRTN	1:75 (Tissue) 1:75 (Cell lines)	Abcam (ab121853)	Rabbit	Used for the expanded cohort (914 patients)

Table 17. Primary antibodies for IHC staining.

Digital pathology (HALO)

The unbiased quantitative tissue staining analysis was performed using HALO, an image analysis platform specialized in the study of pathological tissues which enables tissue segmentation by artificial intelligence [207, 208]. The quality of the tissue segmentation and of the immunohistochemical staining was further assessed and confirmed by Alistair Eason, a THL consultant pathologist.

The preliminary study on the exploratory cohort was performed using HALO Image Analysis Platform version 3.4.2986 (Indica Labs, Inc.), according to the current parameters: TMA Module, Multiplex IHC v3.0.4, DenseNet AI (plugin). This analysis provided a measure of the percentage of SPRTN-positive cells within the malignant epithelium of the 45 CRC patients treated with FOLFIRI.
The final study on the expanded cohort was performed using HALO® Image Analysis

Platform version 3.4.2986 (Indica Labs, Inc.), according to the current parameters: TMA Module, Multiplex IHC v3.1.4., Nuclei Seg (plugin) – BF v1.0.0. This analysis provided a measure of the nuclear optical density (nuclear OD) associated with SPRTN expression within the malignant epithelium of 914 CRC patients. In this second analysis, only the nuclear signal was considered as it provided a more reliable assessment of SPRTN expression, as showed by the staining performed on HCT116 cells expressing different levels of SPRTN (Figure 43). This had not been possible during the preliminary study as those cell lines were not yet available when the exploratory cohort was analysed.

Survival analyses

In order to perform the survival analyses, the unbiased tissue staining assessment provided by HALO was associated with the clinical and molecular data available for each patient. Patients were then divided in tertiles based on SPRTN expression (High, Medium and Low), and the three groups were compared according to their clinical and molecular variables as reported in Table 18.

Kaplan-Meier curves were plotted on GraphPad v9.4.1 and curves were compared using cox proportional hazard regression. In particular, progression free survival (PFS) was considered the primary objective, as it represented the most likely endpoint in which the influence of SPRTN expression could have been observed without interference from the effects of later alternative management strategies. This was agreed by professor Tim Maughan prior to any inspection of the data or analyses. Overall survival (OS) was instead considered a secondary objective, as patients with clinical progression confirmed according to RECIST criteria [209] were then considered for second-line treatments. Analyses were adjusted according to the secondary clinical and molecular variables reported in Table 18.

Cox proportional hazard regressions were also used to investigate possible interactions between SPRTN expression and treatments efficacy, and to perform multivariate analyses to determine the influence of the reported secondary variables on the final outcomes. The proportionality of the hazards within the entire cohort of 914 patients was confirmed by the group of professor Jenny Seligmann at the University of Leeds, through the use of Schoenfeld residuals test (data not shown). As the data obtained by SPRTN immunohistochemical staining was normally distributed (data not shown), it was considered appropriate to divide the patients in tertiles while performing the survival analyses.

Main variables	Category	Options
SPRTN expression	Molecular	High, Medium, Low
Treatment	Clinical	5-Fluorouracil, FOLFIRI, FOLFOX
Progression Free Survival (PFS)	Clinical	-
Overall Survival (OS)	Clinical	-
Tumour response (12 weeks)	Clinical	CR, PR, SD, PD
Secondary variables	Category	Options
Age	Clinical	-
Sex	Clinical	Male, Female
ECOG Performance Status	Clinical	0-1, 2
Previous resection	Clinical	Yes, No
Peritoneal Metastases	Clinical	Yes, No, Unknown
Lung Metastases	Clinical	Yes, No, Unknown
Liver Metastases	Clinical	Yes, No, Unknown
Nodal Metastases	Clinical	Yes, No, Unknown
Primary tumor location	Clinical	Rectum, Left, Right
KRAS status	Molecular	Wild type, Mutated, Unknown
BRAF status	Molecular	Wild type, Mutated, Unknown

Table 18. Clinical and molecular variables of CRC patients.

Clinical and molecular variables taken into consideration for survival analyses. This data is associated to patients enrolled in the MRC FOCUS trial [204]. Tumour response is indicated as complete response (CR), partial response (PR), stable disease (SD) or progressive disease (PD)

Chapter 3: Results

3.1 SPRTN role in mediating the sensitivity to cytotoxic chemotherapy in CRC

Since the role of SPRTN in the repair of DPCs has been first suggested in 2016 [136, 158, 164], different studied have demonstrated SPRTN's potential to influence the efficacy of chemotherapy, both in cell lines and in more complex organisms [136, 158, 165]. In particular, it has been confirmed that SPRTN depletion increases the cytotoxicity of DPC-inducing compounds like formaldehyde, camptothecin and etoposide in both HeLa cells and in lymphoblastoid cells obtained from RJALS patients [158, 165]. The RJALS syndrome, firstly identified in three patients in 2014, is a rare genetic syndrome characterized by biallelic germline mutations of SPRTN, along with increased genomic instability, progeroid features and a predisposition toward the development of early onset hepatocellular carcinoma [165]. These results were also further confirmed in *C. elegans* where, along with an increase sensitivity to formaldehyde, mutant strains lacking functional SPRTN also showed to be more sensitive to cisplatin, a platinum compound frequently used to treat cancer patients [136]. Importantly, hypersensitivity to camptothecin and elevated levels of TOP1ccs in cells from SPRTN hypomorphic mice further suggest the role of SPRTN in the processing of TOP1ccs lesions induced by topoisomerase inhibitors [160].

In order to have a better understanding of the role of SPRNT in patients affected by cancer, a range of publicly available databases were first investigated to determine how frequently SPRTN is altered in human cancers, and to investigate the role that SPRTN can have in assuring the survival of cancer cells. These questions were first addressed by analyzing the data obtained from Project Achilles (https://depmap.org/portal/achilles/), a database containing the results of lentiviral-based pooled RNAi or CRISPR/Cas9 libraries of genome-scaled loss-of-function (LOF) screenings of numerous cancer cell lines [210]. In particular, cancer cell lines belonging to the four most common type of cancer (breast, lung, colorectal and prostate cancer) were analyzed to determine their

SPRTN Chronos dependency score, a measurement of the effect that SPRTN knockout can have on their survival [211]. Unsurprisingly, SPRTN appeared to be a commonly essential gene (Figure 29A), in line with previous reports showing how the complete knockout of SPRTN or morpholino-mediated SPRTN silencing causes embryonic lethality in mice and zebrafish, respectively [160, 165].

Next, the publicly available Cancer Genome Atlas Project (TCGA) dataset, a landmark cancer genomic study originally developed in 2013 [212] was downloaded from cBioPortal (https://www.cbioportal.org/) and investigated to determine if SPRTN is commonly altered in cancer patients and if SPRTN expression is able to influence the survival of colorectal cancer patients, which represents the main focus of this thesis. Initially, SPRTN alterations were investigated based on the genomic data obtained from TCGA regarding the four most common type of cancer (Figure 29B). This allowed to determine that, except in the case of breast cancer patients where the SPRTN gene is amplified in 9% of cases, most cancers do not present significant alterations of SPRTN. Colorectal cancers, in particular, recorded SPRTN mutations in only 2% of the 594 analyzed cases (Figure 29B).

The RNA-seq data collected from the colorectal cancer patients available on TCGA (588 cases) was then used to divide the patients in two main groups, according to their level of SPRTN expression (Figure 29C). RNA-seq represents a better choice to determine SPRTN expression in human tissue when compared to DNA microarrays, as SPRTN transcripts are scarce and generally below the sensitivity allowed by the use of DNA microarrays (data not shown, originally reported by Sylvana Hassanieh from the Maughan Laboratory at the University of Oxford) [213]. The overall survival of CRC patients expressing high or low level of SPRTN mRNA was then compared, and no statistically significant difference could be observed (Figure 29D, p-value=0.89)

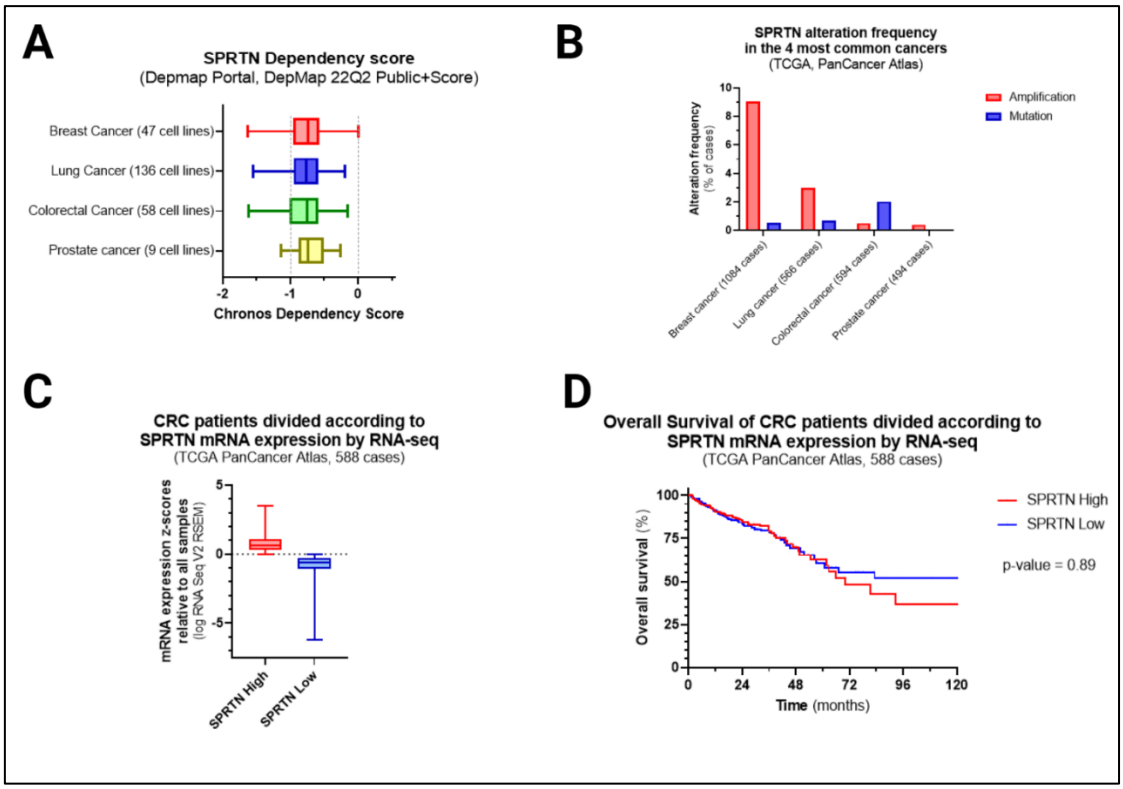

A
SPRTN Dependency score
(Depmap Portal, DepMap 22Q2 Public+Score)
Breast Cancer (47 cell lines)
Lung Cancer (136 cell lines)
Colorectal Cancer (58 cell lines)
Prostate cancer (9 cell lines)
-2
-1
0
Chronos Dependency Score
B
SPRTN alteration frequency
in the 4 most common cancers
(TCGA, PanCancer Atlas)
Alteration frequency
(% of cases)
Amplification
Mutation
Breast cancer (1084 cases)
Lung cancer (566 cases)
Colorectal cancer (594 cases)
Prostate cancer (494 cases)
C
CRC patients divided according to
SPRTN mRNA expression by RNA-seq
(TCGA PanCancer Atlas, 588 cases)
mRNA expression z-scores
relative to all samples
(log RNA Seq V2 RSEM)
SPRTN High
SPRTN Low
D
Overall Survival of CRC patients divided according to
SPRTN mRNA expression by RNA-seq
(TCGA PanCancer Atlas, 588 cases)
Overall survival (%)
SPRTN High
SPRTN Low
p-value = 0.89
Time (months)

Figure 29. SPRTN role and alterations in cancer.

(A) SPRTN CRISPR knockout screen of common cancer cell lines obtained from project Achilles (https://depmap.org/portal/achilles/). A Chronos dependency score of 0 indicates that a gene is not essential, while a value of -1 represent the median of all essential genes. SPRTN is considered a commonly essential gene. (B) Percentage of SPRTN alterations (amplification or mutations) which are registered in the four most common type of cancer. (C) Colorectal cancer patients available from TCGA were divided in two groups, High and Low, according to the level of SPRTN mRNA expression recorded by RNA-seq (D) No statistically significant difference in overall survival can be observed in colorectal cancer patients according to different level of SPRTN expression (p-value=0.89). Survival curves were compared using the log-rank test.

The data used from Figure 29 (B-D) is publicly available from the National Cancer Institute website and can be downloaded through cBioPortal (https://www.cbioportal.org/).

As SPRTN has been reported to influence the efficacy of numerous chemotherapeutic drugs, including topoisomerase inhibitors [136, 158, 164], the next step involved determining if SPRTN depletion could influence the efficacy of the three main chemotherapeutic drugs used to in colorectal cancer patients: 5-Fluorouracil, Irinotecan and Oxaliplatin. This was assessed by colony forming assays, where SPRTN was depleted by siRNA according to the experimental setting described in Figure 30A. Three different colorectal cancer cell lines (HCT116, HT-55 and RKO) were used for these experiments, and SPRTN was initially depleted using siRNA against the 3'UTR of SPRTN's transcript. Unless otherwise specified, siSPRTN refers to the siSPRTN-3'UTR sequence. Two more siSPRTN sequences were used (siSPRTN-Q1, siSPRTN-Q2) to validate the results whenever SPRTN depletion seemed to suggest an increased drug efficacy.

It is worth noticing that the quantification of SPRTN protein expression by western blot can sometimes be challenging, due to scarcity of commercially available antibodies. Furthermore, SPRTN is a low abundant protein which needs to be extracted using urea buffer to aid its unfolding and its separation from chromatin, where it is generally tightly bound [214]. Accordingly, the efficacy of SPRTN depletion by siSPRTN-3'UTR was initially confirmed also through rt-qPCR (Figure 30B).

SPRTN role in mediating the efficacy of 5-Fluorouracil was first investigated in HCT116, HT-55 and RKO cells (Figure 30C-H). Unsurprisingly, SPRTN depletion by siRNA did not hypersensitize cells to 5-Fluorouracil. This is in line with previous reports that 5-Fluorouracil is mostly incorporated into RNA, and it is unlikely to be able to induce relevant amounts of DPCs [94].

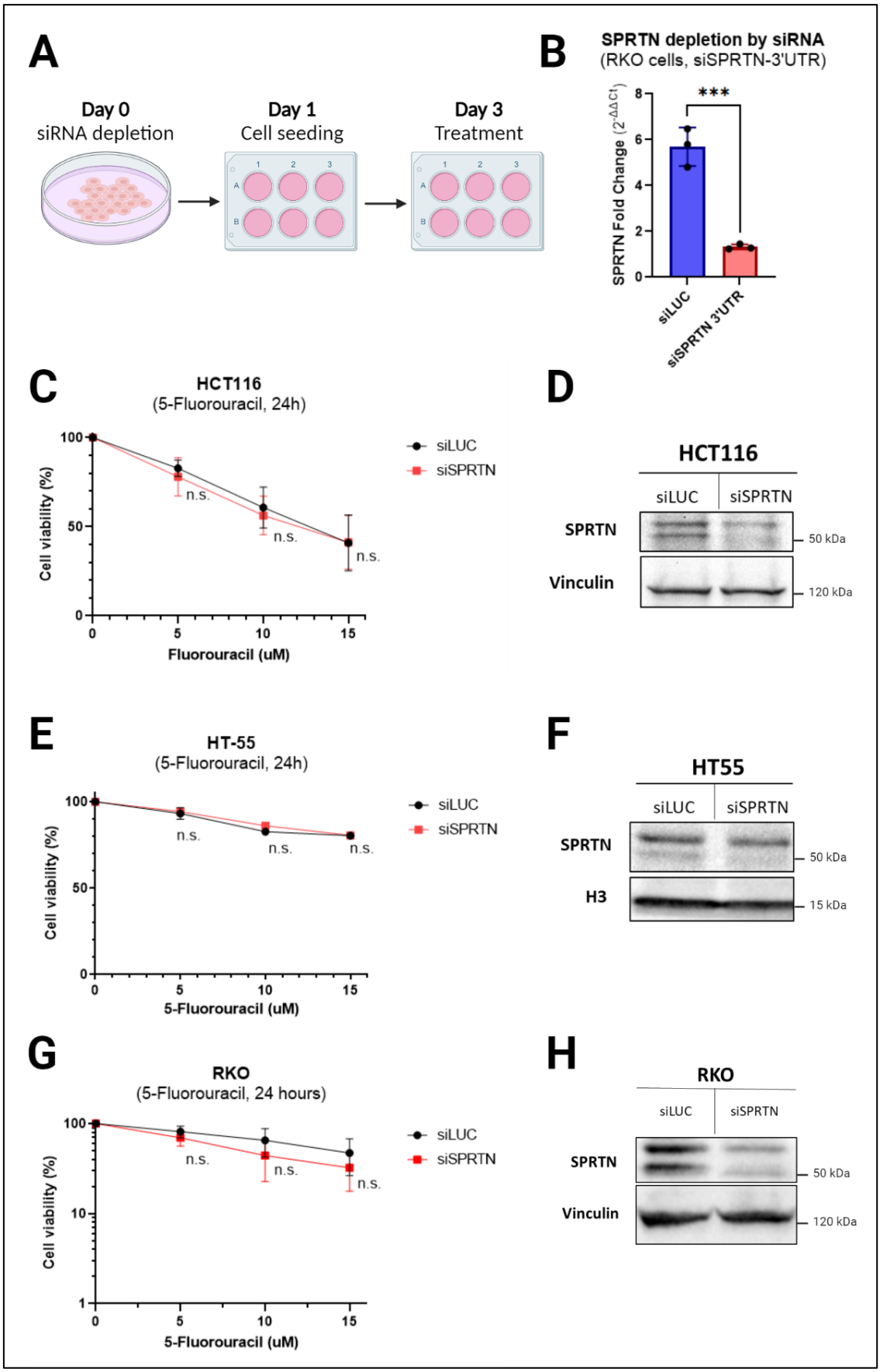

A
Day 0
siRNA depletion
Day 1
Cell seeding
Day 3
Treatment
B
SPRTN depletion by siRNA
(RKO cells, siSPRTN-3'UTR)
SPRTN Fold Change (2^-ΔΔCt)
***
siLUC
siSPRTN 3'UTR
C
HCT116
(5-Fluorouracil, 24h)
Cell viability (%)
Fluorouracil (uM)
siLUC
siSPRTN
n.s.
D
HCT116
siLUC
siSPRTN
SPRTN
Vinculin
50 kDa
120 kDa
E
HT-55
(5-Fluorouracil, 24h)
Cell viability (%)
5-Fluorouracil (uM)
siLUC
siSPRTN
n.s.
F
HT55
siLUC
siSPRTN
SPRTN
H3
50 kDa
15 kDa
G
RKO
(5-Fluorouracil, 24 hours)
Cell viability (%)
5-Fluorouracil (uM)
siLUC
siSPRTN
n.s.
H
RKO
siLUC
siSPRTN
SPRTN
Vinculin
50 kDa
120 kDa

Figure 30. SPRTN role in mediating the efficacy of 5-Fluorouracil.

(A) Cells were depleted with the indicated siRNA against SPRTN for three days, before being treated and analysed by colony forming assay. (B) The efficacy of SPRTN depletion by siSPRTN-3'UTR was confirmed through rt-qPCR. Unless otherwise stated, siSPRTN refers to siSPRTN-3'UTR. (C) HCT116 cells were treated with 5-Fluorouracil for 24 hours. For each condition, the results shows the mean of three independent biological replicate, and they were analyzed by t-test. No difference in survival can be observed between siLUC and siSPRTN. (D) Depletion check performed in HCT116 cells treated with the indicated siRNAs. (E) HT-55 cells were treated with 5-Fluorouracil for 24 hours. For each condition, the results shows the mean of three independent biological replicate, and they were analyzed by t-test. No difference in survival can be observed between siLUC and siSPRTN. (F) Depletion check performed in HT-55 cells treated with the indicated siRNAs. (G) RKO cells were treated with 5-Fluorouracil for 24 hours. For each condition, the results shows the mean of three independent biological replicate, and they were analyzed by t-test. No difference in survival can be observed between siLUC and siSPRTN. (H) Depletion check performed in RKO cells treated with the indicated siRNAs.

After confirming that SPRTN does not have a role in mediating the efficacy of 5-Fluorouracil, the next step aimed to test whether SPRTN inactivation affects the efficacy of topoisomerase-1 (TOP1) inhibitors, like irinotecan. Interestingly, while SPRTN depletion has been associated to an increase efficacy of camptothecin (an alkaloid compound from which Irinotecan is derived), no data has been published regarding the role of SPRTN in the repair of TOP1ccs induced by irinotecan [140, 164, 215]. This represents a potential challenge in a preclinical setting as irinotecan is a prodrug which first needs to be converted into its active metabolite, SN-38, by human carboxylesterase 1 (CES1), an enzyme mainly expressed in the liver tissue. Despite this, some studies suggest that a relevant expression of CES2, an alternative isoform of CES1, can also be found in 88% of colorectal cancer tissues, which suggests the idea that irinotecan can be converted in SN-38 directly by tumour cells [109].

To address the role of SPRTN in repairing Irinotecan-induced TOP1ccs, three colorectal cancer cell lines were depleted by siRNA and treated with the indicated concentrations of Irinotecan for 24 hours (Figure 31A-F). While none of them seemed to present an increased cytotoxicity due to SPRTN depletion, this phenotype could be potentially justified with the fact that Irinotecan conversion in SN-38 might be limited in these cell lines. To address this point, the experiments were repeated while treating the cells with Camptothecin, another TOP1 inhibitor which does not require to be processed by CES2 (Figure 32A-F) [164]. Interestingly, SPRTN depletion seemed to have a mild effect on HCT116 cells treated with Camptothecin (Figure 32A-B), but not on HT-55 and RKO cells (Figure 32C-F).

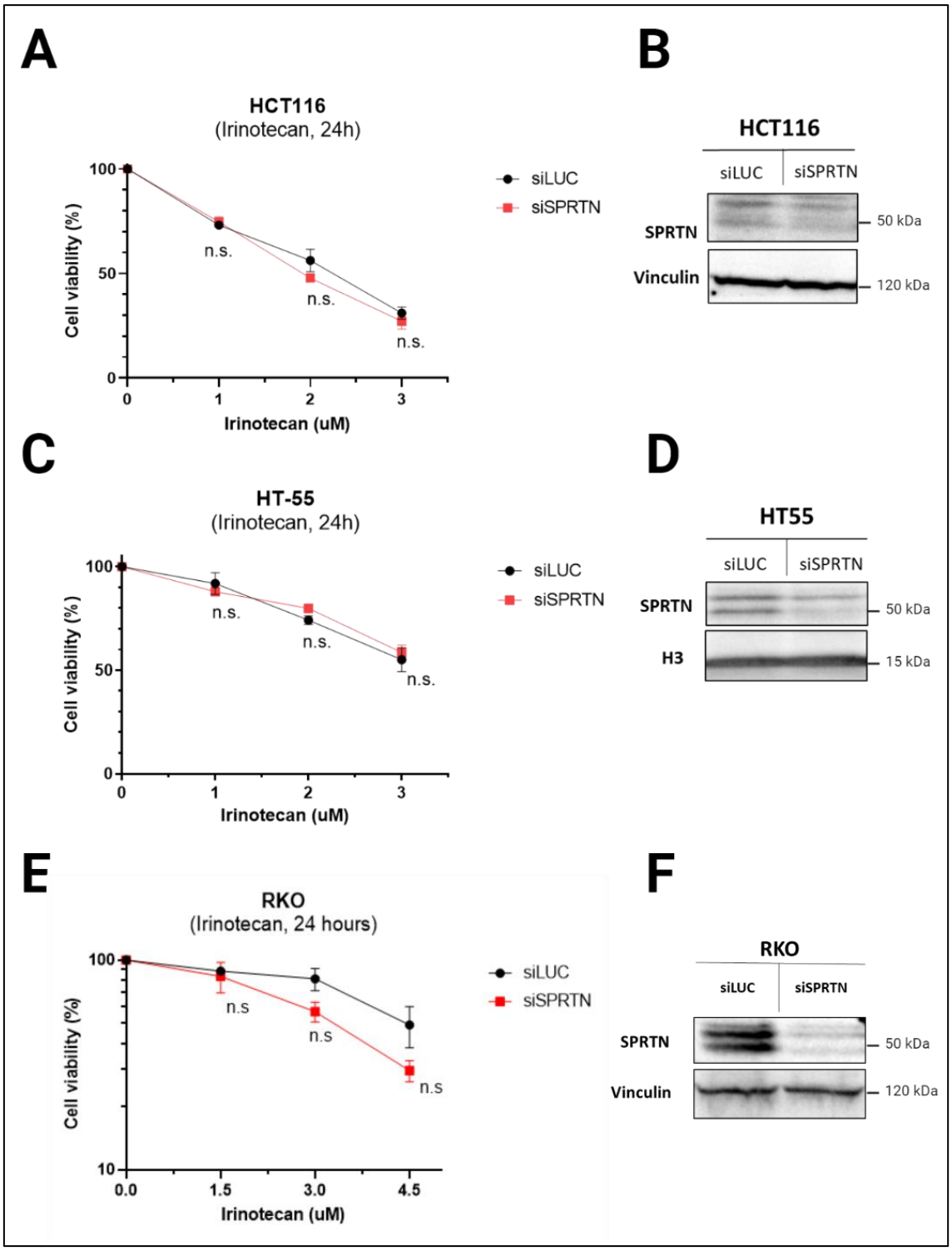

A
HCT116
(Irinotecan, 24h)
Cell viability (%)
Irinotecan (uM)
siLUC
siSPRTN
n.s.
B
HCT116
siLUC
siSPRTN
SPRTN
Vinculin
50 kDa
120 kDa
C
HT-55
(Irinotecan, 24h)
Cell viability (%)
Irinotecan (uM)
siLUC
siSPRTN
n.s.
D
HT55
siLUC
siSPRTN
SPRTN
H3
50 kDa
15 kDa
E
RKO
(Irinotecan, 24 hours)
Cell viability (%)
Irinotecan (uM)
siLUC
siSPRTN
n.s
F
RKO
siLUC
siSPRTN
SPRTN
Vinculin
50 kDa
120 kDa

Figure 31. SPRTN role in mediating the efficacy of Irinotecan.

(A) HCT116 cells were treated with Irinotecan for 24 hours. For each condition, the results shows the mean of three independent biological replicate, and they were analysed by t-test. No difference in survival can be observed between siLUC and siSPRTN. (B) Depletion check performed in HCT116 cells treated with the indicated siRNAs. (C) HT-55 cells were treated with Irinotecan for 24 hours. For each condition, the results shows the mean of three independent biological replicate, and they were analysed by t-test. No difference in survival can be observed between siLUC and siSPRTN. (D) Depletion check performed in HT-55 cells treated with the indicated siRNAs. (E) RKO cells were treated with Irinotecan for 24 hours. For each condition, the results shows the mean of three independent biological replicate, and they were analysed by t-test. No difference in survival can be observed between siLUC and siSPRTN. (F) Depletion check performed in RKO cells treated with the indicated siRNAs.

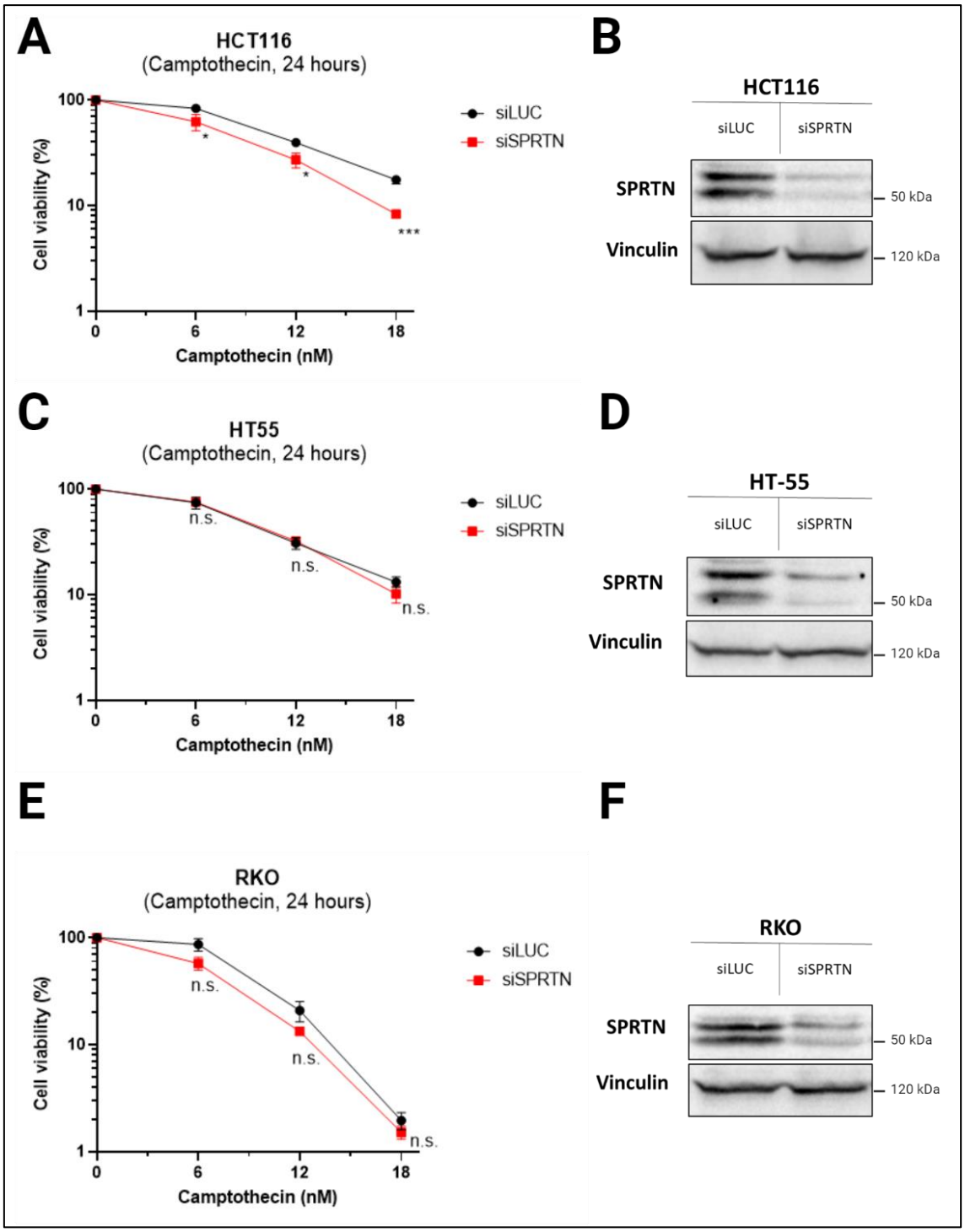

A
HCT116
(Camptothecin, 24 hours)
Cell viability (%)
100
10
1
0
6
12
18
Camptothecin (nM)
*
*
***
siLUC
siSPRTN
B
HCT116
siLUC
siSPRTN
SPRTN
50 kDa
Vinculin
120 kDa
C
HT55
(Camptothecin, 24 hours)
Cell viability (%)
100
10
1
0
6
12
18
Camptothecin (nM)
n.s.
n.s.
n.s.
siLUC
siSPRTN
D
HT-55
siLUC
siSPRTN
SPRTN
50 kDa
Vinculin
120 kDa
E
RKO
(Camptothecin, 24 hours)
Cell viability (%)
100
10
1
0
6
12
18
Camptothecin (nM)
n.s.
n.s.
n.s.
siLUC
siSPRTN
F
RKO
siLUC
siSPRTN
SPRTN
50 kDa
Vinculin
120 kDa

Figure 32. SPRTN role in mediating the efficacy of Camptothecin.

(A) HCT116 cells were treated with Camptothecin for 24 hours. For each condition, the results shows the mean of three independent biological replicate, and they were analysed by t-test. A mild difference in survival can be observed between siLUC and siSPRTN. (B) Depletion check performed in HCT116 cells treated with the indicated siRNAs. (C) HT-55 cells were treated with Camptothecin for 24 hours. For each condition, the results shows the mean of three independent biological replicate, and they were analysed by t-test. No difference in survival can be observed between siLUC and siSPRTN. (D) Depletion check performed in HT-55 cells treated with the indicated siRNAs. (E) RKO cells were treated with Camptothecin for 24 hours. For each condition, the results shows the mean of three independent biological replicate, and they were analysed by t-test. No difference in survival can be observed between siLUC and siSPRTN. (F) Depletion check performed in RKO cells treated with the indicated siRNAs.

Lastly, the role of SPRTN in mediating the cytotoxicity of oxaliplatin was investigated. This represented a particularly interesting area of scientific inquiry, as there are several reports that cisplatin, another platinum compound, is able to induce the formation of DPCs [216]. Moreover, SPRTN depletion has been confirmed to increase the efficacy of cisplatin both in cell lines and in more complex systems, like C. elegans [135, 136].

Previously, only one study has properly assessed the ability of oxaliplatin to induce DPCs. The amount of DPCs formed when isolated cell nuclei were exposed to equimolar amount of oxaliplatin and cisplatin was measured, and cisplatin was determined to be two-fold more efficient in the formation of DPCs when compared to oxaliplatin [105].

In order to assess the role of SPRTN in response to oxaliplatin, HCT116, HT-55 and RKO cells were treated with the indicated Oxaliplatin concentrations for 24 hours (Figure 33A-F). Surprisingly, HCT116 and RKO cells appeared to be more sensitive to Oxaliplatin after SPRTN depletion (Figure 33A-B, E-F), while this phenotype could not be replicated in HT-55 cells (Figure 33C-D). In fact, despite all being colorectal cancer cell lines, HCT116, HT-55 and RKO are characterized by largely different genetic backgrounds, which are summarized in Supplementary Table 1. One of the major differences between them which could have a role in determining this phenotype is represented by their TP53 status, which has been reported to be mutated only in HT-55 (Figure 34A). This is particularly interesting as p53 is normally activated after oxaliplatin-induced DNA damage and TP53 mutations have been associated with resistance to oxaliplatin in colorectal cancer cell lines (Figure 34B) [217].

To determine if this phenotype might be related to TP53 status, HCT116 TP53$^{-/-}$ cells were used [188]. In line with this hypothesis, TP53 knockout cells resulted to be more resistant to oxaliplatin, and insensitive to the effect of SPRTN depletion (Figure 34C-D). However, when the experiment was repeated using different siRNAs (siSPRTN-Q1, siSPRTN-Q2), SPRTN depletion did not seem to sensitize cells to oxaliplatin (Figure 34E-F). While this could be related to an insufficient level of SPRTN depletion induced

by the two alternative siRNAs, the possibility that the phenotype previously observed with siSPRTN-3'UTR might be caused by an off-target siRNA effect should also be considered. Accordingly, another system to assure an efficient SPRTN depletion was developed, and this system will represent the main focus of the next result chapter.

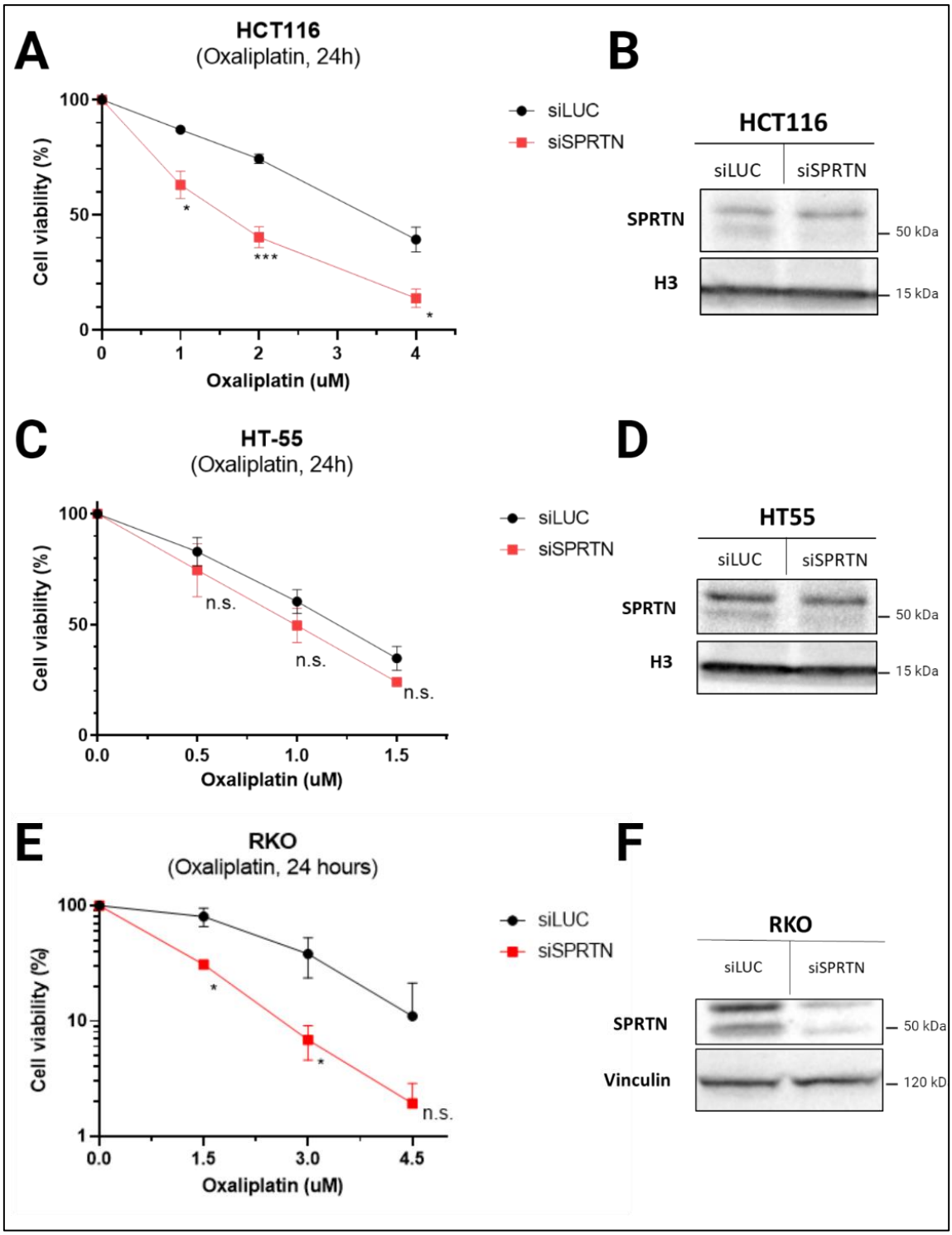

A
HCT116
(Oxaliplatin, 24h)
Cell viability (%)
Oxaliplatin (uM)
B
siLUC
siSPRTN
HCT116
siLUC
siSPRTN
SPRTN
H3
50 kDa
15 kDa
C
HT-55
(Oxaliplatin, 24h)
Cell viability (%)
Oxaliplatin (uM)
n.s.
D
siLUC
siSPRTN
HT55
siLUC
siSPRTN
SPRTN
H3
50 kDa
15 kDa
E
RKO
(Oxaliplatin, 24 hours)
Cell viability (%)
Oxaliplatin (uM)
n.s.
F
siLUC
siSPRTN
RKO
siLUC
siSPRTN
SPRTN
Vinculin
50 kDa
120 kD

Figure 33. SPRTN role in mediating the efficacy of Oxaliplatin.

(A) HCT116 cells were treated with Oxaliplatin for 24 hours. For each condition, the results shows the mean of three independent biological replicate, and they were analysed by t-test. A significant difference can be observed between siLUC and siSPRTN. (B) Depletion check performed in HCT116 cells treated with the indicated siRNAs. (C) HT-55 cells were treated with Oxaliplatin for 24 hours. For each condition, the results shows the mean of three independent biological replicate, and they were analysed by t-test. No difference in survival can be observed between siLUC and siSPRTN. (D) Depletion check performed in HT-55 cells treated with the indicated siRNAs. (E) RKO cells were treated with Oxaliplatin for 24 hours. For each condition, the results shows the mean of three independent biological replicate, and they were analysed by t-test. A significant difference can be observed between siLUC and siSPRTN. (F) Depletion check performed in RKO cells treated with the indicated siRNAs.

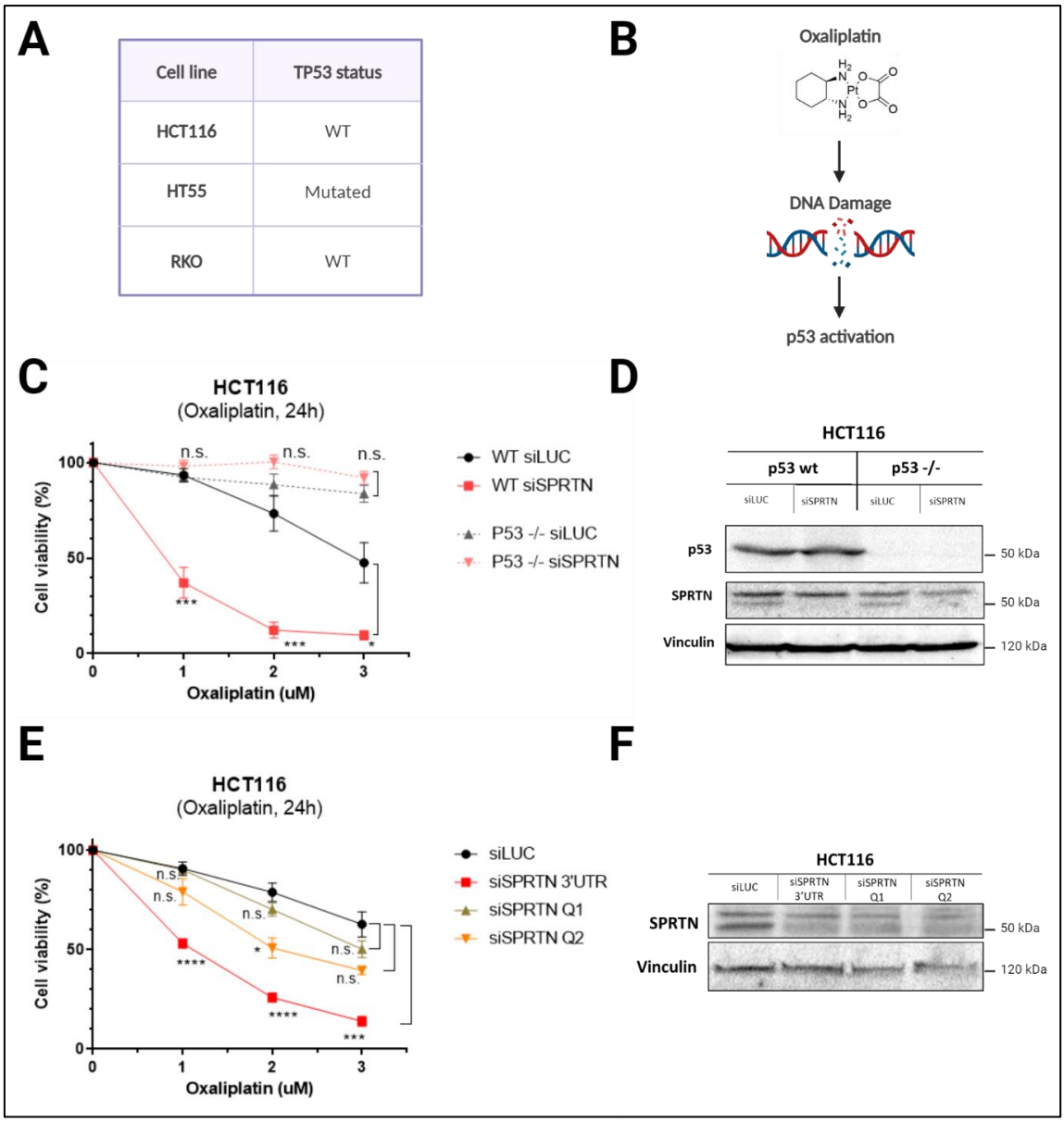

A
Cell line
TP53 status
HCT116
WT
HT55
Mutated
RKO
WT
B
Oxaliplatin
DNA Damage
p53 activation
C
HCT116
(Oxaliplatin, 24h)
Cell viability (%)
Oxaliplatin (uM)
n.s.
WT siLUC
WT siSPRTN
P53 -/- siLUC
P53 -/- siSPRTN
D
HCT116
p53 wt
p53 -/-
siLUC
siSPRTN
p53
SPRTN
Vinculin
50 kDa
120 kDa
E
HCT116
(Oxaliplatin, 24h)
Cell viability (%)
Oxaliplatin (uM)
siLUC
siSPRTN 3'UTR
siSPRTN Q1
siSPRTN Q2
F
HCT116
SPRTN
Vinculin
50 kDa
120 kDa

Figure 34. p53 role in SPRTN depleted cells treated with Oxaliplatin.

(A) HT-55 is the only tested cell line which present an inactivating mutation of TP53. (B) The DNA damage induced by platinum compounds induces the activation of p53 [218]. (C) HCT116 cells $TP53^{+/+}$ and $TP53^{-/-}$ were treated with Oxaliplatin for 24 hours. For each condition, the results shows the mean of three independent biological replicate, and they were analysed by t-test. $TP53^{-/-}$ cells are more resistant to Oxaliplatin and insensitive to the effect of SPRTN depletion. (D) Depletion check performed in HCT116 cells treated with the indicated siRNAs. (E) HCT116 were treated with Oxaliplatin for 24 hours, using different siRNA targeting siSPRTN. For each condition, the results shows the mean of three independent biological replicate, and they were analysed by t-test. While siSPRTN-3'UTR is characterized by a strong phenotype, siSPRTN-Q2 effect is milder and there is no significant difference in survival between siSPRTN-Q1 and siLUC. (F) Depletion check performed in HCT116 cells treated with the indicated siRNAs

Summary

In this initial chapter, the role of SPRTN in mediating the cytotoxicity of relevant chemotherapeutic drugs in colorectal cancer cells was examined. This was performed using siRNAs to deplete SPRTN before treating the cells with nucleotide analogues (5-fluorouracil) platinum compounds (oxaliplatin) or topoisomerase inhibitors (irinotecan, camptothecin). The efficacy of SPRTN depletion was assessed by both RT-qPCR and western blot. Surprisingly, colorectal cancer cells seemed particularly resistant to the effects of SPRTN depletion, and an increased cytotoxicity could be observed only in p53 wild-type colorectal cancer cells treated with oxaliplatin.

Despite the interesting phenotype observed in oxaliplatin-treated cells, it was also important to determine if the this was indeed induced by the effect of SPRTN depletion, as siRNAs are generally known to present some relevant limitations, including the possible depletion of off-target genes [219]. This was partially addressed by the use of two more siRNA sequences to deplete SPRTN (siSPRTN-Q1 and siSPRTN-Q2), but the resulting milder phenotype (Figure 34E-F) did not provide a definitive answer, as it could as well be related to a lower SPRTN depletion efficacy. While a few different approaches could have been used to address this issue, including the supplementation of siRNA-resistant SPRTN, this would not have removed the general limitations which are related to the use of siRNAs. A much better experimental setting is generally given by gene knockouts, but SPRTN full knockout has historically proven to be challenging as SPRTN is an essential gene in most cells (Figure 29A) [220]. This issue was overcome through the development of an auxin-inducible degron system (AID) for SPRTN, which allows the rapid and specific depletion of essential genes after the administration of auxin [221]. The generation of a SPRTN-AID system and its characterization will be the main focus of the next chapter of this thesis.

3.2 Development and characterization of a SPRTN-AID system in HCT116 cells

As described in the last paragraph of Chapter 3.1, SPRTN full knockout did not represents a feasible approach to confirm the phenotypes presented in Figure 33 and Figure 34. SPRTN is an essential gene in most cell lines (as shown in Figure 29A) and, accordingly, its full knockout is generally expected to be lethal [220].

To overcome this issue, the creation of a controlled and inducible depletion of endogenous SPRTN protein by proteolysis-targeting chimera (PROTAC) was attempted [222]. More specifically, an auxin-inducible degron (AID) tag was added to the C-terminal part of the SPRTN gene in HCT116 cells, in order to induce the proteasome-mediated degradation of SPRTN protein every time these cells are exposed to auxin, a plant hormone. Plants have long been known to be able to regulate numerous developmental and physiological processes through the E3-ubiquitin ligase SCF(TIR1/AFB)-mediated proteolysis induced by the auxin family of transcriptional regulators (AUX/IAA proteins) [223]. This process was exploited for the first time by Nishimura et al. in 2009, confirming how it could be transplanted in other cells and used to control the rapid degradation of essential proteins through the exogenous administration of auxin [221].

More in details, in plants auxin mediates the interaction of the AUX/IAA family of transcriptional repressor with the Transport Inhibitor Response 1 (TIR1) protein. TIR1 represents an important component of the E3 ubiquitin ligase SCF(TIR1/AFB), and its auxin-mediated interaction with the AUX/IAA proteins leads to their polyubiquitination and subsequent proteasome-dependent degradation. This system can be adapted in other eukaryotic cells by expressing TIR1 (generally TIR1 from *Oryza sativa* [osTIR1]) in the cell line of interest while also fusing a auxin-inducible degron (AID) tag to the protein which needs to be degraded [224]. Once this has been achieved, the exogenous administration of auxin (500 µM) will permit the interaction of the AID-tagged protein of interest with osTIR1, leading to its ubiquitination and proteasome-mediated degradation (as summarized in Figure 35D).

HCT116 osTIR1, a colorectal cancer cell line already presenting a doxycycline-inducible osTIR1 gene, was originally generated by Natsume T et al. in 2016 [225]. Using this cell line, a CRISPR-Cas9 knock-in approach was used to add the desired tags to the C-terminus of SPRTN gene (AID, V5, TEV and Strep II tags were inserted) as showed in Figure 35A [226] The genome-editing DNAs were provided by the WIMM Genome Engineering Facility and have already been reported in Supplementary Table 2 and Figure 24.

The different clones obtained after CRISPR-Cas9 treatment were initially screened by western blot, to identify potential clones presenting a homozygous SPRTN-AID insertion (Figure 35B). The homozygosity of the most promising clone, HCT116 osTIR1 SPRTN-AID clone 07 (from now on indicated as HCT116 SPRTN-AID), was then also confirmed by PCR (Figure 35C) and by sequencing (Table 19).

To determine the efficacy and the time required to obtain a significant degradation of SPRTN-AID, after induction of osTIR1 expression by doxycycline (1 µg/ml for 24 hours) the selected HCT116 SPRTN-AID clone was exposed to auxin 500 µM for different amounts of time (Figure 35E). As showed, the addition of auxin induced an almost complete SPRTN-AID degradation after 30 minutes and this expression was recovered to normal level within 48 hours after auxin washout.

.

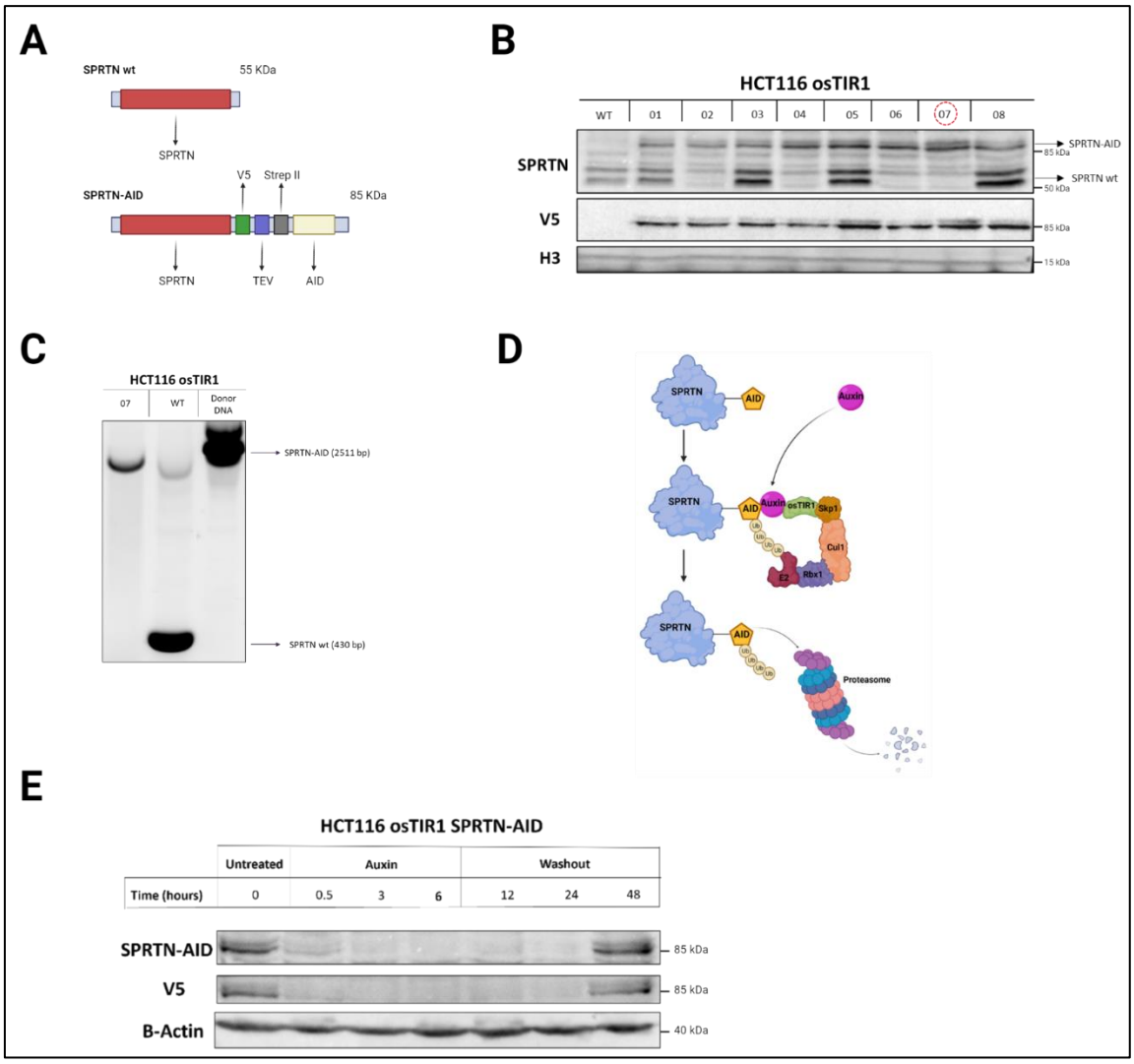

A
SPRTN wt
55 KDa
SPRTN
SPRTN-AID
V5
Strep II
85 KDa
SPRTN
TEV
AID
B
HCT116 osTIR1
WT
01
02
03
04
05
06
07
08
SPRTN
SPRTN-AID
85 kDa
SPRTN wt
50 kDa
V5
85 kDa
H3
15 kDa
C
HCT116 osTIR1
07
WT
Donor DNA
SPRTN-AID (2511 bp)
SPRTN wt (430 bp)
D
SPRTN
AID
Auxin
SPRTN
AID
Auxin
osTIR1
Skp1
Cul1
E2
Rbx1
SPRTN
AID
Proteasome
E
HCT116 osTIR1 SPRTN-AID
Untreated
Auxin
Washout
Time (hours)
0
0.5
3
6
12
24
48
SPRTN-AID
85 kDa
V5
85 kDa
B-Actin
40 kDa

Figure 35. Generation of HCT116 SPRTN-AID cells.

(A) SPRTN-AID gene structure, with list of added tags added to the C-terminus of SPRTN. V5 tag was added for better identification by ELISA, flow cytometry, immunoprecipitation, immunofluorescence, and western blot. TEV tag was introduced for the potential removal of other tags from the recombinant protein. Strep II tag was added to facilitate protein purification. AID tag was inserted to permit the degradation of SPRTN by the addition of auxin. (B) Different clones obtained after FACS selection and clonal growth were tested by western blot to identify possible homozygous SPRTN-AID clones. (C) SPRTN-AID homozygosity for clone 07 was confirmed by PCR. (D) Schematic summarizing how proteins can be degraded by the terminal addition of an auxin inducible degron (AID) tag in osTIR1 positive cell lines. Auxin mediates the interaction of the SCF^{osTIR1} (Skp1-Cullin-F-box) E3 ligase with the endogenous tagged protein, causing its subsequent polyubiquitination and degradation through the proteasome. (E) SPRTN-AID is almost completely degraded 30 minutes after the addition of auxin and cells requires around 48 hours to restore its expression to physiological levels after auxin has been removed. Of notice, cells were also exposed to 1 µg/ml doxycycline from 24 hours before and for the entire duration of the experiment, to assure a sufficient expression of osTIR1.

Sequencing of SPRTN-AID clone As described on Table 8, the PCR reaction was set up to produce a shorter fragment (502 bp) which could then be sequenced to confirm the correct C-terminal insertion.
TCTTGGTAAGCCTATCCCTAACCCTCTCCTCGGTCTCGATTCTACGGAGAATTTGTATTT TCAGGGTAGCGCCTGGAGCCACCCCCAGTTCGAGAAGGGCGGCGGCAGCGGCGGCG GCAGCGGCGGCAGCGCCTGGAGCCACCCCCAGTTCGAGAAGGGAAGCGTGGAACTGA ACCTGCGCGAGACAGAGCTGTGCCTGGGCCTGCCAGGCGGCGATACAGTGGCCCCCG TGACCGGCAACAGAAGAGGCTTCAGCGAGACAGTGGACCTGAAGCTGAACCTGAACAA CGAGCCCGCCAACAAAGAGGGCAGCACCACCCACGACGTGGTGACATTCGACAGCAAA GAGAAGTCCGCCTGCCCCAAGGACCCCGCCAAGCCACCAGCCAAGGCCCAGGTGGTC GGATGGCCTCCCGTGCGCAGCTACAGAAAAGAACGTAAA

Table 19. SPRTN-AID sequencing.

Sequencing of C-terminal insertion for SPRTN-AID clone 07. The sequence corresponding to SPRTN Exon 5 (C-terminus) is indicated in green. The V5 tag is indicated in red. The TEV tag is indicated in yellow. The Strep II tag is indicated in grey. The AID tag is indicated in light blue.

Once the efficacy of SPRTN depletion in the HCT116 SPRTN-AID cell line was confirmed, the next step involved confirming the phenotypes obtained in the previous chapter and summarized in Figure 32-34. HCT116 SPRTN-AID cells ± 500 μM auxin were treated for 24 hours with different concentrations of oxaliplatin (Figure 36A), cisplatin (Figure 36B) and camptothecin (Figure 36C). Cisplatin was introduced to determine if the increased efficacy of oxaliplatin after SPRTN depletion by siRNA, observed in the previous chapter, could also be confirmed with other platinum compounds. Surprisingly, SPRTN depletion by auxin in HCT116 SPRTN-AID cells did not significantly improved the efficacy of any of the three different chemotherapeutic drugs (Figure 36A-D). This strongly suggests that the previous phenotype observed in oxaliplatin-treated colorectal cancer cells, which was proved to be p53-dependent, was most likely due to an unfortunate off-target effect of siSPRTN 3'UTR.

The results collected until this point are particularly intriguing, as they are suggesting that SPRTN depletion does not increase the efficacy of cisplatin, oxaliplatin, irinotecan and camptothecin in colorectal cancer cells. Despite SPRTN depletion has been proved to increase the efficacy camptothecin [158, 164] and cisplatin [135, 136] both in several cell lines and in more complex systems like *C. elegans*, these results were never replicated using colorectal cancer cells. In order to address if the observed resistance to SPRTN depletion was limited to colorectal cancer cell, or if it could instead be due to technical issues related to the colony forming assays, the experiments were repeated using HeLa cells as a positive control.

HeLa wild-type cells (HeLa WT) were compared to HeLa SPRTN haploinsufficient cells (HeLa SPRTN $^{-/-/+}$), previously produced in our laboratory by John Fielden [158]. As expected, the reduced levels of SPRTN increased the efficacy of oxaliplatin, cisplatin and camptothecin in this cell line (Figure 36E-H), excluding a technical issue with the assay itself.

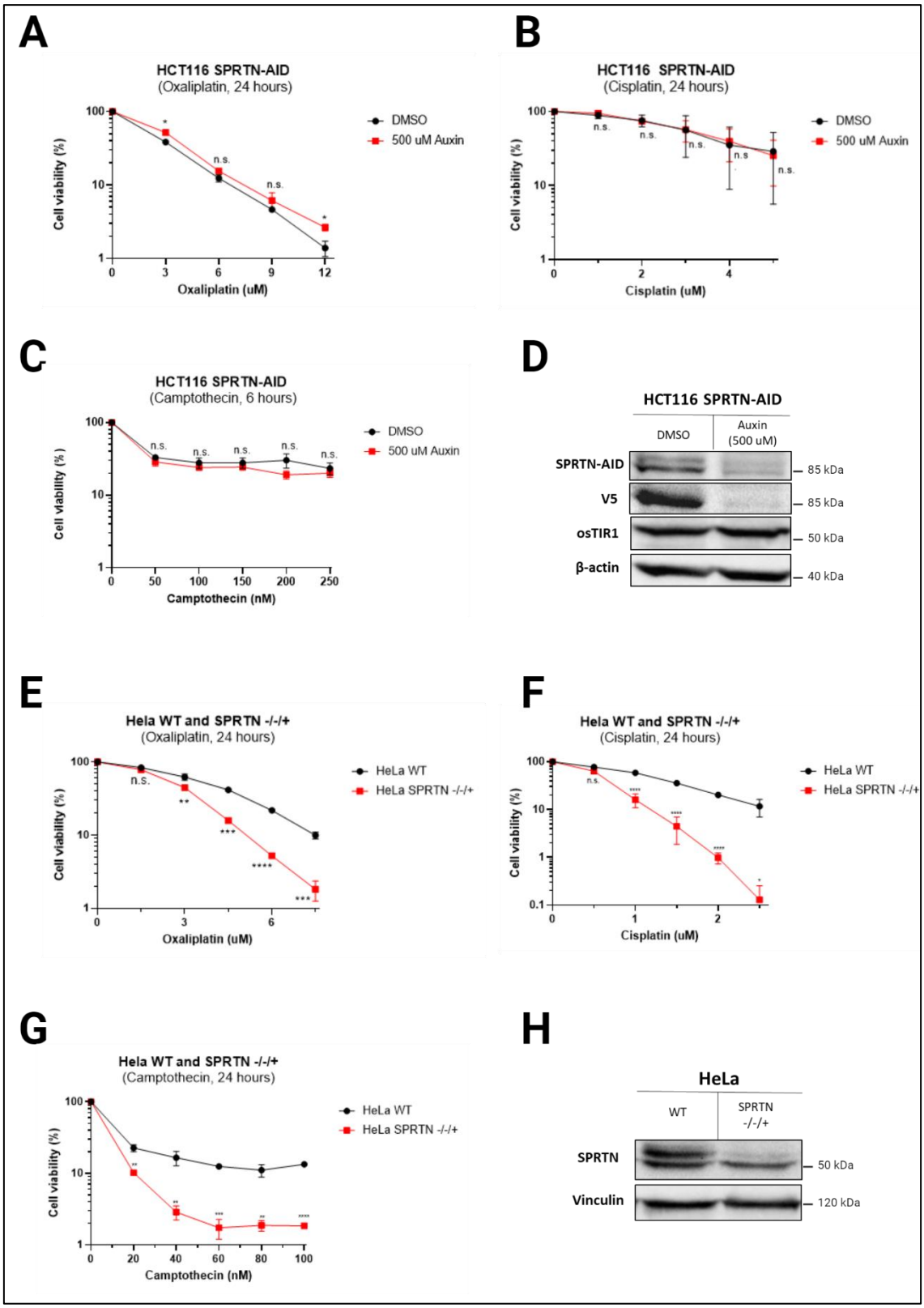

A
HCT116 SPRTN-AID
(Oxaliplatin, 24 hours)
Cell viability (%)
Oxaliplatin (uM)
DMSO
500 uM Auxin
n.s.
B
HCT116 SPRTN-AID
(Cisplatin, 24 hours)
Cell viability (%)
Cisplatin (uM)
DMSO
500 uM Auxin
C
HCT116 SPRTN-AID
(Camptothecin, 6 hours)
Cell viability (%)
Camptothecin (nM)
DMSO
500 uM Auxin
D
HCT116 SPRTN-AID
DMSO
Auxin (500 uM)
SPRTN-AID
V5
osTIR1
β-actin
85 kDa
85 kDa
50 kDa
40 kDa
E
Hela WT and SPRTN -/-/+
(Oxaliplatin, 24 hours)
Cell viability (%)
Oxaliplatin (uM)
HeLa WT
HeLa SPRTN -/-/+
F
Hela WT and SPRTN -/-/+
(Cisplatin, 24 hours)
Cell viability (%)
Cisplatin (uM)
HeLa WT
HeLa SPRTN -/-/+
G
Hela WT and SPRTN -/-/+
(Camptothecin, 24 hours)
Cell viability (%)
Camptothecin (nM)
HeLa WT
HeLa SPRTN -/-/+
H
HeLa
WT
SPRTN -/-/+
SPRTN
Vinculin
50 kDa
120 kDa

Figure 36. SPRTN-depleted cells survival to DPC-inducing drugs.

(A) HCT116 SPRTN-AID cells were treated with oxaliplatin for 24 hours, with and without the addition of auxin. For each condition, the results shows the mean of three independent biological replicate, and they were analysed by t-test. No difference in survival can be observed. (B) HCT116 SPRTN-AID cells were treated with cisplatin for 24 hours, with and without the addition of auxin. For each condition, the results shows the mean of three independent biological replicate, and they were analysed by t-test. No difference in survival can be observed. (C) HCT116 SPRTN-AID cells were treated with camptothecin for 24 hours, with and without the addition of auxin. For each condition, the results shows the mean of three independent biological replicate, and they were analysed by t-test. No difference in survival can be observed. (D) Depletion check performed in HCT116 SPRTN-AID cells (E) HeLa WT and HeLa SPRTN $^{-/-/+}$ cells were treated with oxaliplatin for 24 hours. For each condition, the results shows the mean of three independent biological replicate, and they were analyzed by t-test. HeLa SPRTN $^{-/-/+}$ cells present an increased sensitivity to the indicated drug. (F) HeLa WT and HeLa SPRTN $^{-/-/+}$ cells were treated with cisplatin for 24 hours. For each condition, the results shows the mean of three independent biological replicate, and they were analyzed by t-test. HeLa SPRTN $^{-/-/+}$ cells present an increased sensitivity to the indicated drug. (G) HeLa WT and HeLa SPRTN $^{-/-/+}$ cells were treated with camptothecin for 24 hours. For each condition, the results shows the mean of three independent biological replicate, and they were analysed by t-test. HeLa SPRTN $^{-/-/+}$ cells present an increased sensitivity to the indicated drug. (H) Western blot confirming different level of SPRTN expression in HeLa WT and HeLa SPRTN $^{-/-/+}$ cells.

The next step aimed to determine if SPRTN depletion influences the kinetics of DNA damage repair observed in HCT116 cells treated with chemotherapy, despite their survival not being affected. To this aim, the accumulation of 53BP1 and the formation of phosphorylated H2AX (γH2AX) foci after treatment with cisplatin, oxaliplatin and camptothecin was measured. Both 53BP1 and γH2AX are generally recognized as markers for double strand breaks (DSBs), the most cytotoxic type of DNA damage [227-232].

Initially, HCT116 SPRTN-AID cells ± 500 μM auxin were exposed to cisplatin, oxaliplatin or camptothecin for 2 hours. Then, after washout of the drugs, cells were allowed to recover for 24 and 48 hours before being fixed. 53BP1 and γH2AX foci were analysed by immunofluorescence and quantified as described in “Chapter 2: Material and Methods”. These experiments, summarized in Figure 37A-I, were originally performed by Susan Kilgas, a previous PhD student in the Ramadan Laboratory. SPRTN depletion induced by the administration of auxin significantly delayed the kinetics of DSBs repair when compared to the control cells in response to cisplatin, oxaliplatin and camptothecin treatment, as observed in Figure 37A,D,G and quantified in Figure 37B-C,E-F,H-I. Furthermore, even 48 hours after the exposure to these chemotherapeutic drugs was interrupted, SPRTN-depleted cells exhibited a substantial accumulation of DSBs lesions. This represents a particularly interesting finding, as DBSs should normally present a direct correlation with cytotoxicity [233]. However, as previously showed in Figure 36A-D, SPRTN depletion did not increase the cytotoxicity of these drugs in HCT116 SPRTN-AID cells.

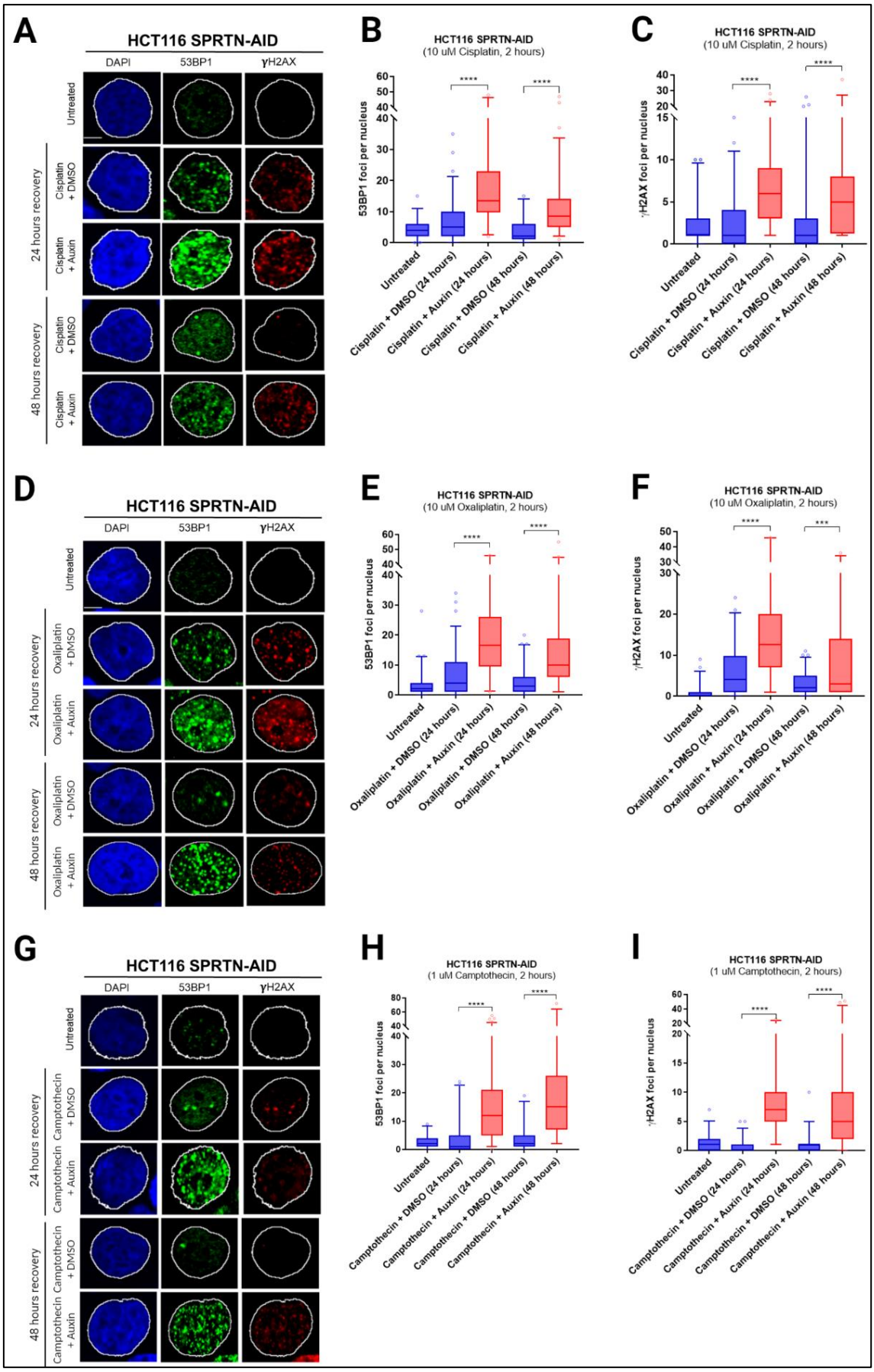

A
HCT116 SPRTN-AID
DAPI
53BP1
γH2AX
Untreated
24 hours recovery
Cisplatin + DMSO
Cisplatin + Auxin
48 hours recovery
Cisplatin + DMSO
Cisplatin + Auxin
B
HCT116 SPRTN-AID
(10 uM Cisplatin, 2 hours)
53BP1 foci per nucleus
****
Untreated
Cisplatin + DMSO (24 hours)
Cisplatin + Auxin (24 hours)
Cisplatin + DMSO (48 hours)
Cisplatin + Auxin (48 hours)
C
HCT116 SPRTN-AID
(10 uM Cisplatin, 2 hours)
γH2AX foci per nucleus
****
Untreated
Cisplatin + DMSO (24 hours)
Cisplatin + Auxin (24 hours)
Cisplatin + DMSO (48 hours)
Cisplatin + Auxin (48 hours)
D
HCT116 SPRTN-AID
DAPI
53BP1
γH2AX
Untreated
24 hours recovery
Oxaliplatin + DMSO
Oxaliplatin + Auxin
48 hours recovery
Oxaliplatin + DMSO
Oxaliplatin + Auxin
E
HCT116 SPRTN-AID
(10 uM Oxaliplatin, 2 hours)
53BP1 foci per nucleus
****
Untreated
Oxaliplatin + DMSO (24 hours)
Oxaliplatin + Auxin (24 hours)
Oxaliplatin + DMSO (48 hours)
Oxaliplatin + Auxin (48 hours)
F
HCT116 SPRTN-AID
(10 uM Oxaliplatin, 2 hours)
γH2AX foci per nucleus
****
***
Untreated
Oxaliplatin + DMSO (24 hours)
Oxaliplatin + Auxin (24 hours)
Oxaliplatin + DMSO (48 hours)
Oxaliplatin + Auxin (48 hours)
G
HCT116 SPRTN-AID
DAPI
53BP1
γH2AX
Untreated
24 hours recovery
Camptothecin + DMSO
Camptothecin + Auxin
48 hours recovery
Camptothecin + DMSO
Camptothecin + Auxin
H
HCT116 SPRTN-AID
(1 uM Camptothecin, 2 hours)
53BP1 foci per nucleus
****
Untreated
Camptothecin + DMSO (24 hours)
Camptothecin + Auxin (24 hours)
Camptothecin + DMSO (48 hours)
Camptothecin + Auxin (48 hours)
I
HCT116 SPRTN-AID
(1 uM Camptothecin, 2 hours)
γH2AX foci per nucleus
****
Untreated
Camptothecin + DMSO (24 hours)
Camptothecin + Auxin (24 hours)
Camptothecin + DMSO (48 hours)
Camptothecin + Auxin (48 hours)

Figure 37. 53BP1 and yH2AX foci in SPRTN-depleted cells.

(A) HCT116 SPRTN-AID cells were treated with 10 μM cisplatin for 2 hours, with and without the addition of auxin. The formation of 53BP1 and γH2AX foci was analysed by immunofluorescence after 24 and 48 hours of recovery. (B) Quantification of 53BP1 foci formation in HCT116 SPRNT-AID cells, based on the images presented on Figure 37A. Their statistical significance was determined by one-way anova. (C) Quantification of yH2AX foci formation in HCT116 SPRNT-AID cells, based on the images presented on Figure 37A. Their statistical significance was determined by one-way anova. (D) HCT116 SPRTN-AID cells were treated with 10 μM oxaliplatin for 2 hours, with and without the addition of auxin. The formation of 53BP1 and γH2AX foci was analysed by immunofluorescence after 24 and 48 hours of recovery. (E) Quantification of 53BP1 foci formation in HCT116 SPRNT-AID cells, based on the images presented on Figure 37D. Their statistical significance was determined by one-way anova. (F) Quantification of yH2AX foci formation in HCT116 SPRNT-AID cells, based on the images presented on Figure 37D. Their statistical significance was determined by one-way anova. (G) HCT116 SPRTN-AID cells were treated with 1 μM camptothecin for 2 hours, with and without the addition of auxin. The formation of 53BP1 and γH2AX foci was analysed by immunofluorescence after 24 and 48 hours of recovery. (H) Quantification of 53BP1 foci formation in HCT116 SPRNT-AID cells, based on the images presented on Figure 37G. Their statistical significance was determined by one-way anova. (I) Quantification of γH2AX foci formation in HCT116 SPRNT-AID cells, based on the images presented on Figure 37G. Their statistical significance was determined by one-way anova.

As previously discussed, SPRTN is generally considered to be an essential gene in most cell lines (Figure 29A). In line with this, SPRTN depletion significantly affects cell cycle progression and SPRTN-deficient cells present DPC-induced replication stress and G2/M checkpoint activation [234-237].

To determine if SPRTN depletion significantly affected the progression through the different phases of the cell cycle in colorectal cancer cells, HCT116 SPRTN-AID cells were equally split and kept in fresh medium ± 500 µM auxin for 48 hours. Then, cells were fixed and processed as described in "Chapter 2: Materials and Methods – Cell cycle analysis". The efficacy of SPRTN depletion by auxin was confirmed by a western blot (Figure 38D), but this did not significantly affect the cell cycle progression (Figure 38A-C). Interestingly, this phenotype seemed once again to be limited to colorectal cancer cells. Once the experiment was repeated with HeLa WT and HeLa SPRTN$^{-/-/+}$, SPRTN depletion determined instead a severe phenotype, with an increase of the percentage of cells within the S and G2/M phases of the cell cycle (Figure 38E-H). While this could theoretically suggest an increased replication rate, data previously published by our laboratory confirms that SPRTN-depleted cells are actually characterized by a slower growth [172]. DPCs are mostly repaired during the S phase, when SPRTN is recruited to the replication fork by directly interacting with PCNA, until being subsequently degraded through an APC/C^{Cdh1}-dependent mechanism during the G1 phase of the cell cycle [143, 238, 239]. Accordingly, SPRTN depletion is accompanied by the accumulation of DPCs which severely limits cell proliferation due to the stalling of replication forks, as confirmed by DNA combing [220].

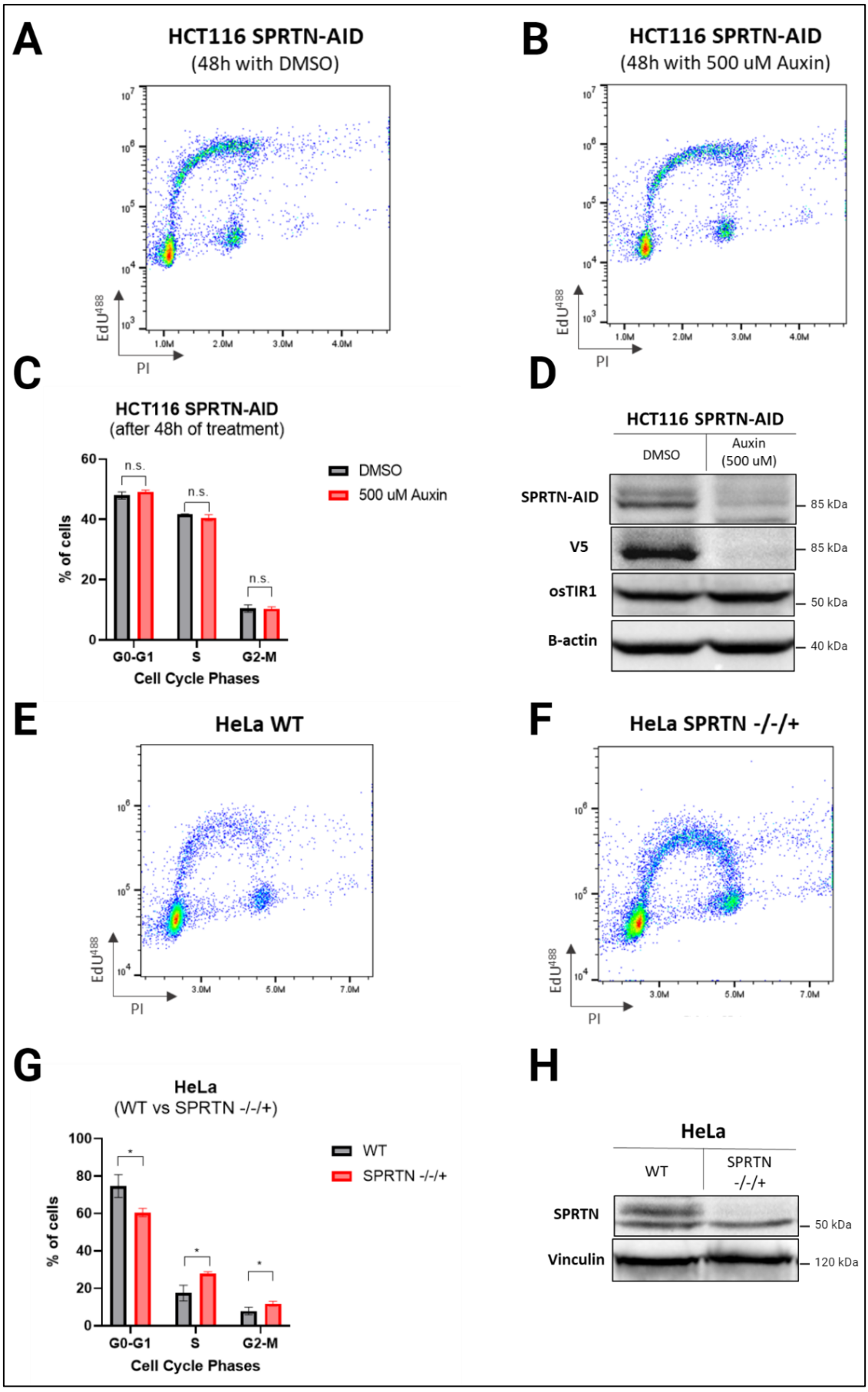

A
HCT116 SPRTN-AID
(48h with DMSO)
EdU488
PI
B
HCT116 SPRTN-AID
(48h with 500 uM Auxin)
EdU488
PI
C
HCT116 SPRTN-AID
(after 48h of treatment)
% of cells
n.s.
n.s.
n.s.
G0-G1
S
G2-M
Cell Cycle Phases
DMSO
500 uM Auxin
D
HCT116 SPRTN-AID
DMSO
Auxin
(500 uM)
SPRTN-AID
V5
osTIR1
B-actin
85 kDa
85 kDa
50 kDa
40 kDa
E
HeLa WT
EdU488
PI
F
HeLa SPRTN -/-/+
EdU488
PI
G
HeLa
(WT vs SPRTN -/-/+)
% of cells
G0-G1
S
G2-M
Cell Cycle Phases
WT
SPRTN -/-/+
H
HeLa
WT
SPRTN
-/-/+
SPRTN
Vinculin
50 kDa
120 kDa

Figure 38. Cell cycle of SPRTN-depleted cells.

(A) The cell cycle of HCT116 SPRTN-AID cells without the addition of auxin was analysed after 48 hours from seeding. (B) The cell cycle of HCT116 SPRTN-AID cells with the addition of auxin was analysed after 48 hours from seeding. (C) Quantification of the percentage of HCT116 SPRTN-AID cells in different phases of the cell cycles, with and without the addition of auxin, based on the images presented in Figure 4.4A and Figure 4.4B. The percentage of cells in each phase of the cell cycle was compared by T-test to determine their statistical significance. No significant difference can be observed. (D) Depletion check of SPRTN-AID in HCT116 cells after the addition of auxin. (E) The cell cycle of HeLa WT cells was analysed after 48 hours from seeding. (F) The cell cycle of HeLa SPRTN $^{-/-/+}$ cells was analysed after 48 hours from seeding. (G) Quantification of the percentage of HeLa WT and Hela SPRTN $^{-/-/+}$ cells in different phases of the cell cycles, based on the images presented in Figure 4.4E and Figure 4.4F. The percentage of cells in each phase of the cell cycle was compared by T-test to determine their statistical significance. In HeLa cells, after the depletion of SPRTN, a higher number of cells can be observed to be in the S and G2/M phase of the cell cycle after 48 hours from seeding. (H) Western blot confirming different level of SPRTN expression in HeLa WT and HeLa SPRTN $^{-/-/+}$ cells.

Summary

In this chapter, the generation and characterization of a SPRTN-AID system was described for the first time. The successful creation of this system in the HCT116 cell line was demonstrated by the rapid depletion of SPRTN protein levels within 30 minutes upon the administration of auxin (Figure 35). Then, the HCT116 SPRTN-AID system was used to confirm the role of SPRTN in mediating the cytotoxicity of different chemotherapeutic compounds (oxaliplatin, cisplatin, camptothecin) (Figure 36). HeLa WT and HeLa $SPRTN^{-/-/+}$ cells, previously reported to be significantly affected by SPRTN depletion, were used a positive control [158].

First, the previously reported increased efficacy of oxaliplatin in TP53 wild-type SPRTN-depleted colorectal cancer cells was confirmed to be likely due to an unfortunate siRNA-mediated off target effect (Figure 36). Despite their remarkable survival profiles, SPRTN-depleted colorectal cells still presented a significant increase in the amount of 53BP1 and γH2AX foci generated while treated with oxaliplatin, cisplatin or camptothecin (Figure 37). 53BP1 and γH2AX foci are generally considered to be useful markers of DSBs, the most cytotoxic DNA lesion. However, the increased amount of DNA damage did not seem to affect HCT116 cells progression through the cell cycle, which was instead severely impacted in HeLa cells (Figure 38). To obtain more insights in the mechanisms responsible for the resistance of colorectal cancer cells to SPRTN protein depletion, a whole genome transcriptomic analysis of HCT116 and HeLa cells was performed. This will be the focus of the next result chapter.

3.3 Whole genome transcriptomic analysis of SPRTN-depleted cells (RNA-seq)

The specific resistance of SPRTN-depleted colorectal cancer cells to various chemotherapeutic agents could be potentially explained by the fact that colorectal cancer might rely on the overexpression of alternative DPC proteases, like GCNA, FAM111A/B and DDI1 [170, 178-180]. This assumption represents however a significant challenge as DPC repair has not yet been fully characterized, and others DPC proteases are likely to be identified in the future. Furthermore, the overexpression of other DPC proteases in colorectal cancer cells could be constitutive or could happen as a response to SPRTN depletion, to assure the survival of cancer cell.

In order address this hypothesis, the whole transcriptomes of HCT116 and HeLa wild type and SPRTN depleted cells, respectively, were investigated and compared using RNA-seq [240, 241]. To induce SPRTN depletion in HCT116 cells, HCT116 SPRTN-AID cells were originally exposed to doxycycline 1 µg/ml for 1 day before beginning of the experiment (to assure adequate levels of osTIR1), and then continuously treated with 500 µM auxin and 1 µg/ml doxycycline for 4 days. As previously shown in Figure 35E, SPRTN protein level was almost completely removed after 30 minutes after auxin addition. In contrary, SPRTN inactivation in HeLa was obtained by CRISPR/Cas9 knock-out to create haploinsufficient SPRTN$^{-/-/+}$ cells as previously published by the Ramadan laboratory [158]. RNA extraction was performed as described in “Chapter 2: Materials and Methods – RNA-seq”. RNA samples were then stored at -80 ºC and shipped to an external company (BGI Genomics) for analysis. The list of differentially expressed genes (DEGs) is reported in Supplementary Tables 3-6, according to the following:

- List of DEGs in HCT116 DMSO vs HCT116 Auxin (Supplementary Table 3)
- List of DEGs in HeLa WT vs HeLa SPRTN $^{-/-/+}$ [Top 100 out of 4575] (Supplementary Table 4)

- List of DEGs present in both HCT116 Auxin and HeLa SPRTN $^{-/-/+}$ (Supplementary Table 5)
- List of DEGs present in HCT116 Auxin, but not in HeLa SPRTN $^{-/-/+}$ (Supplementary Table 6)

Firstly, the amount of SPRTN transcript depletion was assessed and, as expected, HeLa SPRTN $^{-/-/+}$ cells were characterized by significantly lower levels of SPRTN expression when compared with HeLa wild type (Figure 39A). At the same time, no difference in SPRTN gene expression was identified in HCT116 SPRTN-AID cells before and after the addition of auxin (Figure 39A). This is expected, as auxin induces the degradation of SPRTN only after it has been translated from mRNA into a protein [221].

Next, the gene expression of alternative DPC proteases (GCNA, FAM111A/B, DDI1) was also assessed, to determine if these proteases are upregulated after SPRTN depletion (Figure 39B). This analysis showed that, while neither GCNA or DDI1 are significantly expressed in HCT116 and HeLa, FAM111A and FAM111B are instead upregulated in SPRTN-depleted cells. Furthermore, it can be observed that HCT116 cells are more likely to rely on FAM111A as an alternative DPC protease, while HeLa cells preferentially express FAM111B (Figure 39B). Despite FAM111A and FAM111B share over 45% of sequence homology [184], their respective contribution to DPC repair in SPRTN-deficient cells has never been properly assessed. In fact, these results suggest that the resistance to the effect of SPRTN depletion of colorectal cancer cells, largely described in this thesis, has the potential to be due to their reliance on FAM111A as their main DPC protease, even in physiological conditions.

The total number of DEGs which were upregulated or downregulated after SPRTN depletion in HCT116 and HeLa cells was reported in Figure 39C. A Venn diagram was also used to compared the number of DEGs which are shared between HCT116 and

HeLa, and 118 genes were identified (Figure 39D). The full list of these 118 genes has been reported in Supplementary Table 5.

In order to allow the rapid identification of the DEGs with the largest fold change and their significance in HCT116 and HeLa, the results presented in Supplementary Table 3 and Supplementary Table 4 were also shown as volcano plots (Figure 40A-B). Interestingly, as it can be observed in Figure 40, some of the main genes which were upregulated in HCT116 cells after the addition of auxin included aryl hydrocarbon receptor repressor (AHRR) and TCDD Inducible Poly-ADP-Ribose Polymerase (TIPARP). The upregulation of both these genes likely represents an adaptive response to the excessive activation of aryl hydrocarbon receptor (AHR), which has previously been associated to the presence of indole compounds like auxin [242-245]. In fact, AHR is a transcription factor regulating cellular detoxification, and several of his target genes (ALDH1A3, CYP1A1, ADGRF1) were found to be upregulated in HCT116 cells after the addition of auxin (Figure 40A, Supplementary Table 3) [246].

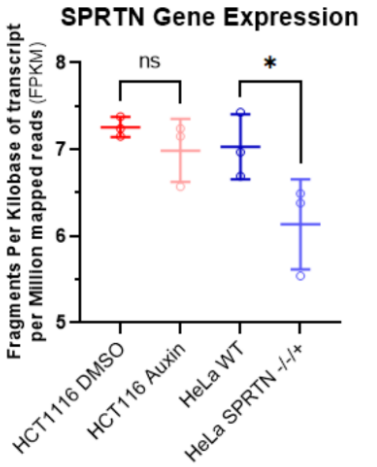
The effect of SPRTN depletion on the cell cycle, previously studied in Figure 38, was once again investigated in both HCT116 and HeLa cells (Supplementary Table 5). Interestingly, both SPRTN-depleted cell lines were characterized by an upregulation of CDK6 (log_2 [HCT116 Auxin / HCT116 DMSO] = 0.462391, q-value = < 0.0000001; log_2 [HeLa SPRTN $^{-/-/+}$ / HeLa WT] = 0.44619, q-value < 0.0000001) and a downregulation of CDKN2A (log2 [HCT116 Auxin / HCT116 DMSO] = -0.40612, q-value < 0.05; log2 [HeLa SPRTN $^{-/-/+}$ / HeLa WT] = -0.39438, q-value < 0.00001). Cyclin D-bound CDK6 has a well characterized role in phosphorylating and deactivating the tumor suppressor protein RB, which in turn dissociates from E2F transcription factors an allows them to stimulate G1-S cell cycle progression and entry into S phase [247-249]. This process is expected to be further facilitated by the CDKN2A downregulation observed in SPRTN-depleted cells. CDKN2A is a gene involved in the production of several proteins, including $p16^{INK4A}$ and $p14^{ARF}$. $p16^{INK4A}$ has been recognized as an important protein responsible for inhibiting

the activity of CDK6 [250], so a downregulation of CDKN2A is likely to further increase CDK6 activity and stimulate G1-S transition. Furthermore, a concomitant reduction in $p14^{ARF}$ expression, considering its role in binding the MDM2-p53 complex and preventing p53 degradation [251, 252], has the potential to reduce p53 activation in SPRTN-depleted cells and inhibit p53-dependend cell cycle arrest and induction of apoptosis [253-255].

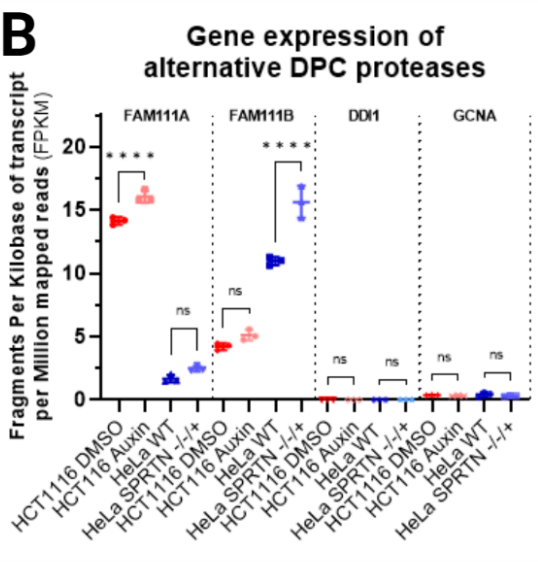
Along with CDKN2A downregulation and CDK6 upregulation, both SPRTN-depleted cell lines were characterized by an upregulation of CDC25B (log2 [HCT116 Auxin / HCT116 DMSO] = 0.173221, q-value < 0.05; log2 [HeLa SPRTN $^{-/-/+}$ / HeLa WT] = 0.231839, q-value < 0.01). This gene encodes for a phosphatase responsible for activating CDK1, which in turn stimulates G2-M cell cycle transition [256-258].

The combined roles of CDK6, CDKN2A and CDC25B are likely to explain the results observed in Figure 38E-H, where HeLa $SPRTN^{-/-/+}$ cells are more likely to be observed in S or G2-M phase when compared to HeLa WT cells. Despite an increased G1-S and G2-M transition, however, SPRTN-depleted cells have previously been reported to present a slower replication rate [165]. Together, these results seem to suggest that the persistence of unrepaired DPCs in SPRTN-depleted cells is likely to prolong the time required for DNA replication and cell mitosis. Despite these common characteristics, however, the cell cycle progression in HCT116 cells remains largely unaffected by SPRTN depletion (Figure 38A-D). This topic will be further expanded in “Chapter 4: Discussion” and, in order to address this, a list of DEGs exclusively present in HCT116 cells has been collected in Supplementary Table 6.

A



B



C

HCT116 and HeLa cells

Differentially expressed genes (DEGs)	HCT116 Auxin vs HCT116 DMSO	HeLa SPRTN -/-/+ vs HeLa WT
Total	253	4574
Upregulated	128	2263
Downregulated	125	2311

D

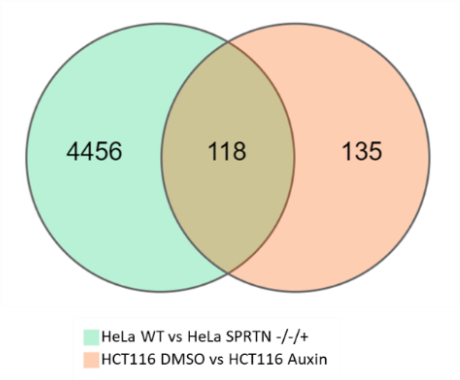


Figure 39. Gene expression in SPRTN-depleted cells.

(A) SPRTN gene expression levels in HCT116 and HeLa cells, with and without SPRTN depletion. The statistical significance between different experimental conditions was determined by T-test. The gene expression is not reduced in HCT116, as SPRTN is degraded only after being translated into protein by the addition of auxin (B) Alternative DPC proteases (FAM111A, FAM111B, DDI1, GCNA) gene expression levels in HCT116 and HeLa cells, with and without SPRTN depletion. The statistical significance between different experimental conditions was determined by T-test. (C) Number of differentially expressed genes identified after SPRTN depletion in HCT116 and HeLa Cells (D) Venn diagram reporting the number of DEGs in HeLa and HCT116 cells after SPRTN depletion. As observed, HeLa cells transcriptome is more severely affected by the depletion of SPRTN and only a minority of DEGs are shared between HeLa and HCT116 cells. The list of DEGs is reported in Supplementary Tables 3-6.

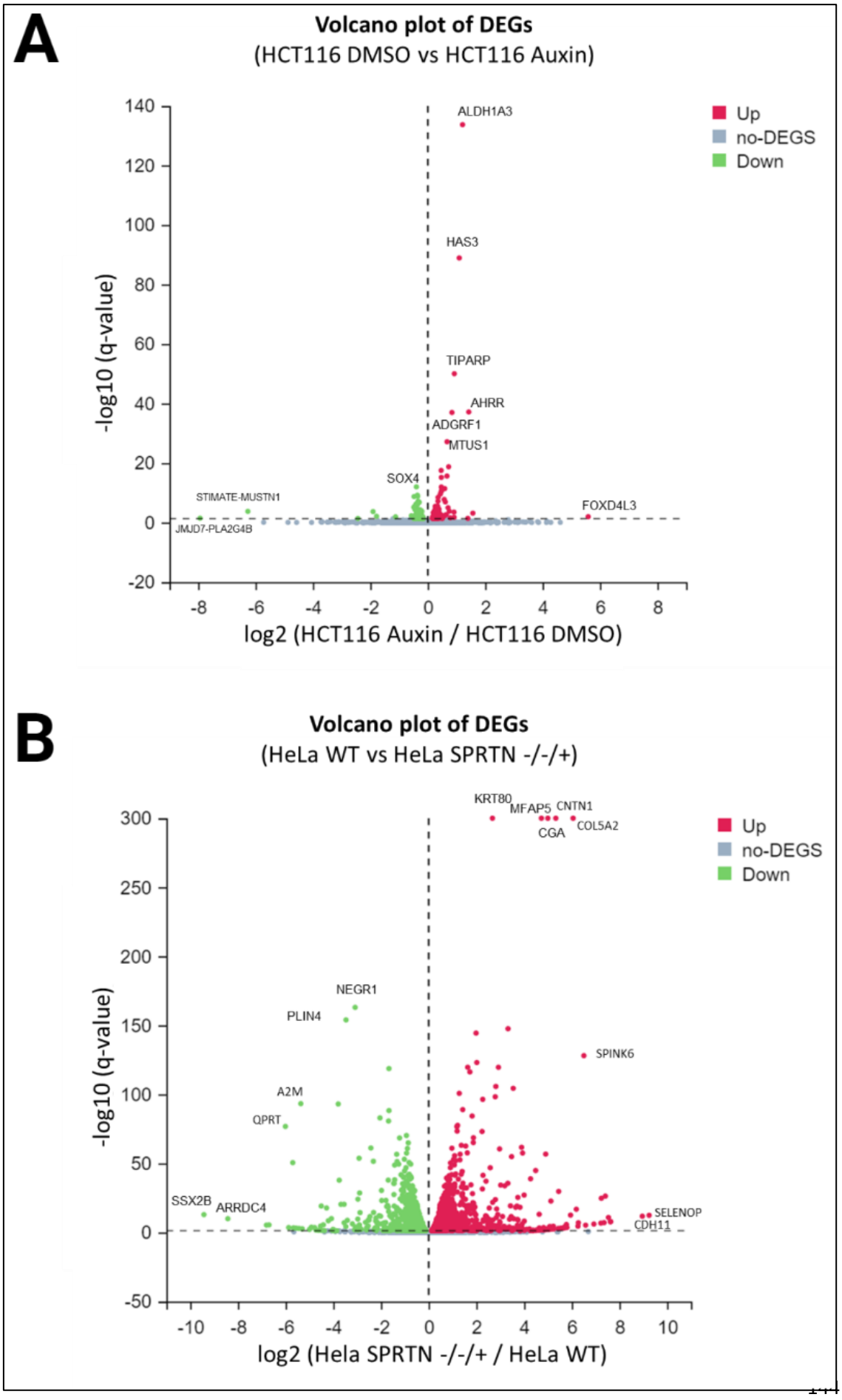

A
Volcano plot of DEGs
(HCT116 DMSO vs HCT116 Auxin)
Up
no-DEGS
Down
ALDH1A3
HAS3
TIPARP
AHRR
ADGRF1
MTUS1
SOX4
STIMATE-MUSTN1
JMJD7-PLA2G4B
FOXD4L3
-log10 (q-value)
log2 (HCT116 Auxin / HCT116 DMSO)
B
Volcano plot of DEGs
(HeLa WT vs HeLa SPRTN -/-/+)
Up
no-DEGS
Down
KRT80
MFAP5
CNTN1
CGA
COL5A2
NEGR1
PLIN4
SPINK6
A2M
QPRT
SSX2B
ARRDC4
SELENOP
CDH11
-log10 (q-value)
log2 (Hela SPRTN -/-/+ / HeLa WT)

Figure 40. Volcano plots of DEGs in SPRTN-depleted cells.

(A) Volcano plot of DEGs in HCT116 SPRTN-AID cells with or without the addition of auxin for 4 days. (B) Volcano plot of DEGs in HeLa WT and HeLa SPRTN -/-/+.

One of the most interesting findings obtained from the analysis of the effect of SPRTN depletion on the transcriptome of both HCT116 and HeLa cells was the fact that these cancers, despite presenting profoundly different phenotypes, were both characterized by a downregulation of multiple ribosomal genes like RPL11, RPL17, RPL24, and RPS16 (Supplementary Table 5). As the ribosomes are the machinery in charge or regulating the synthesis of new proteins in the cells, their genes are generally considered to be some of the most highly expressed genes in most cell types [259, 260]. This finding suggests that DPCs, beside representing a significant obstacle to DNA replication, are also likely to hinder RNA transcription, especially for frequently transcribed genes like the ribosomal ones.

To determine which cellular pathways are strongly affected by SPRTN depletion, a KEGG pathway analysis of the RNA-seq data was also performed (Figure 41). KEGG (Kyoto Encyclopedia of Genes and Genomes) was a project originally initiated in 1995 under the Human Genome Program in Japan and currently contains an extensive database of genes involved in different higher-level cellular pathways [261, 262]. For each cellular pathway, the Rich Ratio (ratio of DEGs in a specific pathway compared to all genes annotated in that pathway) was determined. The higher the Rich Ratio, the more the cellular pathway is enriched. The q-value associated with a specific pathway was also reported in a colour gradient scale. Interestingly, as previously discussed, the ribosome pathway was once again confirmed to represents one of the top hits in both HCT116 and HeLa cells (Figure 41A-B). HCT116 SPRTN-AID DMSO and HeLa WT cells, when compared to their SPRTN-depleted counterparts, both presented a higher expression of multiple ribosomal genes.

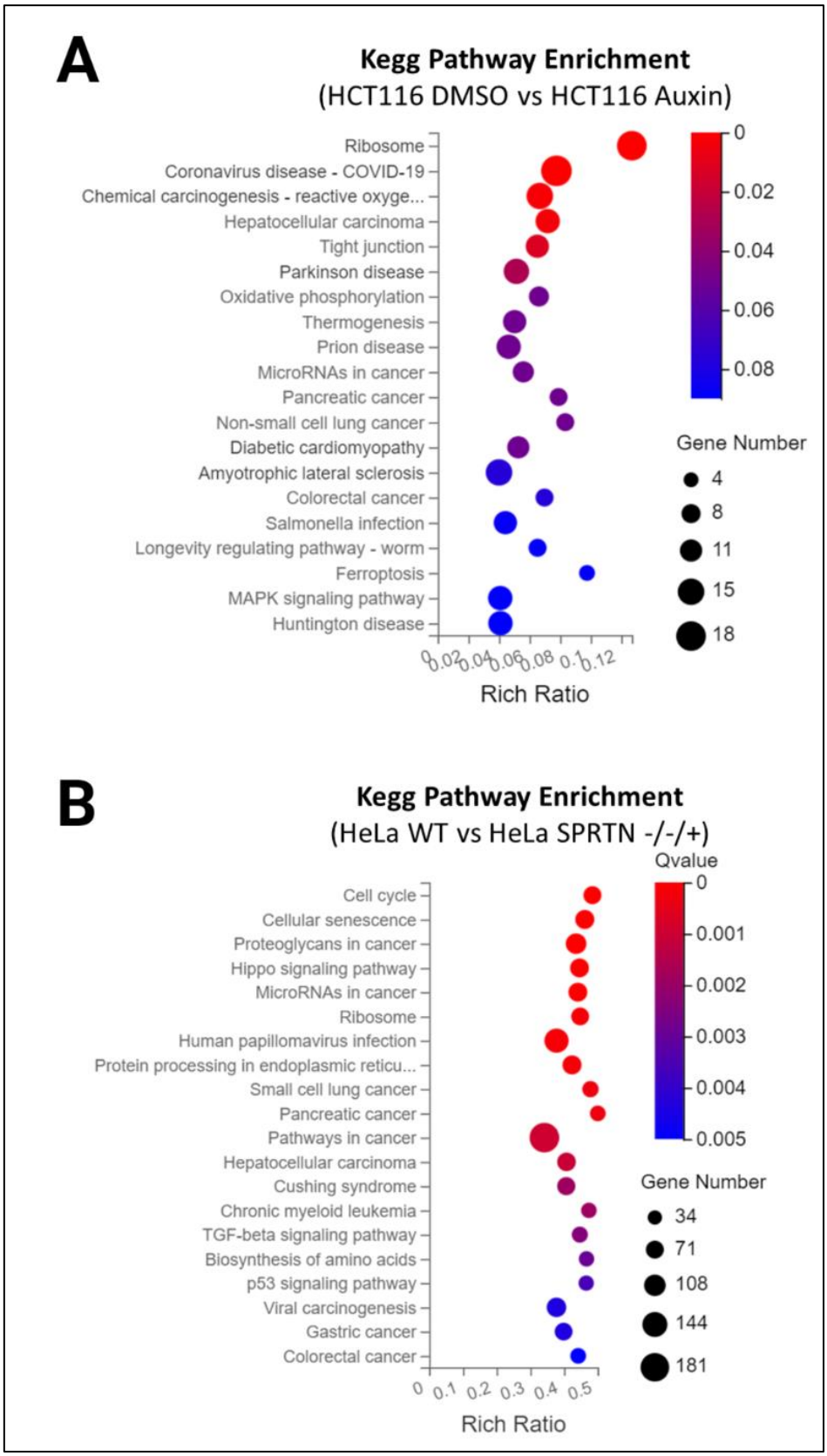

A
Kegg Pathway Enrichment
(HCT116 DMSO vs HCT116 Auxin)
Ribosome
Coronavirus disease - COVID-19
Chemical carcinogenesis - reactive oxyge...
Hepatocellular carcinoma
Tight junction
Parkinson disease
Oxidative phosphorylation
Thermogenesis
Prion disease
MicroRNAs in cancer
Pancreatic cancer
Non-small cell lung cancer
Diabetic cardiomyopathy
Amyotrophic lateral sclerosis
Colorectal cancer
Salmonella infection
Longevity regulating pathway - worm
Ferroptosis
MAPK signaling pathway
Huntington disease
0.02 0.04 0.06 0.08 0.1 0.12
Rich Ratio
0
0.02
0.04
0.06
0.08
Gene Number
4
8
11
15
18
B
Kegg Pathway Enrichment
(HeLa WT vs HeLa SPRTN -/-/+)
Cell cycle
Cellular senescence
Proteoglycans in cancer
Hippo signaling pathway
MicroRNAs in cancer
Ribosome
Human papillomavirus infection
Protein processing in endoplasmic reticu...
Small cell lung cancer
Pancreatic cancer
Pathways in cancer
Hepatocellular carcinoma
Cushing syndrome
Chronic myeloid leukemia
TGF-beta signaling pathway
Biosynthesis of amino acids
p53 signaling pathway
Viral carcinogenesis
Gastric cancer
Colorectal cancer
0 0.1 0.2 0.3 0.4 0.5
Rich Ratio
Qvalue
0
0.001
0.002
0.003
0.004
0.005
Gene Number
34
71
108
144
181

Figure 41. Kegg pathway analysis of SPRTN-depleted cells.

(A) HCT116 SPRTN-AID DMSO cells presents a significant upregulation of the ribosomal pathway when compared to HCT116 SPRTN-AID 500 µM Auxin. (B) HeLa WT cells presents a partially divergent phenotype to the one observed in HCT116 SPRTN-AID DMSO cells. In particular, HeLa WT not only are characterized by an upregulation of ribosomal genes, but also by a significant upregulation of genes involved in the cell cycle when compared to HeLa SPRTN $^{-/-/+}$.

Summary

In this chapter, the transcriptome of SPRTN-depleted cells is analysed for the first time. In order to obtain more insights into the mechanisms responsible for the unexpected SPRTN-resistant phenotypes observed in colorectal cancer cells, both HCT116 and HeLa transcriptomes were analysed by RNA-seq, before and after SPRTN depletion.

Once again, the results suggested that, while SPRTN depletion is deeply affecting the transcriptome of HeLa cells, HCT116 cells remains largely undisturbed (Figure 39A-D). A possible explanation of this phenomenon might be related to the overexpression of another DPC protease, FAM111A, which was observed in HCT116 cells and that could be the main protease responsible for DPC repair in colorectal cancer cells (Figure 39B). While this result is certainly encouraging, further experimental data should be collected to determine if FAM111A is indeed responsible for the resistance to SPRTN-depletion observed in colorectal cancer cells.

One of the few common characteristics observed on SPRTN-depleted HCT116 and HeLa cells has been the downregulation of multiple ribosomal genes (Figure 41, Supplementary Table 5). This finding suggests that DPCs, beside representing a significant obstacle to DNA replication, might also hinder RNA transcription, especially for frequently transcribed genes like the ribosomal ones.

3.4: SPRTN role as a biomarker in CRC patients

The idea that SPRTN expression might have the potential to be used as a biomarker to predict the efficacy of DPC-inducing drugs like TOP1-inhibitors in CRC patients was originally explored by the Ramadan Laboratory at the University of Oxford in 2019. A preliminary analysis was performed on 12 CRC patients, and SPRTN expression assessed by IHC seemed to correlated to the clinical response obtained in patients treated with FOLFIRI. In particular, patients not responding to FOLFIRI resulted more likely to present higher level of SPRTN expression. Based on these early findings, a patent regarding the possible use of SPRTN as a biomarker was filed in 2020 [263].

To further expand and confirm this early and preliminary data, access to adequate clinical samples was required. Formalin-fixed, paraffin-embedded (FFPE) tissue samples previously obtained from patients enrolled in the MRC FOCUS trial were kindly made available by Professor Tim Maughan at the University of Oxford and by the FOCUS Trial Investigators. This trial enrolled 2135 unpretreated CRC patients, which were randomized and allocated to three different possible treatments (5-Fluorouracil alone, FOLFIRI or FOLFOX) (Figure 42A) [204].

Initially, another small exploratory analysis was performed, to investigate if this hypothesis was indeed worth pursuing. This exploratory analysis included only 45 patients and focused exclusively on patients treated with FOLFIRI, as SPRTN has been previously associated to the resolution of TOP1ccs induced by topoisomerase inhibitors like irinotecan (Figure 42B) [140]. Then, based on the encouraging results which were obtained, the analysis was expanded by requesting access to the tissue microarrays of 914 CRC patients treated with 5-Fluorouracil, FOLFIRI or FOLFOX (Figure 42C). Two different SPRTN antibodies (SPRTN Atlas HPA025073, SPRTN Abcam ab121853) were used for the immunohistochemical staining required during these two analyses. The reliability of their staining was compared both in cell lines expressing different level of SPRTN and in patients' tissue, and was considered to be satisfactory (Figure 43A-C).

Then, the obtained staining was quantified using a digital pathology platform (HALO). The entire procedure was performed with the supervision of the Translational Histopathology Laboratory (THL) and Dr. Alistair Easton at the University of Oxford. Further information can be found in the chapter entitled “Material and methods - Immunohistochemistry and staining assessment”.

To provide a better understanding of the result obtained from the different analyses, the results presented in this chapter will be divided in three main sections:

- Exploratory analysis → 45 CRC patients treated with FOLFIRI.
- Oxford cohort → 400 CRC patients treated with 5-Fluorouracil, FOLFIRI or FOLFOX.
- Final analysis (Oxford cohort + Leeds cohort) → 914 CRC patients treated with 5-Fluorouracil, FOLFIRI or FOLFOX.

Of notice, out of the 914 CRC patients tested in the final analysis, the University of Oxford had access to the survival data (progression free survival, overall survival) of only 400 patients. These 400 patients will be indicated as the “Oxford cohort”. The survival data of the remaining 514 patients was instead under the management of professor Jenny Seligmann at the University of Leeds. These 514 patients will be indicated as the “Leeds cohort”. As the access to clinical data between different institutes presents significant bureaucratic challenges, the survival analyses of the entire cohort (Oxford + Leeds) were performed by Faye Elliott, a statistician at the University of Leeds, after the results of SPRTN immunohistochemical staining were shared with them. The statistical analyses of the patient enrolled in the Oxford cohort were instead performed under the supervision of Enric Domingo, a senior bioinformatician at the University of Oxford.

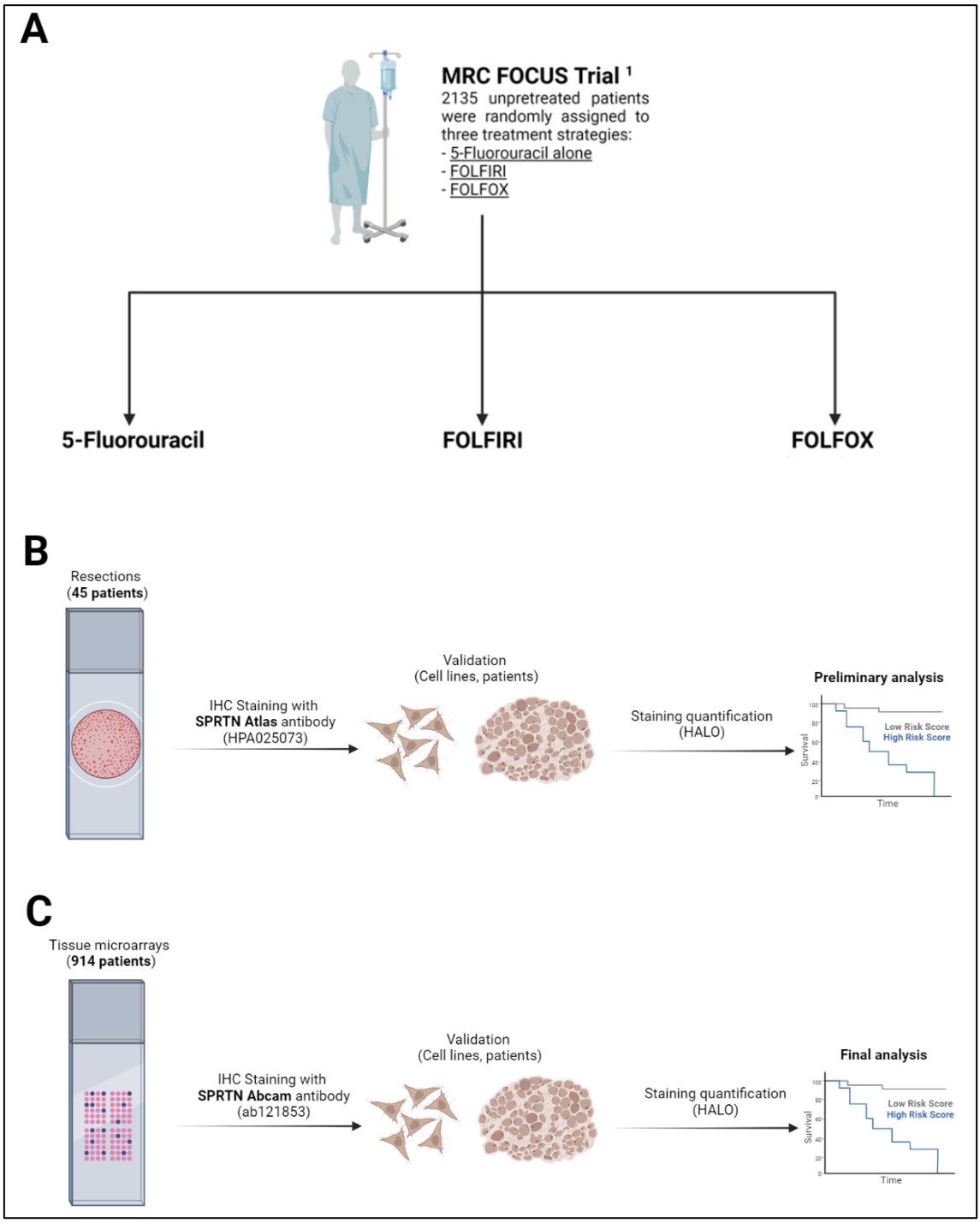

A
MRC FOCUS Trial [1]
2135 unpretreated patients were randomly assigned to three treatment strategies:
- 5-Fluorouracil alone
- FOLFIRI
- FOLFOX
5-Fluorouracil
FOLFIRI
FOLFOX
B
Resections
(45 patients)
IHC Staining with
SPRTN Atlas antibody
(HPA025073)
Validation
(Cell lines, patients)
Staining quantification
(HALO)
Preliminary analysis
Low Risk Score
High Risk Score
Survival
Time
C
Tissue microarrays
(914 patients)
IHC Staining with
SPRTN Abcam antibody
(ab121853)
Validation
(Cell lines, patients)
Staining quantification
(HALO)
Final analysis
Low Risk Score
High Risk Score
Survival
Time

Figure 42. MRC focus trial and SPRTN IHC staining.

(A) Summary of the MRC FOCUS trial. 2135 patients were randomized to three possible first line treatments: 5-fluorouracil, FOLFIRI, FOLFOX. Of these, 914 patients were used for the analyses presented on this chapter, based on the availability of tissue samples for immunohistochemical analysis. [204] (B) Summary of the overall procedures performed in the exploratory analysis. FFPE tissue resections obtained from 45 patients treated with FOLFIRI were stained with the SPRTN Atlas HPA025073 antibody. The staining was quantified using HALO, a digital pathology platform, to assess the percentage of SPRTN-positive cells in the malignant epithelium. Patients were divided in tertiles based on the percentage of SPRTN-positive cells and the survival data was compared between different tertiles. (C) Summary of the overall procedures performed in the Oxford cohort and Leeds cohort. FFPE tissue microarrays obtained from 914 patients treated with 5-Fluorouracil, FOLFIRI or FOLFOX were stained with the SPRTN Abcam ab121853 antibody. The staining was quantified using HALO, a digital pathology platform, to assess the average SPRTN nuclear optical density (OD) in the malignant epithelium. Patients were divided according to the treatment they received (5-Fluorouracil, FOLFIRI, FOLFOX) and according to SPRTN tertiles (SPRTN High, SPRTN Medium and Low).

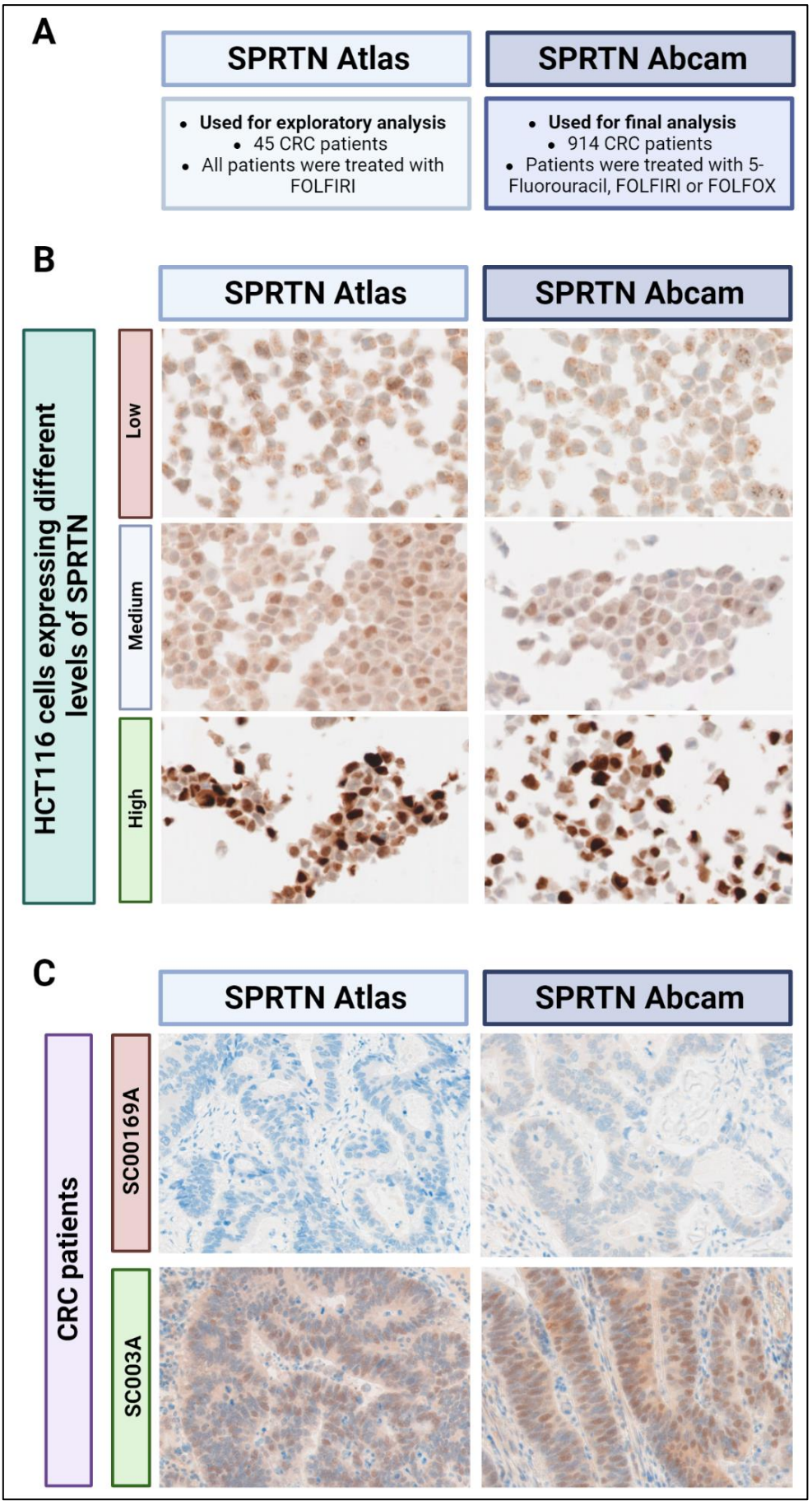

A
SPRTN Atlas
SPRTN Abcam
Used for exploratory analysis
45 CRC patients
All patients were treated with FOLFIRI
Used for final analysis
914 CRC patients
Patients were treated with 5-Fluorouracil, FOLFIRI or FOLFOX
B
SPRTN Atlas
SPRTN Abcam
HCT116 cells expressing different levels of SPRTN
Low
Medium
High
C
SPRTN Atlas
SPRTN Abcam
CRC patients
SC00169A
SC003A

Figure 43. SPRTN antibodies and their validation for IHC.

(A) Summary of the two different antibodies used in the immunohistochemical analyses (SPRTN Atlas HPA025073, SPRTN Abcam ab121853). SPRTN Atlas was used in the exploratory analyses, while SPRTN Abcam was used in the final analyses. Further information regarding the antibody concentrations and the staining procedures can be found in the chapter entitled “Material and methods - Immunohistochemistry and staining assessment”. (B) HCT116 cells expressing different levels of SPRTN (High, Medium, Low) were stained with the two different SPRTN antibodies. HCT116 SPRTN-AID, which was generated in the Ramadan Laboratory and described in the previous result chapter, was used to provide an example of medium SPRTN expression (HCT116 SPRTN-AID) and low SPRTN expression (HCT116 SPRTN-AID + 500 µM). A doxycycline-inducible lentiviral SPRTN vector was used in HCT116 cells to provide an example of cells expressing high level of SPRTN. (C) FFPE tissue samples obtained from two CRC patients (SC00169A, SC003A) were stained with the two different SPRTN antibodies. As observed, the result of the staining is similar with both antibodies.

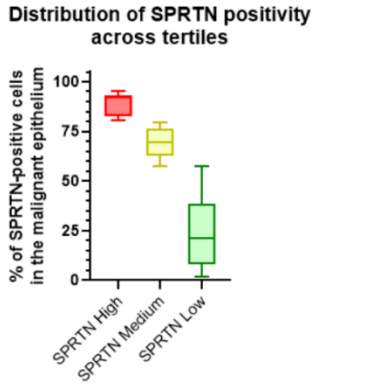
Exploratory analysis (45 CRC patients treated with FOLFIRI)

45 tissue samples obtained from FFPE resections of CRC patients from the MRC FOCUS trial were stained with SPRTN Atlas HPA025073 antibody, as previously described [204]. The staining was quantified using HALO, a digital pathology platform, in order to assess the percentage of SPRTN-positive cells in the malignant epithelium of each samples in an unbiased approach. The data which was obtained from this analysis is summarized in Supplementary Table 7. The 45 patients were then divided in tertiles (15 patients each), according to their SPRTN expression (Figure 44A,B). The progression free survival (PFS) of each group was plotted using Kaplan-Meier curves, and the curves were compared using cox proportional hazard regression (Figure 44C,D). Interestingly, the median PFS was found to be inversely correlated with SPRTN expression in this cohort (5.56 months in SPRTN High, 8.55 months in SPRTN Medium, 9.52 months in SPRTN Low). In particular, patients expressing low levels of SPRTN were characterized by a significantly longer progression free survival when compared to patients expressing high levels of SPRTN (SPRTN Low vs SPRTN High, HR = 0.1547, 95% CI = 0.06233 – 0.3683, p-value < 0.0001). This is in line with the initial hypothesis according to which tumor expressing high SPRTN levels might be able to better repair the TOP1ccs induced by topoisomerase inhibitors like irinotecan, and consequently being more resistant to irinotecan-based chemotherapy regimens like FOLFIRI.

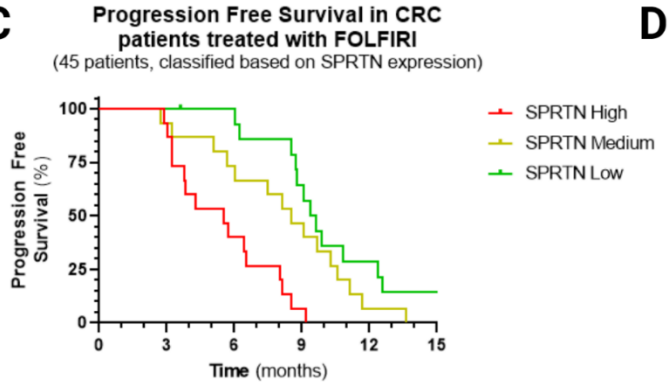
A

	Number of patients	Treatment	Median PFS (months)
SPRTN High	15	FOLFIRI	5.56
SPRTN Medium	15	FOLFIRI	8.55
SPRTN Low	15	FOLFIRI	9.52

B



C



D

Statistical analysis

Comparison	Hazard Ratio (HR)	Confidence interval (95% CI)	p-value
SPRTN Medium vs SPRTN High	0.2711	0.1172 - 0.6122	**0.0018**
SPRTN Low vs SPRTN High	0.1547	0.06233 - 0.3683	**<0.0001**
SPRTN Medium vs SPRTN Low	1.753	0.8151 - 3.853	0.1521

Figure 44. SPRTN as a biomarker in FOLFIRI-treated patients (45 patients).

A) Summary and median progression free survival of the 45 CRC patients in the exploratory cohort, divided in tertiles according to SPRTN expression (B) Percentage of SPRTN-positive cells in the malignant epithelium of the 45 CRC patients in the exploratory cohort, divided in tertiles (C) Progression free survival of the 45 patients enrolled in the exploratory cohort. Patients were divided in tertiles according to SPRTN expression, and their survival was plotted using Kaplan Meyer curves. (D) Statistical analysis of the data presented in Figure 44C.

Oxford Cohort (400 patients)

The detailed survival data of the 400 patients analyzed as the University of Oxford, along with their SPRTN expression assessed by immunohistochemistry and digital pathology, is reported in Supplementary Table 8. As summarized in Figure 45A, 272 patients were treated with 5-Fluorouracil alone, 67 were treated with FOLFIRI and 61 were treated with FOLFOX. As expected, the median Progression Free Survival (PFS) was found to be shorter in the patients treated with 5-Fluorouracil alone (7.27 months) when compared with patients treated with FOLFIRI (8.81 months) or FOLFOX (9.14 months). The survival data was plotted using Kaplan Meyer curves, and the different groups were compared by Cox proportional hazards regression (Figure 45B,C). As expected, FOLFOX was able to provide a significantly better progression free survival when compared to 5-Fluorouracil alone (HR = 0.6917, 95% CI = 0.5156-0.9120, p-value = 0.0111), and no statistically significant difference was observed between FOLFOX and FOLFIRI (HR = 0.8178, 95% CI = 0.5729 – 1.164, p-value = 0.2650). This is in line with what has been previously reported in the scientific literature, where FOLFIRI and FOLFOX are considered to present a similar level of efficacy [25].

After confirming the effect of the three different treatment on PFS, the overall survival (OS) was considered instead (Figure 45D,E). As expected, no difference in overall survival was found between FOLFOX and FOLFIRI (HR = 1.074, 95% CI = 0.7455 – 1.545, p-value = 0.6995). Interestingly, however, even the patients treated with 5-Fluorouracil alone were found to present a comparable overall survival to the ones treated with FOLFIRI and FOLFOX, and no statistically significant difference could be observed (Figure 45D,E). This is most likely related to the fact that, while OS is generally been considered as the best clinical endpoint, it presents potential pitfalls. In particular, it can be influenced by later alternative management strategies [264]. As the patients treated with 5-Fluorouracil alone were treated with either FOLFIRI or FOLFOX after clinical progression [204], the only survival parameter not influenced by these second-

line treatment choices is represented by progression free survival. Accordingly, the PFS was used as the primary endpoint to determine the role of SPRTN as a biomarker to predict the efficacies of different chemotherapy regimens in colorectal cancer. An exploratory analysis was also conducted to evaluate SPRTN possible role as a prognostic biomarker (Supplementary Figure 1), and CRC patients expressing lower levels of SPRTN seemed to be characterized by a slightly longer overall survival within the Oxford Cohort. However, while interpreting this data, it should be taken into account that in this analysis patients were not separated according to the treatment they received.

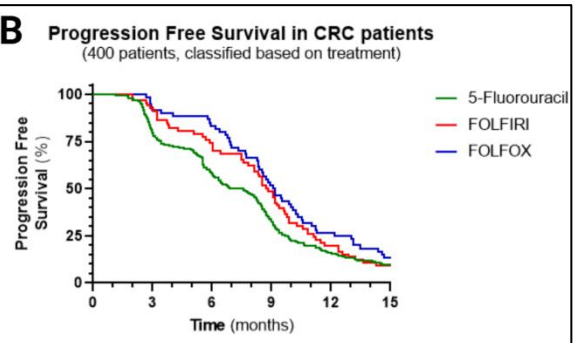
To investigate the possible role of SPRTN as a predictive biomarker, the 400 CRC patients in the Oxford Cohort were divided in tertiles according to the average SPRTN nuclear optical density quantified by HALO in their malignant epithelium (Figure 46A). The progression free survival of the 124 patients expressing the highest levels of SPRTN was compared, according to the treatment that they received (Figure 46B-D). Then, the same was done for the 276 patients expressing medium and low level of SPRTN (Figure 46E-F). Strikingly, the PFS of patients expressing high levels of SPRTN and treated with FOLFIRI was found to be significantly worse than the patients treated with FOLFOX (FOLFOX vs FOLFIRI, HR = 0.4213, 95% CI = 0.1875 – 0.8910, p-value = 0.0278). The PFS of patients expressing high level of SPRTN and treated with FOLFIRI was also found to be not different from the patients treated with 5-Fluorouracil alone (HR = 0.8374, 95% CI = 0.4934 – 1.349, p-value = 0.4865). This is in line with the original hypothesis that tumor expressing high levels of SPRTN might be able to better repair TOP1ccs induced by Irinotecan. Accordingly, the addition of Irinotecan to 5-Fluorouracil did not provide any additional benefit in survival to these patients and they should be treated with FOLFOX instead (Figure 46B-D). This idea was further supported by the fact that patients expressing medium and low levels of SPRTN, instead, did not present any significant difference in PFS while treated with FOLFOX or FOLFIRI (HR = 1.111, 95% CI = 0.7235 – 1.708, p-value = 0.6300).

While these results represented an extremely exciting finding, it is important to underline that the number of patients in this analysis expressing high levels of SPRTN and treated with either FOLFIRI (21 patients) or FOLFOX (16 patients) was modest. Accordingly, the analysis was further expanded including also the 514 CRC patients from the University of Leeds.

A

	Number of patients	Treatment	Median PFS (months)	Median OS (months)
FOLFIRI	67	Irinotecan + 5-Fluorouracil	8.81	14.73
FOLFOX	61	Oxaliplatin + 5-Fluorouracil	9.14	14.83
5-Fluorouracil	272	5-Fluorouracil	7.27	16.08

B

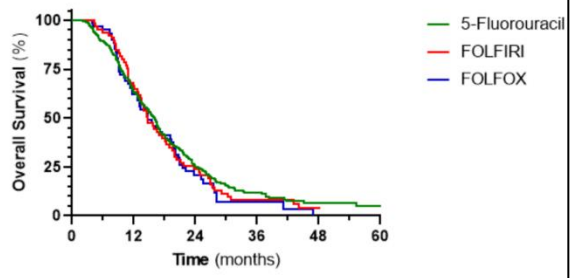


C

Statistical analysis

Comparison	Hazard Ratio (HR)	Confidence interval (95% CI)	p-value
FOLFIRI vs 5-Fluorouracil	0.8458	0.6403 - 1.1010	0.2252
FOLFOX vs 5-Fluorouracil	0.6917	0.5156 - 0.9120	**0.0111**
FOLFOX vs FOLFIRI	0.8178	0.5729 - 1.164	0.2650

D



E

Statistical analysis

Comparison	Hazard Ratio (HR)	Confidence interval (95% CI)	p-value
FOLFIRI vs 5-Fluorouracil	1.081	0.8085 - 1.422	0.5894
FOLFOX vs 5-Fluorouracil	1.161	0.8587 - 1.542	0.3170
FOLFOX vs FOLFIRI	1.074	0.7455 - 1.545	0.6995

Figure 45. Survival analyses of CRC patients based on treatment (Oxford cohort – 400 patients)

(A) Summary of the 400 CRC patients in the Oxford cohort, divided according to the first-line treatment they received (B) Progression free survival of the 400 patients enrolled in the Oxford cohort. Patients were divided according to treatment, and their survival was plotted using Kaplan Meyer curves. (C) Statistical analysis of the data presented in Figure 45B. (D) Overall survival of the 400 patients enrolled in the Oxford cohort. Patients were divided according to treatment, and their survival was plotted using Kaplan Meyer curves. (E) Statistical analysis of the data presented in Figure 45D.

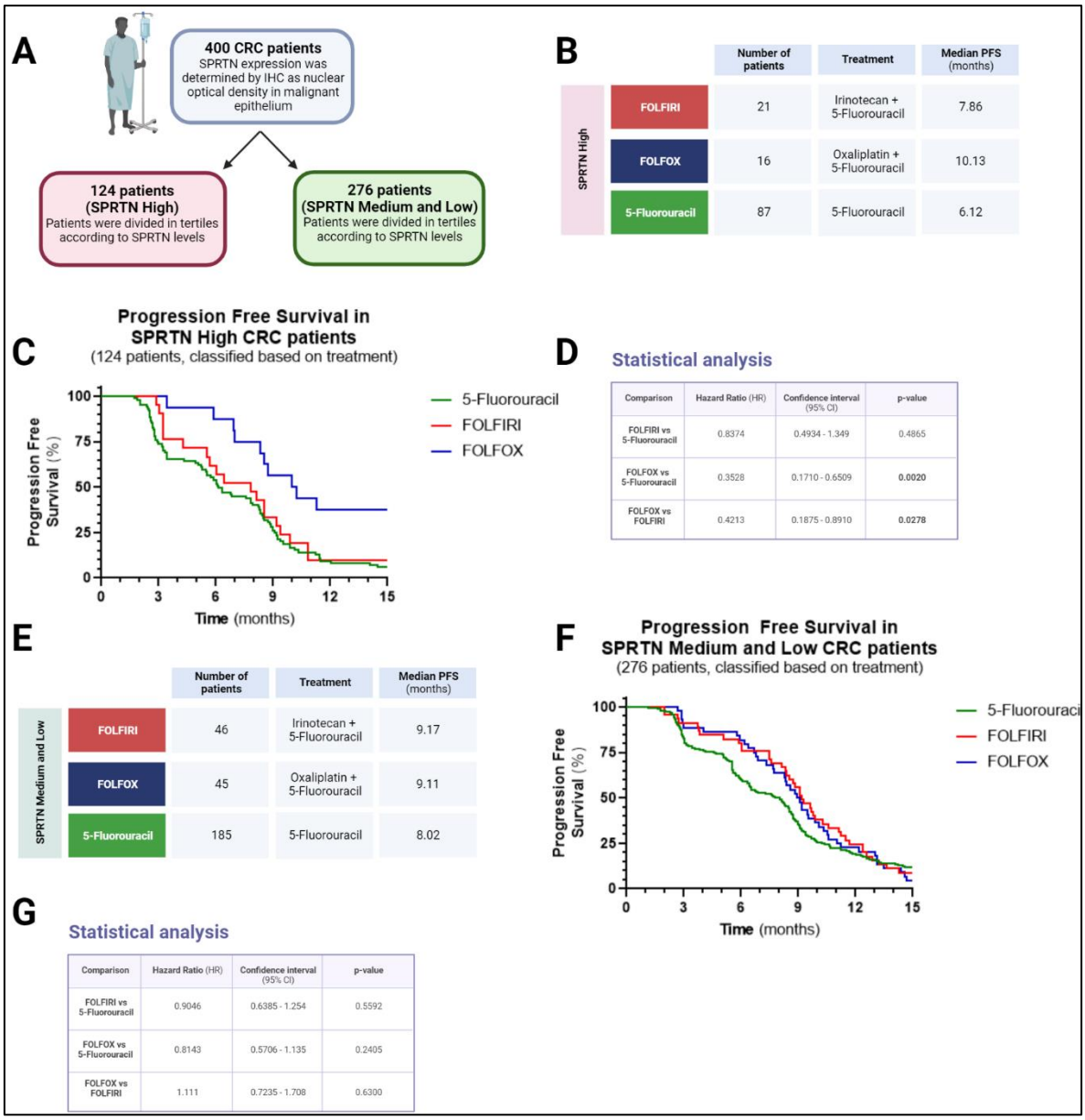

A
400 CRC patients
SPRTN expression was determined by IHC as nuclear optical density in malignant epithelium
124 patients (SPRTN High)
Patients were divided in tertiles according to SPRTN levels
276 patients (SPRTN Medium and Low)
Patients were divided in tertiles according to SPRTN levels
B
SPRTN High
FOLFIRI 21 Irinotecan + 5-Fluorouracil 7.86
FOLFOX 16 Oxaliplatin + 5-Fluorouracil 10.13
5-Fluorouracil 87 5-Fluorouracil 6.12
C
Progression Free Survival in SPRTN High CRC patients
(124 patients, classified based on treatment)
Progression Free Survival (%)
Time (months)
5-Fluorouracil
FOLFIRI
FOLFOX
D
Statistical analysis
Comparison | Hazard Ratio (HR) | Confidence interval (95% CI) | p-value
FOLFIRI vs 5-Fluorouracil | 0.8374 | 0.4934 - 1.349 | 0.4865
FOLFOX vs 5-Fluorouracil | 0.3528 | 0.1710 - 0.6509 | 0.0020
FOLFOX vs FOLFIRI | 0.4213 | 0.1875 - 0.8910 | 0.0278
E
SPRTN Medium and Low
FOLFIRI 46 Irinotecan + 5-Fluorouracil 9.17
FOLFOX 45 Oxaliplatin + 5-Fluorouracil 9.11
5-Fluorouracil 185 5-Fluorouracil 8.02
F
Progression Free Survival in SPRTN Medium and Low CRC patients
(276 patients, classified based on treatment)
Progression Free Survival (%)
Time (months)
5-Fluorouracil
FOLFIRI
FOLFOX
G
Statistical analysis
Comparison | Hazard Ratio (HR) | Confidence interval (95% CI) | p-value
FOLFIRI vs 5-Fluorouracil | 0.9046 | 0.6385 - 1.254 | 0.5592
FOLFOX vs 5-Fluorouracil | 0.8143 | 0.5706 - 1.135 | 0.2405
FOLFOX vs FOLFIRI | 1.111 | 0.7235 - 1.708 | 0.6300

Figure 46. SPRTN as a predictive biomarker in CRC patients (Oxford cohort – 400 patients).

(A) The 400 CRC patients in the Oxford cohort were divided in two groups (SPRTN High, SPRTN Medium and Low), according to the SPRTN nuclear optical density quantified by HALO in their malignant epithelium (B) Summary of the 124 CRC patients expressing high levels of SPRTN in the Oxford cohort, divided according to the first-line treatment they received (C) Progression free survival of the 124 patients expressing high levels of SPRTN. Patients were divided according to treatment, and their survival was plotted using Kaplan Meyer curves. (D) Statistical analysis of the data presented in Figure 46C. (E) Summary of the 276 CRC patients expressing medium or low levels of SPRTN in the Oxford cohort, divided according to the first-line treatment they received (F) Progression free survival of the 276 patients expressing medium or low levels of SPRTN. Patients were divided according to treatment, and their survival was plotted using Kaplan Meyer curves. (G) Statistical analysis of the data presented in Figure 46F.

Final analysis (Oxford cohort + Leeds cohort) (914 patients)

As previously discussed, the final analysis included a total of 914 colorectal cancer patients treated with either 5-Fluorouracil, FOLFIRI or FOLFOX. However, the University of Oxford had access to the clinical data of only 400 of these patients, while the remaining 514 patients' clinical data was under the management of the University of Leeds. As the sharing of clinical data has the potential to increase the risk of confidentiality breaches for trial participants, the direct access to the clinical data of the 514 patients was originally denied. This issue was addressed by deciding to share the results obtained from the assessment of the SPRTN immunohistochemical staining with the research group of professor Jenny Seligmann at the University of Leeds. Then, the final analysis was performed by Faye Elliot, an experienced biostatistician, who completed the analysis including the remaining 514 patients from Leeds.

As described in Figure 47A, the 914 patients were once again divided in two main groups, according to their SPRTN expression tertiles: SPRTN High (304 patients), SPRTN Medium and Low (610 patients).

The progression free survival of the 304 CRC patients expressing high levels of SPRTN, divided according to the treatment they received, was then analysed (Figure 47B-D). Unfortunately, the statistically significant difference in PFS observed in SPRTN High patients enrolled in the Oxford cohort and treated with FOLFIRI or FOLFOX was not confirmed in the final analysis (Figure 46B-D, Figure 47B-D). In order to address which reasons might be responsible for this difference, the main clinical and molecular characteristics of the Oxford cohort (400 patients) and the Leeds cohort (514 patients) were analysed through a multivariate analysis (Table 20). As showed in Table 20, the main statistically significant differences between the two populations were the presence of lung (p-value =0.03) or liver metastases (p-value < 0.0005), the primary tumour location (p-value < 0.0005), and their clinical response after 12 weeks (p-value = 0.002). Strikingly, only 6.7% of CRC cases in the Oxford Cohort actually originated from the

rectum, while over 46% of cases in the remaining population from Leeds were rectal cancers. This difference is due to the fact that the patients in the Oxford cohort were originally selected in order to develop possible stratifiers for FOLFOX or FOLFIRI chemotherapy, and most rectal cases were excluded as they are usually treated with neoadjuvant radiotherapy.

A



914 CRC patients
SPRTN expression was determined by IHC as nuclear optical density in malignant epithelium

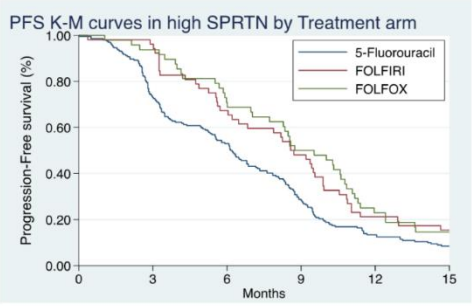
304 patients (SPRTN High)
Patients were divided in tertiles according to SPRTN levels

610 patients (SPRTN Medium and Low)
Patients were divided in tertiles according to SPRTN levels

B

		Number of patients	Treatment	Median PFS (months)
SPRTN High	FOLFIRI	52	Irinotecan + 5-Fluorouracil	8.6
	FOLFOX	48	Oxaliplatin + 5-Fluorouracil	8.7
	5-Fluorouracil	204	5-Fluorouracil	6.1

C



D

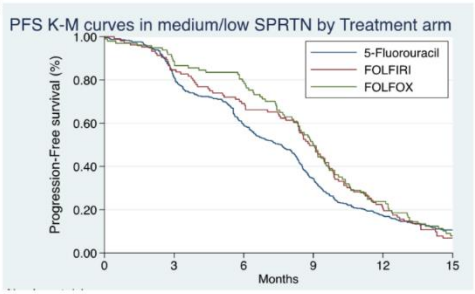
Statistical analysis

Comparison	Hazard Ratio (HR)	Confidence interval (95% CI)	p-value
FOLFIRI vs 5-Fluorouracil	0.67	0.48 - 0.93	0.02
FOLFOX vs 5-Fluorouracil	0.62	0.44 - 0.87	0.006
FOLFOX vs FOLFIRI	No observable difference between FOLFIRI and FOLFOX		

E

		Number of patients	Treatment	Median PFS (months)
SPRTN Medium and Low	FOLFIRI	105	Irinotecan + 5-Fluorouracil	8.8
	FOLFOX	97	Oxaliplatin + 5-Fluorouracil	9.0
	5-Fluorouracil	408	5-Fluorouracil	7.6

F



G

Statistical analysis

Comparison	Hazard Ratio (HR)	Confidence interval (95% CI)	p-value
FOLFIRI vs 5-Fluorouracil	0.88	0.70 - 1.10	0.26
FOLFOX vs 5-Fluorouracil	0.82	0.65 - 1.03	0.09
FOLFOX vs FOLFIRI	No observable difference between FOLFIRI and FOLFOX		

Figure 47. SPRTN as a predictive biomarker in CRC patients (Oxford and Leeds cohorts – 914 patients).

(A) 914 CRC patients were divided in two groups (SPRTN High, SPRTN Medium and Low), according to the SPRTN nuclear optical density quantified by HALO in their malignant epithelium (B) Summary of the 304 CRC patients expressing high levels of SPRTN, divided according to the first-line treatment they received (C) Progression free survival of the 304 patients expressing high levels of SPRTN. Patients were divided according to treatment, and their survival was plotted using Kaplan Meyer curves. (D) Statistical analysis of the data presented in Figure 46C. (E) Summary of the 610 CRC patients expressing medium or low levels of SPRTN, divided according to the first-line treatment they received (F) Progression free survival of the 610 patients expressing medium or low levels of SPRTN. Patients were divided according to treatment, and their survival was plotted using Kaplan Meyer curves. (G) Statistical analysis of the data presented in Figure 46F.

		Oxford cohort (n=400)	**Leeds cohort** (n=514)	**p-value**	**Total population** (n=914)
Age mean (sd)		63.4 (9.4)	62.8 (9.0)	0.32	63.1 (9.2)
Sex	Male	279 (66.7)	349 (70.4)		628 (68.7)
	Female	139 (33.3)	147 (29.6)	0.24	286 (31.3)
ECOG PS	0-1	392 (93.8)	449 (90.5)		841 (92.0)
	2	26 (6.2)	47 (9.5)	0.07	73 (8.0)
Previous resection	Yes	396 (94.7)	471 (95.0)		867 (94.9)
	No	22 (5.3)	25 (5.0)	0.88	47 (5.1)
Peritoneal mets	Yes	59 (14.1)	62 (12.5)		121 (13.2)
	No	356 (82.2)	424 (85.5)		780 (85.3)
	Unknown	3 (0.7)	10 (2.0)	0.52	13 (1.4)
Lung mets	Yes	139 (33.3)	197 (39.7)		336 (36.8)
	No	276 (66.0)	289 (58.3)		565 (61.8)
	Unknown	3 (0.7)	10 (2.0)	**0.03**	13 (1.4)
Liver mets	Yes	339 (81.1)	335 (67.5)		674 (73.7)
	No	77 (18.4)	151 (30.4)		228 (25.0)
	Unknown	2 (0.5)	10 (2.0)	**<0.0005**	12 (1.3)
Nodal mets	Yes	173 (41.4)	197 (39.7)		370 (40.5)
	No	242 (57.9)	289 (58.3)		531 (58.1)
	Unknown	3 (0.7)	10 (2.0)	0.73	13 (1.4)
Primary tumour location	Rectum	28 (6.7)	230 (46.4)		258 (28.2)
	Left	202 (48.3)	163 (32.9)		365 (39.9)
	Right	157 (37.6)	79 (15.9)		236 (25.8)
	Unknown	31 (7.4)	24 (4.8)	**<0.0005**	55 (6.0)
KRAS status	WT	204 (48.8)	173 (34.9)		377 (41.2)
	MUT	199 (47.6)	148 (29.8)		347 (38.0)
	Unknown	15 (3.6)	175 (35.3)	0.38	190 (20.8)
BRAF status	WT	362 (86.6)	297 (59.9)		659 (72.1)
	MUT	41 (9.8)	24 (4.8)		65 (7.1)
	Unknown	15 (3.6)	175 (35.3)	0.21	190 (20.8)
Median PFS (months)		8.3 (4.4-10.4)	7.7 (3.4-10.4)	0.52	8.0 (3.7-10.4)
Median OS (months)		15.7 (9.7-24.0)	15.2 (7.9-23.1)	0.45	15.4 (9.0-23.6)
Clinical response (12 weeks)	Complete response	4 (1.0)	5 (1.5)		9 (1.0)
	Partial response	161 (39.2)	94 (28.8)		255 (27.9)
	Stable disease	155 (37.7)	167 (51.2)		322 (35.2)
	Progressive disease	91 (22.1)	60 (18.4)	**0.002**	151 (16.5)

Table 20. Main differences between Oxford and Leeds cohorts.

Summary

In this final chapter, the possible role of SPRTN as a biomarker to predict the efficacy of different chemotherapeutic regiments in colorectal cancer was investigated for the first time. In particular, the work aimed to identify if colorectal cancers expressing higher levels of SPRTN could be characterized by a worse prognosis while treated with an irinotecan-based regimen like FOLFIRI. This could be due to the fact that higher levels of SPRTN have the potential to better repair the TOP1ccs induced by irinotecan, assuring the survival of cancer cells and the subsequent cancer progression.

In order to test this hypothesis, the FFPE tissue previously obtained from patients enrolled in the MRC FOCUS trial was immunohistochemically stained with two different SPRTN antibodies (SPRTN Atlas and SPRTN Abcam). The signal provided by the two different antibodies was proved to be consistent with each other, as confirmed in Figure 43.

First, one small exploratory cohort including 45 CRC patients treated with FOLFIRI was stained with the SPRTN Atlas antibody, and the prognostic role of SPRTN was investigated (Figure 44). In line with the original hypothesis, patients expressing higher levels of SPRTN and treated with FOLFIRI were found to be characterized by a worse progression free survival.

After the initial encouraging result, the analysis was expanded to include 914 CRC patients treated with FOLFIRI, FOLFOX or 5-Fluorouracil. As FOLFIRI and FOLFOX are generally characterized recognized to provide a similar benefit in progression free survival, this setting offered the opportunity to evaluate the role of SPRTN as a predictive biomarker in CRC chemotherapy. This analysis had to be divided in two different cohorts, due to some bureaucratic challenges in sharing clinical data between different institutes: Oxford cohort (400 patients) and Leeds cohort (514 patients).

Patients expressing high levels of SPRTN in the Oxford cohort and treated with FOLFIRI were characterized by a worse progression free survival than the patients treated with FOLFOX, with a difference in median PFS of 2.27 months (Figure 46). This result was once again in line with the potential role of SPRTN as a biomarker to predict the efficacy of irinotecan-based chemotherapy. However, when this analysis was expanded by including also the Leeds cohort, no statistically significance difference in PFS could be observed in CRC patients expressing high SPRTN levels and treated with FOLFIRI or FOLFOX (Figure 47). In order to determine which factors might have been responsible of the difference results observed in the full cohort, the main clinical and molecular differences between the Oxford cohort and Leeds cohort were compared. Interestingly, one of the main differences which could be observed was the fact that most of the patients from Leeds were actually CRC originating from the rectum. While the difference susceptibility to DPCs in colon and rectal cancers has never been investigated before, it is important to understand that these two entities are characterized by significant differences in anatomy, embryological origins and most common mutations [11, 265]. For example, the incidence of BRAF mutations and MSI-high tumours has been showed to gradually decrease from the proximal colon to the rectum [266]. While it remains unclear if rectal cancer might indeed be less reliant on SPRTN in order to repair the TOP1ccs induced by irinotecan-based chemotherapy, this represents an interesting area of scientific inquiry.

Chapter 4: Discussion

The work presented in this thesis aimed to investigate the possible role of SPRTN protease in mediating the response to DNA damage induced by commonly used chemotherapeutic drugs in colorectal cancer, including 5-fluorouracil, oxaliplatin and irinotecan. Since SPRTN was originally described as a gene and protein involved in the development of a progeroid syndrome characterized by premature ageing, early onset hepatocellular carcinoma and genomic instability (Ruijs-Aalfs syndrome) in 2014, our knowledge of SPRTN and its role in DNA-repair has progressively started to expand [165, 267]. In particular, SPRTN has been identified as a DNA-dependent metalloprotease involved in DPC repair in 2016, and SPRTN-depleted cells have been confirmed to be more sensitive to the cytotoxic effects of numerous DPC-inducing drugs like formaldehyde, etoposide, cisplatin and camptothecin [136, 158, 164].

In order to better understand the potential role of SPRTN in colorectal cancer, it was important to investigate how SPRTN inactivation affects colorectal cancer cell lines' response to chemotherapeutic drugs commonly used in this setting. In particular, SPRTN was known to have a role in the repair of TOP1ccs, an enzymatic type of DPC induced by topoisomerase poisons like camptothecin and irinotecan [160]. In addition, previous work in our laboratory identified that SPRTN collaborates with the ATPase p97 and its cofactor TEX264 in the initial proteolysis of TOP1ccs. This proteolysis is essential for the subsequent cleavage of the remnant peptide from DNA by TDP1 [140]. While all these results were obtained while studying the TOP1ccs induced by camptothecin, another topoisomerase inhibitor, we believe the SPRTN role on TOP1ccs induced by irinotecan, a camptothecin-derivate drug used in clinical settings, should be the same. As irinotecan is frequently used in combination with 5-fluorouracil in colorectal cancer patients, this represented an ideal setting to evaluate the possible role of SPRTN as a prognostic and predictive biomarker. FOLFIRI (5-Fluorouracil + Irinotecan) and FOLFOX (5-Fluorouracil + Oxaliplatin), the two mains chemotherapeutic regimens used to treat metastatic

colorectal cancer, are generally consider to present similar benefits on progression free survival [25, 26]. Accordingly, if SPRTN expression could predict which patients are more likely to respond to FOLFIRI or FOLFOX, this could represent a significant improvement in the clinical management of colorectal cancer patients.

Therefore, to address my scientific questions, I focused my DPhil research into two main areas, namely:

1. Investigating the role of SPRTN inactivation in CRC cells response to chemotherapy (5-fluorouracil, irinotecan, oxaliplatin)
2. Investigating the role of SPRTN as a biomarker in metastatic CRC patients.

In short, my results demonstrated that:

1. SPRTN inactivation causes severe defects in DNA repair in CRC cells, consistent with previous reports using non-CRC cell lines (Figure 37) [136, 158, 164]. However, the survival of CRC cell lines to chemotherapy is not dependent on SPRTN (Figure 30-34, Figure 36). This is likely related to the fact that CRC cells present an upregulation of alternative DPC proteases like FAM111A (Figure 39B), along with an increase in anti-apoptotic signaling after SPRTN depletion (Supplementary Table 6).
2. SPRTN inactivation causes the downregulation of multiple ribosomal genes (Figure 41, Supplementary Table 5). This is in line with the idea that DPCs, beside representing significant obstacles to DNA replication, also have the potential to significantly hinder DNA transcription.
3. SPRTN is unlikely to represent an useful biomarker to predict the efficacy of chemotherapy in CRC patients (Figure 47). However, SPRTN could still represent an interesting biomarker in other cancer patients treated with alternative DPC-inducing drugs. For example, small cell lung cancer patients are

frequently treated with cisplatin and etoposide, and both these drugs have been reported to be able to induce DPCs [216, 268, 269].

The role of SPRTN in mediating the survival of CRC cells treated with chemotherapy

Before being able to start addressing the role of SPRTN in polychemotherapeutic regimens like FOLFIRI or FOLFOX in patients, the initial step included trying to understand if SPRTN depletion could influence the efficacy of the single chemotherapeutic agents (5-fluorouracil, irinotecan and oxaliplatin) in colorectal cancer cell lines. Addressing the contribution of SPRTN in repairing the DNA damage induced by the each drug was considered important because, while SPRTN depletion has been associated to increased sensitivity to topoisomerase inhibitors and platinum compounds [136], its possible role mediating the response to DNA-protein or RNA-protein crosslinks which could be induced by nucleotide analogues like 5-Fluorouracil remained largely unexplored.

Three different colorectal cancer cells (HCT116, HT55, RKO) were initially analyzed in this thesis due to their different genetic background (summarized in Supplementary Table 1). SPRTN was transiently inactivated either by siRNA or by auxin-inducible degradation, and the efficacy of 5-fluorouracil, irinotecan and oxaliplatin was investigated by colony forming assays (Figure 30-34, Figure 36). These initial results were particularly intriguing because, while SPRTN-mediated DPC-repair has been showed to be important in assuring the survival of numerous cell lines and lower eukaryotes like C. elegans [158, 164], colorectal cancer cells survival remained largely unaffected by SPRTN depletion. This was originally considered to be potentially associated to an insufficient level of SPRTN depletion obtained through siRNAs, despite the use of those siRNA was previously validated in other papers and was also confirmed by RT-qPCR (Figure 30B) [158]. Furthermore, the development of a SPRTN-AID system, assuring higher levels of SPRTN depletions, still confirmed the observed phenotypes (Figure 36A-D). Another hypothesis to explain this phenomenon was related to the fact that SPRTN is not the only DPC-protease. Recently identified proteases, such FAM111A/B, GCNA,

and DDI1, might also have the ability to proteolyze DNA-protein crosslinks induced by the investigated drugs in colorectal cancer cells. This hypothesis was supported by the result of the RNA-seq performed in HCT116 cells, which confirmed a higher basal expression of FAM111A, further increased following SPRTN depletion (Figure 39B). However, the increased expression of FAM111A might still not be sufficient to guarantee an efficient DPC-repair, as suggested by the analysis of 53BP1 and γH2AX foci formed after chemotherapy in HCT116 SPRTN-AID cells, which confirmed the presence of higher levels of DNA damage in SPRTN-depleted cells (Figure 37). Overall, these results suggest that colorectal cancer cells exhibit significant resistance to SPRTN depletion. They are able to survive SPRTN inactivation, despite experiencing a substantial increase in DNA damage when treated with topoisomerase inhibitor or platinum compounds (Figure 48A-C).

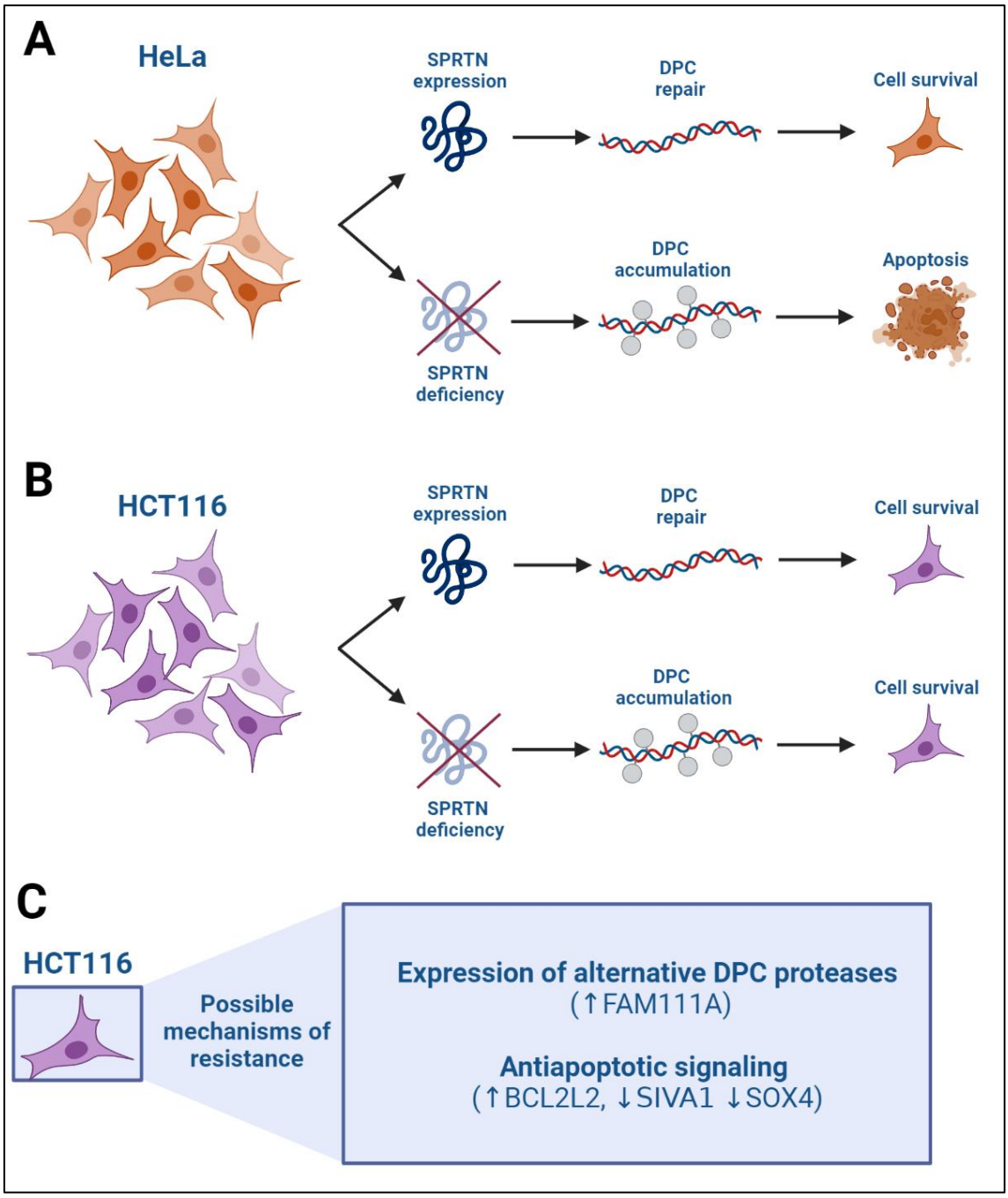

A
HeLa
SPRTN expression
DPC repair
Cell survival
DPC accumulation
Apoptosis
SPRTN deficiency
B
HCT116
SPRTN expression
DPC repair
Cell survival
DPC accumulation
Cell survival
SPRTN deficiency
C
HCT116
Possible mechanisms of resistance
Expression of alternative DPC proteases
(↑FAM111A)
Antiapoptotic signaling
(↑BCL2L2, ↓SIVA1 ↓SOX4)

Figure 48. Consequences of SPRTN-depletion in different cell lines.

(A) Following SPRTN depletion, HeLa cells present an increased sensitivity to DPC-inducing drugs and an activation of apoptosis. (B) Following SPRTN depletion, HCT116 cells do not become more sensitive to most DPC-inducing drugs and they continue to proliferate despite increased levels of DNA damage. (C) Possible mechanisms of resistance to SPRTN depletion which have been identified in HCT116 cells.

The consequences of SPRTN depletion on the transcriptome of CRC cells

To further investigate the mechanisms which might be responsible for the SPRTN-resistant phenotype in colorectal cancer cells, the list of differentially expressed genes (DEGs) obtained from the RNA-seq performed on HCT116 SPRTN-AID cells was investigated and compared to the data obtained from HeLa cells. Interestingly, after SPRTN inactivation both HCT116 and HeLa cells presented an upregulation of CDK6, an important positive cell cycle regulator of the G1-S transition. CDK6 phosphorylates the Retinoblastoma protein (Rb), and thus promotes the release and activation of the transcription factor E2F, the main driver of G1-S transition (Supplementary Table 5, Figure 49) [247, 270]. At the same time, both cell lines downregulated the expression of cyclin-dependent kinase inhibitor 2A (CDKN2A), a gene that encodes for two distinct kinase inhibitor and tumor suppressor proteins, $p16^{INK4A}$ and $p14^{ARF}$ (Supplementary Table 5, Figure 49). $p16^{INK4A}$, in particular, is involved in arresting cells in G1 by preventing binding of CDK4/6 to cyclin D1, consequently suppressing Rb phosphorylation and inhibiting the release of the transcription factor E2F [271]. $p14^{ARF}$, instead, has a role in mediating cell cycle arrest through p53 stabilization and activation, which is secondary to the inhibitory effect of $p14^{ARF}$ on two ubiquitin ligases normally targeting p53, Mdm2 and ARF-BP1 [272]. Taken together, both the upregulation of CDK6 and the downregulation of CDKN2A promote G1-S transition. Another interesting consequence of SPRTN depletion which can be observed in both cell lines consist in the upregulation of CDC25B (Supplementary Table 5, Figure 49). CDC25B is a gene encoding for a phosphatase involved in the dephosphorylation and consequent activation of CDK1, stimulating G2-M transition and the beginning of mitosis [273-275]. Despite these similarities, only the cell cycle progression of HeLa cells was significantly impaired after SPRTN depletion, as observed in Figure 38, while colorectal cancer cells remained largely unaffected. This phenotype could be potentially explained by the fact that only HeLa cells were characterized by an upregulation of ATM after SPRTN

depletion (log_2 [HeLa SPRTN $^{-/-/+}$ / HeLa WT] = 0.21569, q-value < 0.05), which is generally recognized to play a dominant role in cell cycle delay after DNA damage, known as DNA damage checkpoint, especially in case of DSBs [276].

An interesting scientific question to be addressed involves how colorectal cancer cells are able to sustain the increased amount of DNA damage they present after SPRTN depletion, especially while being treated with DNA-damaging drugs like cisplatin, oxaliplatin or camptothecin (Figure 37). To address this question, the list of DEGs in HCT116 and HeLa cells was investigated to determine if they presented any significant alteration of the apoptosis pathway (Supplementary Table 3-4). Strikingly, HCT116 cells were found to present a stronger antiapoptotic signal after SPRTN depletion, mostly led by the upregulation of BCL2L2 (log_2 [HCT116 Auxin / HCT116 DMSO] = 0.293997, q-value < 0.05) and the downregulation of SOX4 (log_2 [HCT116 Auxin / HCT116 DMSO] = -0.39885, q-value < 0.00001) and SIVA1 (log_2 [HCT116 Auxin / HCT116 DMSO] = -0.3129, q-value < 0.05). In comparison, HeLa cells were characterized by a classic proapoptotic signature instead, due to the upregulation of caspases like CASP8 (log_2 [HeLa SPRTN $^{-/-/+}$ / HeLa WT] = 0.216043, q-value < 0.05), CASP9 (log_2 [HeLa SPRTN $^{-/-/+}$ / HeLa WT] = 0.933999, q-value < 0.00001) and CASP3 (log_2 [HeLa SPRTN $^{-/-/+}$ / HeLa WT] = 0.324303, q-value < 0.05). The effect of SPRTN depletion on the apoptosis pathway of HeLa and HCT116 cells has been summarized in Figure 50A-B.

The upregulation of BCL2L2 observed in colorectal cancer cells after SPRTN depletion, a gene encoding for BCL-W, is particularly intriguing. This is because BCL-W is an anti-apoptotic protein which has been previously associated to reduced cell apoptosis under cytotoxic conditions and to the survival of drug-resistant cells [277-279]. Along with it, the downregulation of SOX4, a DNA damage sensor critical for the activation of p53, is likely to further contribute to the overall resistance to the consequences of SPRTN depletion in HCT116 cells. SOX4 has been reported to stabilize p53 by blocking Mdm2-mediated p53 degradation, and to promote apoptosis and cell cycle arrest in a p53-

dependent manner [280]. Lastly, SIVA1 is a gene encoding for a pro-apoptotic E3 ubiquitin ligase that regulates apoptosis by stimulating the proteasome-mediated degradation of Bcl-x(L), an anti-apoptotic member of the BCL2 family of proteins [281, 282]. Accordingly, SIVA1 downregulation also has the potential to indirectly contribute to the increased anti-apoptotic signalling observed in SPRTN-depleted colorectal cancer cells.

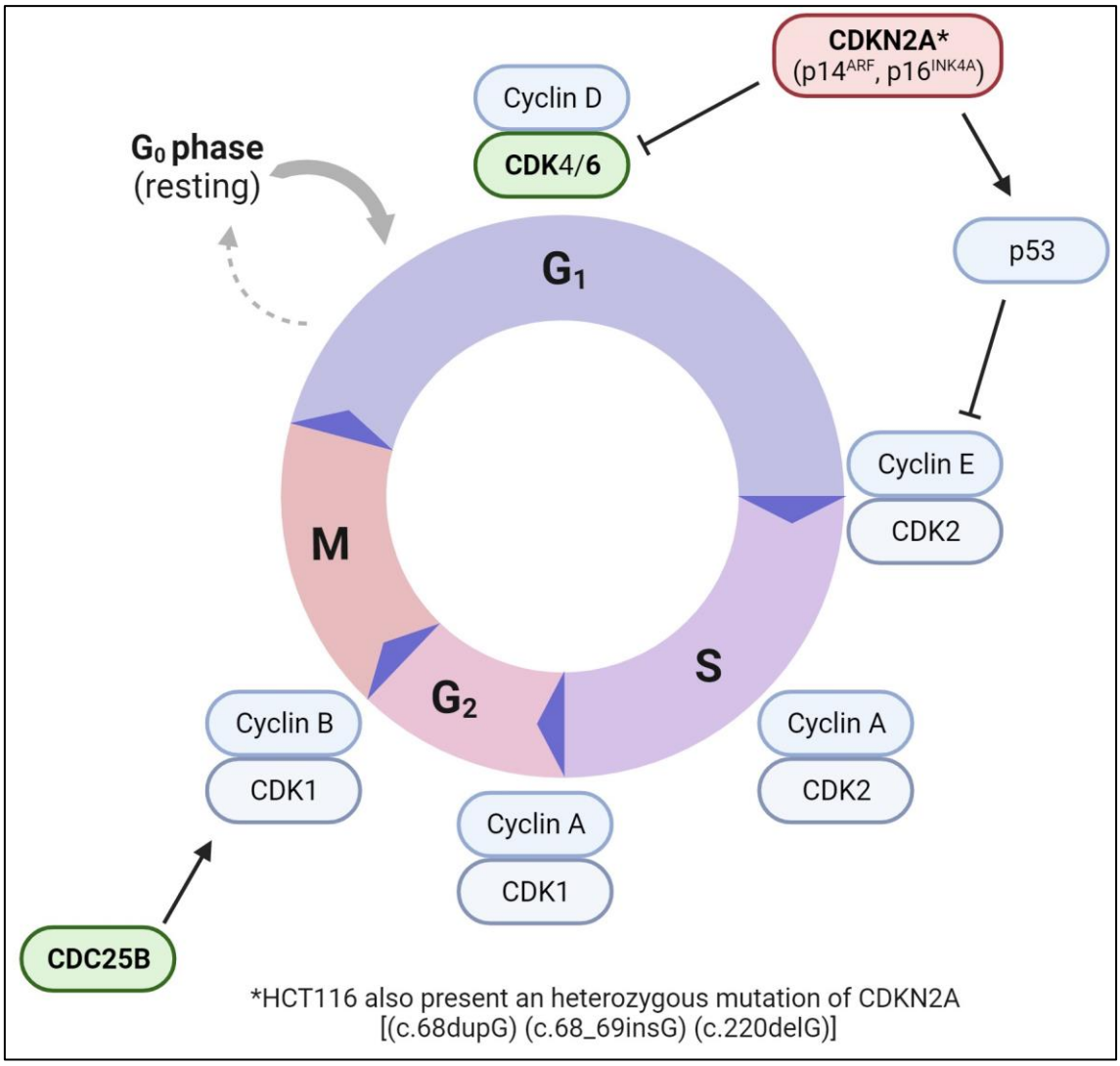

CDKN2A*
(p14ARF, p16INK4A)
Cyclin D
CDK4/6
G0 phase
(resting)
G1
p53
Cyclin E
CDK2
M
S
G2
Cyclin B
CDK1
Cyclin A
CDK2
Cyclin A
CDK1
CDC25B
*HCT116 also present an heterozygous mutation of CDKN2A
[(c.68dupG) (c.68_69insG) (c.220delG)]

Figure 49. SPRTN-depletion and its consequences on the cell cycle.

Effect of SPRTN depletion on the cell cycle of HCT116 and HeLa cells. Both cell lines were characterized by an upregulation of CDK6 and CDC25B (green) and a downregulation of CDKN2A (red). However, only HeLa cells presented an ATM upregulation (not shown).

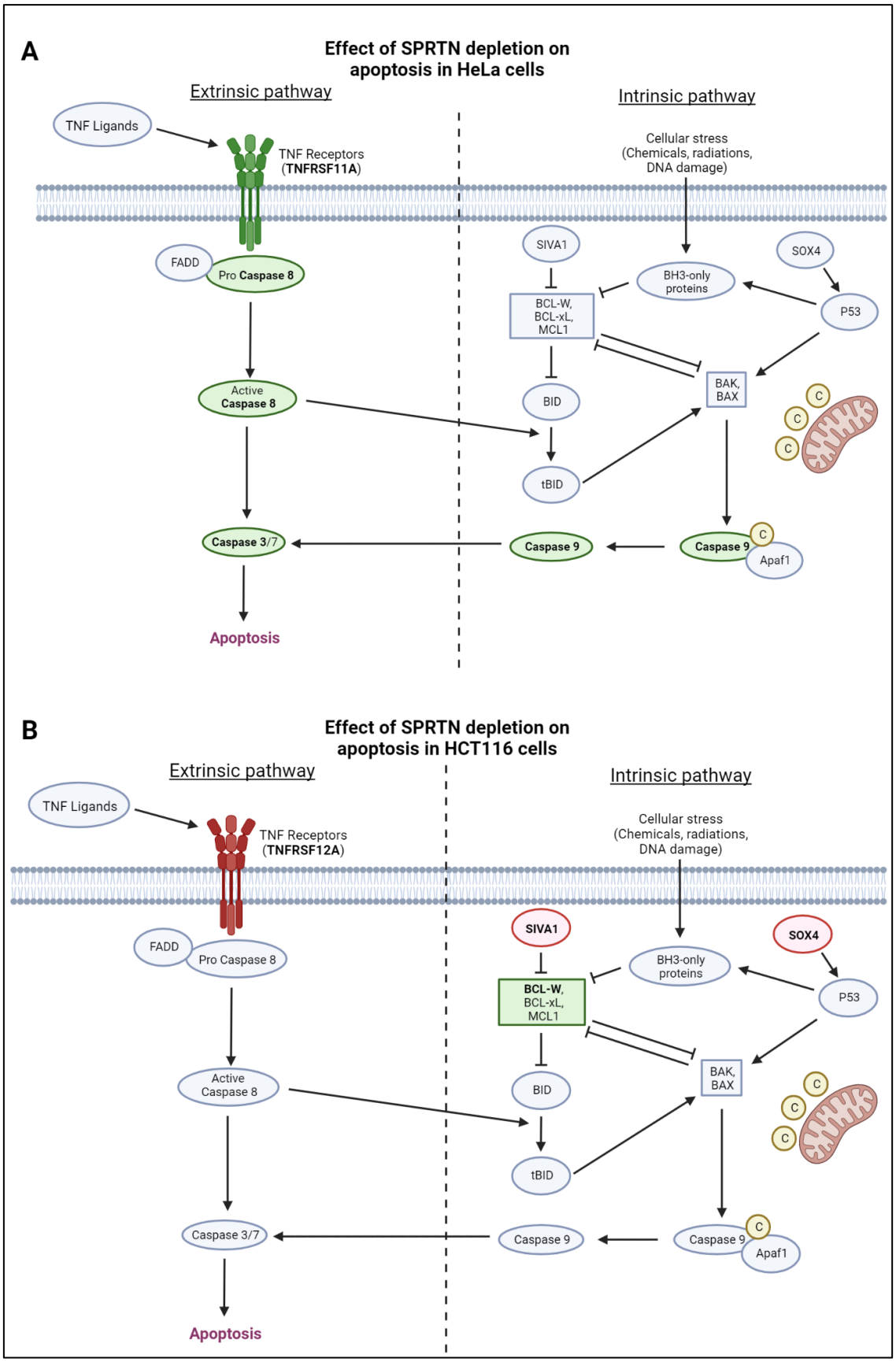

A
Effect of SPRTN depletion on apoptosis in HeLa cells
Extrinsic pathway
Intrinsic pathway
TNF Ligands
TNF Receptors (TNFRSF11A)
Cellular stress (Chemicals, radiations, DNA damage)
FADD
Pro Caspase 8
SIVA1
SOX4
BH3-only proteins
BCL-W, BCL-xL, MCL1
P53
Active Caspase 8
BID
BAK, BAX
C
C
C
tBID
Caspase 3/7
Caspase 9
Caspase 9
C
Apaf1
Apoptosis
B
Effect of SPRTN depletion on apoptosis in HCT116 cells
Extrinsic pathway
Intrinsic pathway
TNF Ligands
TNF Receptors (TNFRSF12A)
Cellular stress (Chemicals, radiations, DNA damage)
FADD
Pro Caspase 8
SIVA1
SOX4
BH3-only proteins
BCL-W, BCL-xL, MCL1
P53
Active Caspase 8
BID
BAK, BAX
C
C
C
tBID
Caspase 3/7
Caspase 9
Caspase 9
C
Apaf1
Apoptosis

Figure 50. Regulation of apoptosis in SPRTN-depleted cells.

HeLa and HCT116 cells regulation of apoptosis after SPRTN depletion. (A) HeLa cells were characterized by the upregulation of pro-apoptotic genes after SPRTN depletion, involving both the extrinsic and intrinsic apoptotic pathway. (B) HCT116 cells were characterized by an increased anti-apoptotic response after SPRTN depletion, which could justify their resistant phenotype. Upregulated genes are indicated in green, while downregulated genes are indicated in red.

The role SPRTN as a possible biomarker in CRC patients

Despite the results presented in this thesis, which indicated that colorectal cancer cell lines are able to resist to the deleterious consequences of SPRTN depletion, this might not necessarily be the case in more complex system like the human body. In fact, the contribution of the microenvironment and the increased cellular heterogeneity of patient tissues could play a significant role in determining clinical outcomes [283-286]. To support this idea, a preliminary analysis on 12 CRC patients was performed by the Ramadan Laboratory in 2019, in collaboration with Professor Tim Maughan at the University of Oxford. SPRTN protein expression, assessed by IHC, seemed to directly correlate with the poor response to TOP1-inhibitors like Irinotecan, and a patent establishing the possible role of SPRTN as a predictive biomarker was filed in 2020 [263]. According to these early results, SPRTN overexpression in tumor cells might allow a better repair of TOP1ccs induced by Irinotecan, worsening patient prognosis and influencing their clinical response to FOLFIRI chemotherapy. The early findings originally obtained in 2019 were significantly expanded in this thesis, by including 914 CRC patients originally enrolled in the MRC FOCUS trial, and investigating the possible role of SPRTN as a predictive biomarker [204, 287, 288].

SPRTN as a possible biomarker for FOLFIRI chemotherapy (Exploratory Cohort)

A preliminary analysis aimed to investigate the role of SPRTN as a predictive biomarker in CRC patients treated with Irinotecan-based chemotherapy (FOLFIRI), to confirm the early data filed in the aforementioned patent [263], was initially performed. This time, the analysis was expanded to 45 CRC patients treated with FOLFIRI (Figure 44). Patient were divided on tertiles according to their SPRTN expression assessed by immunohistochemistry and digital pathology, and their progression free survival (PFS) was analyzed. Interestingly, in line with the original hypothesis that SPRTN overexpression has the potential to improve TOP1ccs repair in cancer cells and worsen patient's prognosis, patients expressing high levels of SPRTN and treated with FOLFIRI

showed a 3.96 months reduction in their median PFS when compared with patients expressing low levels of SPRTN (Figure 44, Supplementary Table 7).

SPRTN as a predictive biomarker in FOLFIRI vs FOLFOX (Oxford and Leeds Cohorts)

Once the possible role of SPRTN to determine the prognosis of patients treated with FOLFIRI was confirmed, the next step included determining if SPRTN expression could also be used as a predictive biomarker to determine which patients are more likely to benefit from FOLFIRI or FOLFOX chemotherapy. This represents a particularly interesting clinical question, as FOLFIRI and FOLFOX are both frequently used as first-line treatment for metastatic CRC and they are both presenting similar level of efficacy. In particular, FOLFIRI presents a 31-56% response rate (RR) and 7.0-8.5 months median PFS, while FOLFOX is characterized by 34-54% RR and 7.0-8.0 months median PFS, depending on the trial [25, 26].

To determine the role of SPRTN as a predictive biomarker, a total of 914 CRC patients treated with FOLFIRI, FOLFOX or 5-fluorouracil alone were divided in tertiles according to their SPRTN expression. These patients were also further divided in two different cohorts (Oxford cohort [400 patients] and Leeds cohort [514 patients]), according to the University which had access to their clinical data. The initial analysis aiming to establish the role of SPRTN as a predictive biomarker included only the patients enrolled in the Oxford cohort, and this was then further expanded in a final analysis including all 914 patients (Oxford cohort + Leeds cohort), performed by the research group of professor Jenny Seligmann at the University of Leeds.

The initial analysis performed in the Oxford cohort provided some particularly exciting results, as patients expressing high levels of SPRTN and treated with FOLFIRI seemed to present a significantly worse PFS when compared to patients treated to FOLFOX (7.86 months for FOLFIRI, 10.13 months for FOLFOX) (Figure 46B-D). However, this result was not confirmed in the final analysis including all 914 CRC patients, where no statistically significance difference in PFS could be observed between patients treated

with FOLFIRI or FOLFOX (8.6 months for FOLFIRI, 8.7 months for FOLFOX) (Figure 47B-D). Furthermore, no significance difference in PFS could be observed in CRC patients treated with FOLFIRI and expressing different levels of SPRTN (8.6 months for SPRTN High, 8.8 months for SPRTN Medium and Low) (Figure 47B-G).To determine which factors might contribute to the different observed results, the main clinical and molecular variables between the Oxford cohort and the Leeds cohort were compared in Table 20. Interestingly, one of the most striking differences between the two cohorts was the fact that most patients from Leeds presented cancers originating for the rectum, which have the potential to significantly differ from cancers originating from the colon. Colon and rectal cancers are respectively characterized by a different incidence of the most common mutations responsible for CRC pathogenesis and, for example, BRAF mutations in colorectal cancer increase gradually from the rectum to ascending colon subsites [266]. Despite this, no significant difference in the incidence of BRAF or KRAS mutations was reported between the Oxford and Leeds cohorts (Table 20).

Another possible explanation of the divergent results between the Oxford and Leeds cohort was considered to be the fact that rectal cancers are frequently treated with neoadjuvant chemoradiotherapy, and this might have the potential to influence the response to further lines of chemotherapy within the Leeds cohort. This idea was discussed with the research team in Leeds and, to address this hypothesis, the analysis was repeated while including only the previously untreated colon cancer patients. However, even in this case, no statistically significant association was found between SPRTN expression and response to FOLFIRI or FOLFOX (data not shown). Taking into account the limitations mentioned in the Oxford cohort, due to the limited number of patients treated with either FOLFIRI or FOLFOX, and the overall results obtained from the final analysis of 914 CRC patients (Figure 47), SPRTN was then considered unlikely to represent a valuable predictive biomarker for selecting the most appropriate chemotherapeutic regimen in colorectal cancer.

Chapter 5: Future work

In this thesis, the possible role of SPRTN in mediating the repair of DPCs induced by chemotherapy in colorectal cancer has been investigated for the first time. Interestingly, CRC cells seemed to be particularly resistant to the deleterious effect of SPRTN depletion on the cell cycle and on cell survival after DNA damage (Figure 30-34, Figure 36, Figure 38), which have previously been reported by different research groups in other cell lines and systems [158, 164]. This resistance could be caused by different mechanisms, including:

1. Reliance on alternative DPC proteases

FAM111A seems to be the alternative DPC protease mainly expressed in colorectal cancer cells (Figure 39B). In order to understand if these cells are mostly reliant on FAM111A as the main DPC protease involved in DPC repair, the sensitivity of CRC cells to different DPC-inducing drugs could be assessed after FAM111A depletion by siRNA or after the generation of FAM111A$^{-/-}$ cell lines in SPRTN-inactivated cells. Despite this, the results obtained in Figure 37 suggests that CRC cells indeed present higher levels of DNA damage after SPRTN depletion, so the upregulation of alternative DPC proteases does not appear to be sufficient to efficiently repair DPCs and avoid the generation of DSBs.

2. Evasion of apoptosis

The induction of apoptosis has long been considered to be a major player in determining cell death after DNA-damage, and colorectal cancer is known to develop various mechanisms to evade the apoptotic process [289]. The tumor suppressor p53, in particular, has been reported to be mutated in over 40-50% of colorectal cancer and it is recognized as one of the major players responsible for inducing apoptosis [290-292]. Despite p53 is not mutated in HCT116 cells (Supplementary Table 1), the downregulation of SOX4, a DNA damage sensor responsible for stabilizing and

activating p53, observed in these cells after SPRTN depletion has the potential to significantly reduce p53-mediated apoptosis (Figure 50) [280]. This finding, based on RNA-seq data, should be validated by measuring p53 activation in HCT116 cells before and after SPRTN depletion, either by western-blot (determining the expression of p53 pro-apoptotic transcriptional targets like BBC3 [PUMA]) or by ELISA (using oligonucleotides capable of binding to activated p53) [293, 294]. However, the BBC3 gene is not downregulated in SPRTN-depleted HCT116 cells (Supplementary Table 3), so the evasion of apoptosis observed in this cell line is unlikely to be only reliant on a p53-mediathed mechanism. As an alternative, the upregulation of BCL2L2, an anti-apoptotic gene encoding for BCL-W, could play an important role in mediating the evasion of apoptosis in HCT116 cells (Supplementary Table 3, Figure 50). BCL-W has previously been associated to the survival of cells under cytotoxic conditions, and its overexpression after SPRTN depletion could be validated by western blot [277, 295]. Overall, even before trying to understand which gene is more important in mediating the evasion to apoptosis in SPRTN-depleted colorectal cancer cells, this evasion should first be confirmed by using different assays. A number of techniques have been developed over the years to reliably measure the activation of the apoptotic pathway, from caspase activity assays to annexin V assays [296-298], and they could be applied to HCT116 cells before and after SPRTN depletion.

Beside trying to understand the molecular mechanisms responsible for mediating the resistance to the deleterious effects of SPRTN depletion in colorectal cancer cells, the results presented in this thesis suggest at least a couple of other areas of potential scientific interest. They include:

1. Transcription of ribosomal DNA

A particularly interesting finding which was observed from the RNA-seq performed in HCT116 and HeLa cells was the fact that both cells presented a significant downregulation of multiple ribosomal genes after SPRTN depletion (Supplementary Table 5, Figure 51). Ribosomal genes are generally considered to be the most expressed genes in most cell types. For example, it is estimated that in cells over 80-90% of total RNA transcription is dedicated to the production of ribosomal RNA [299]. As unrepaired DNA-protein crosslinks represent significant physical obstacles to both DNA replication and transcription, it is reasonable to expect ribosomal genes transcription to be the one most likely to be affected by defects in DPC-repair [260, 300]. To guarantee an efficient production of ribosomal RNA (rRNA), the ribosomal genes are concentrated in specific nuclear domains characterized by the highest RNA synthesis, the nucleoli [301, 302]. This characteristic can be taken advantage of while planning an experiment to confirm the presence of unrepaired DNA-protein crosslinks in the nucleoli of SPRTN-depleted cells. In particular, the nucleoli of cultured cells can be isolated [303, 304] and the nucleolar DNA could then be analyzed by rapid approach to DNA adduct recovery (RADAR) or advanced recovery of K^+-SDS precipitates (ARK) assays to confirm the presence of DPCs located within this specific region of the genome [305, 306].

2. Role of SPRTN as a biomarker in other cancers

As previously discussed, based on the results presented in this thesis (Figure 47), SPRTN is not likely to represent a useful biomarker to predict the efficacy of FOLFIRI vs FOLFOX chemotherapy. However, numerous other drugs frequently used in the clinical

setting have been recognized to be able to induce DPCs. This includes cisplatin and etoposide, which are commonly combined to treat patients affected by small cell lung cancer (SCLC) [216, 268, 269]. In this case, for example, SPRTN could be investigated as a possible prognostic biomarker in cisplatin and etoposide chemotherapy. Furthermore, the recent approval of immunotherapeutic agents to treat SCLC patients also makes it possible to investigate the role of SPRTN as a predictive biomarker while comparing the efficacy of immunotherapy and standard chemotherapy (cisplatin and etoposide) in this patients [307]. Both these aim could be achieved by applying the same type of experiments which have been presented on this thesis, by staining with an appropriate SPRTN antibody tissue samples obtained from SCLC patients and investigating their survival according to SPRTN protein expression.

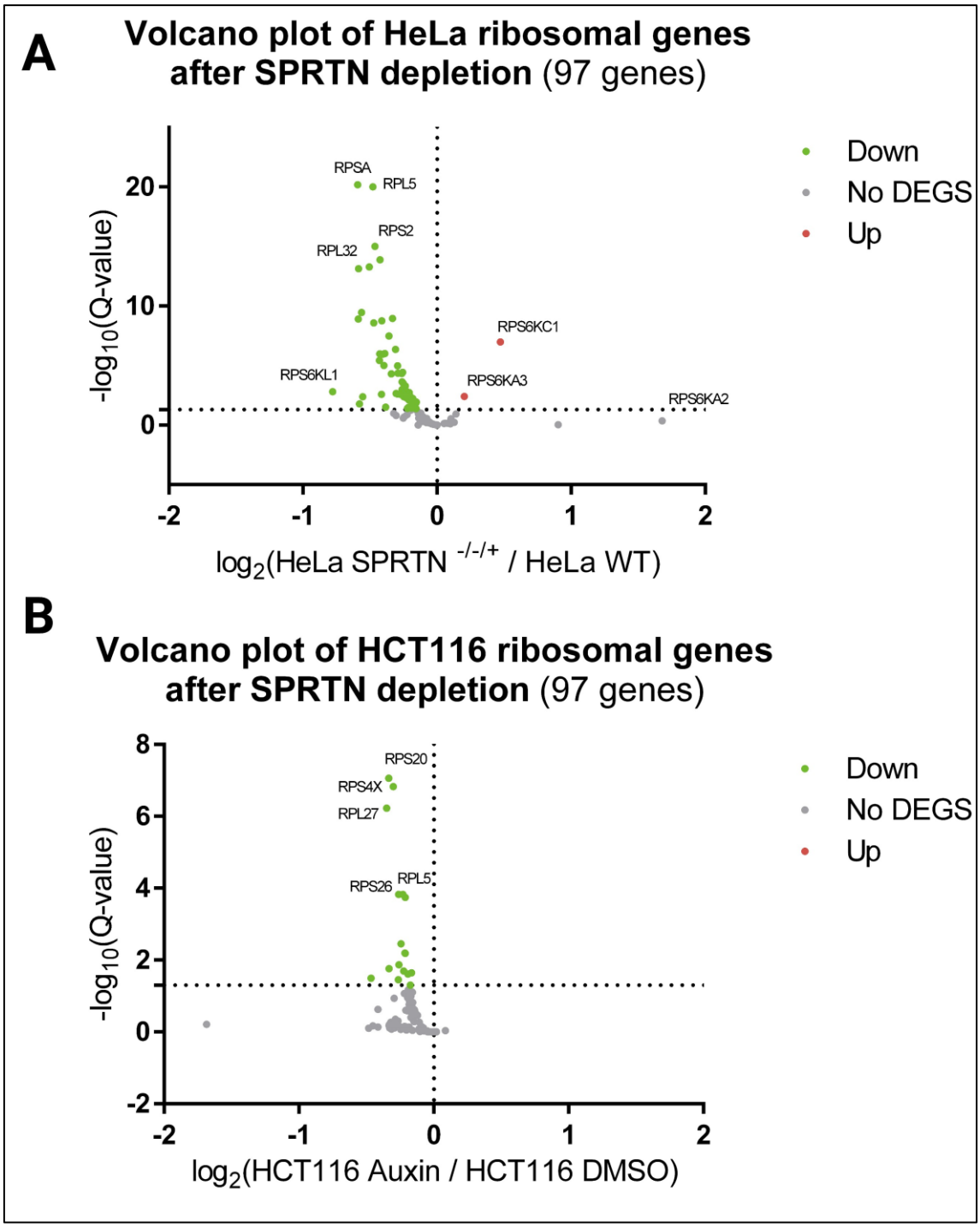

A
Volcano plot of HeLa ribosomal genes after SPRTN depletion (97 genes)
RPSA
RPL5
RPS2
RPL32
RPS6KC1
RPS6KL1
RPS6KA3
RPS6KA2
Down
No DEGS
Up
-log10(Q-value)
log2(HeLa SPRTN -/-/+ / HeLa WT)
B
Volcano plot of HCT116 ribosomal genes after SPRTN depletion (97 genes)
RPS20
RPS4X
RPL27
RPS26
RPL5
Down
No DEGS
Up
-log10(Q-value)
log2(HCT116 Auxin / HCT116 DMSO)

Figure 51. Volcano plots of ribosomal DEGs in SPRTN-depleted cells.

(A) Volcano plot of ribosomal DEGs in HeLa WT and HeLa SPRTN $^{-/-/+}$. (B) Volcano plot of ribosomal DEGs in HCT116 SPRTN-AID cells with or without the addition of auxin for 4 days.

Supplementary tables

Cell Line	Disease	Doubling time	Classification and main mutations
HCT116	Colon carcinoma	17 hours [308]	Microsatellite instability: Instable (MSI-high) [309] CMS classification: CMS1 [310] or CMS4 [311] CRIS classification: CRIS-D [310] Heterozygous mutations: - BRCA2 (c.8021dupA) (c.8021_8022insA) [312] - CDKN2A (c.68dupG) (c.68_69insG) (c.220delG) [312] - CTNNB1 (c.133_135delTCT) [313] - EP300 (c.5099delA) [314] - KRAS (c.38G>A) [315] - PIK3CA (c.3140A>G) [316] Homozygous mutations: - ACVR2A homozygous mutation (c.1310delA) [317] - TGFBR2 homozygous mutation (c.383delA) [317] No TP53 mutations [318]
HT55	Colon carcinoma	28 hours [319]	Microsatellite instability: Stable (MSS) [320] CMS classification: CMS2 [321] CRIS classification: Unknown Heterozygous mutations: - APC (c.3391C>T) (c.3907C>T) (c.4389A>T) [322] Homozygous mutations: - TP53 (c.638G>T) [323]
RKO	Colon carcinoma	21 hours [319]	Microsatellite instability: Instable (MSI-high) [324] CMS classification: CMS1 [310] CRIS classification: CRIS-D [310] Heterozygous mutations: - ACVR2A (c.1310delA) [317] - BRAF (c.1799T>A) [315] - PIK3CA (c.3140A>G) [324] - TGFBR2 (c.383delA) (c.1355T>C) [317] No TP53 mutations [323]

HeLa	HPV-related endocervical adenocarcinoma	30 hours [325]	Microsatellite instability: Stable (MSS) [326] Human papillomavirus status: Positive (HPV18) [325] Chromosome abnormalities: - Hypertriploid (3n+), with a variable total number of chromosomes (76-80) [327, 328] Relevant information for this DPhil Thesis: - HeLa cells have three copies of the gene encoding for SPRTN, and a SPRTN haploinsufficient cell line (HeLa SPRTN $^{-/-/+}$) was previously created in our laboratory [158].

Supplementary table 1. Main characteristic of cell lines.

Name	Sequence	Purpose
sgRNA	CGAGTCGGTGCTTTTTT	To allow the Cas9-mediated targeting of SPRTN C-terminus (Exon 5)
Donor DNA	GGTCTCCCCTGCTGTTAGTAACAGTCACCAAAATGTTCTAA GCAACTACTTTCCTAGAGTATCATTTGCCAACCAAAAGGCTT TCAGAGGTGTGAATGGATCTCCAAGGATAAGTGTAACAGTT GGCAACATCCCTAAAAACTCAGTCTCTTCTAGTTCTCAGAG AAGGGTTTCATCTTCTAAGATATCCCTAAGAAATTCTTCAAA AGTAACGGAATCAGCATCTGTGATGCCATCCCAGGATGTGA GTGGGTCTGAAGATACATTCCCAAATAAACGACCTAGGCTA GAAGATAAGACTGTTTTTGACAATTTTTTTATCAAGAAAGAG CAAATAAAAAGCAGTGGTAATGATCCAAAGTATAGTACAAC CACAGCTCAGAATTCCAGCAGTTCATCCAGTCAGAGCAAAA TGGTTAATTGCCCAGTTTGTCAGAATGAAGTTCTGGAGTCT CAGATTAATGAGCACTTGGACTGGTGCCTTGAAGGTGACAG CATCAAAGTCAAAAGCGAAGAAAGTCTTGGTAAGCCTATCC CTAACCCTCTCCTCGGTCTCGATTCTACGGAGAATTTGTATT TTCAGGGTAGCGCCTGGAGCCACCCCCAGTTCGAGAAGGG CGGCGGCAGCGGCGGCGGCAGCGGCGGCAGCGCCTGGA GCCACCCCCAGTTCGAGAAGGGAAGCGTGGAACTGAACCT GCGCGAGACAGAGCTGTGCCTGGGCCTGCCAGGCGGCGA TACAGTGGCCCCCGTGACCGGCAACAGAAGAGGCTTCAGC GAGACAGTGGACCTGAAGCTGAACCTGAACAACGAGCCCG CCAACAAAGAGGGCAGCACCACCCACGACGTGGTGACATT CGACAGCAAAGAGAAGTCCGCCTGCCCCAAGGACCCCGCC AAGCCACCAGCCAAGGCCCAGGTGGTCGGATGGCCTCCC GTGCGCAGCTACAGAAAGAACGTGATGGTGTCCTGCCAGA AGTCCAGCGGCGGACCAGAGGCCGCTGCCTTCGTGAAGGT GTCCATGGACGGCGCTCCCTACCTGAGAAAGATCGACCTG AGGATGTACAAGAGCTACGACGAGCTGAGCAACGCCCTGA GCAACATGTTCAGCAGCTTCACCATGGGCAAGCACGGCGG CGAGGAAGGCATGATCGACTTCATGAACGAGCGCAAGCTG ATGGACCTGGTGAACTCCTGGGACTACGTGCCCAGCTACG AGGACAAGGACGGCGACTGGATGCTCGTGGGAGATGTGCC CTGGCCCATGTTCGTGGATACCTGCAAGAGACTGAGACTG ATGAAGGGCAGCGACGCCATCGGCCTGGCCCCCAGAGCC ATGGAAAAGTGCAAGAGCAGAGCGACTAGTTGAAAAAGGTT TCAAAGTCTCAAGTACCATTTGTATTATCTCACTAAAGTCAG GAAGTTCTGGTTAATACTAAGATTTGTAGGTTATAATCTAGT TCACATAACCAATAGAAAGTGTCCTATTTTATATATACGCAT ATAAGATTGTAATTTTAAGATGTTTTGTGTCTCAGGGAACTT GTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCAT CACAAATTTCACAAATAAAGCATTTTTTTCACTGCATTCTAGT TGTGGTTTGTCCAAACTCATCAATGTATCTTATCATGTCTGA ATGTGCTATGTCAGATAACTTCGTATAATGTATGCTATACGA AGTTATGGGCAGAGCGCACATCGCCCACAGTCCCCGAGAA GTTGGGGGGAGGGGTCGGCAATTGAACCGGTGCCTAGAG AAGGTGGCGCGGGGTAAACTGGGAAAGTGATGTCGTGTAC TGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCGTATA TAAGTGCAGTAGTCGCCGTGAACGTTCTTTTTCGCAACGGG TTTGCCGCCAGAACACAGCCACCATGACCGAGTACAAGCC CACGGTGCGCCTCGCCACCCGCGACGACGTCCCCAGGGC CGTACGCACCCTCGCCGCCGCGTTCGCCGACTACCCCGCC ACGCGCCACACCGTCGATCCGGACCGCCACATCGAGCGG GTCACCGAGCTGCAAGAACTCTTCCTCACGCGCGTCGGGC	To provide a DNA repair template containing the desired knock-in, as showed in Figure 20

	TCGACATCGGCAAGGTGTGGGTCGCGGACGACGGCGCCG CGGTGGCGGTCTGGACCACGCCGGAGAGCGTCGAAGCGG GGGCGGTGTTCGCCGAGATCGGCCCGCGCATGGCCGAGT TGAGCGGTTCCCGGCTGGCCGCGCAGCAACAGATGGAAG GCCTCCTGGCGCCGCACCGGCCCAAGGAGCCCGCGTGGT TCCTGGCCACCGTCGGCGTCTCGCCCGACCACCAGGGCAA GGGTCTGGGCAGCGCCGTCGTGCTCCCCGGAGTGGAGGC GGCCGAGCGCGCCGGGGTGCCCGCCTTCCTGGAGACCTC CGCGCCCCGCAACCTCCCCTTCTACGAGCGGCTCGGCTTC ACCGTCACCGCCGACGTCGAGGTGCCCGAAGGACCGCGC ACCTGGTGCATGACCCGCAAGCCCGGTGCCTGACTGTGCC TTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGC CTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCC TAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTG TCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAG GGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATG CGGTGGGCTCTATGGATAACTTCGTATAATGTATGCTATAC GAAGTTATTGCTACATTCACTCTTGTATTAGGTATACTGTAA CCCATATTCTGCCTGTCGTGTATAAGTTTTAGATACTTTTGT TCTTTCTTGCTCTTAAGGATTTTAAAAACCTGTTAATCTTTTT ATTTGTATACTTTCCTAAAAATATTCATATGGGGAATCCTGT CAGGTGTTTGGTTATATTGACTATTTATTAATAGTATTAGAAC TCATTCCCTGAACTGATGTAAATCTTCATAGTGTCAGACATA CTGACCAAAACCACAATCTAGACTACAAAGTATATTGTTTTA GAGTACTCAAATTGTATTATTTATTAATTTTTTTGTTTGCAAA ATCTTAACAGGAACTGTATTTTCTATATTTTAAAGAATTTTAT TTGTCCCACTTTTACTAAACAGTGGCAGCAGATTTTAAGTTA AAGAATATGGAATATAGTAAAATAAGTAAATTTCTTTTGGAAT ATTTTTAGTAACAAATAGCCACTATAATTCTGTAGGCCAAAT TTTATATTGAGTTTAGCTGTTTTCTCAAAATTTAGCAGAGTG G	

Supplementary table 2. Genome-editing plasmids used for HCT116 SPRTN-AID. The sequence corresponding to SPRTN Exon 5 (C-terminus) is indicated in green. The V5 tag is indicated in red. The TEV tag is indicated in yellow. The Strep II tag is indicated in grey. The AID tag is indicated in light blue.

List of DEGs in HCT116 DMSO vs HCT116 Auxin		
Gene	**log2** (HCT116 Auxin / HCT116 DMSO)	**Q- value** (HCT116 Auxin / HCT116 DMSO)
'ABCC3'	0.378644965	0.040201698
'ACSL5'	0.355584479	3.43E-09
'ACTB'	-0.203160064	0.010906784
'ACTG1'	-0.338559201	7.13E-08
'ACTN4'	-0.170863332	0.021002954
'ADAM19'	0.232949208	0.00150978
'ADGRF1'	0.840094407	1.01E-37
'ADORA2B'	0.371336862	2.80E-04
'AGRN'	-0.27711446	0.001803423
'AHRR'	1.424863371	6.85E-38
'AKAP12'	0.341628214	3.06E-06
'ALDH1A3'	1.209544458	2.02E-134
'ALDOA'	-0.180719529	0.04269235
'AMOTL2'	-0.309362516	0.009649624
'ANAPC11'	-0.34646731	0.017560469
'ANKRD12'	0.336246297	0.035776428
'ANXA1'	0.204464863	0.005350686
'ARHGAP12'	0.219583572	0.00389655
'ARHGAP26'	0.327775462	6.16E-05
'ARHGAP29'	0.217717213	0.001860421
'ARHGDIA'	-0.169417112	0.018929182
'ATF6B'	0.271437484	0.022653592
'ATL3'	0.208831527	0.042285329
'ATP5MG'	-0.29052561	3.55E-04
'ATP5PF'	-0.276843445	0.017952335
'ATP9A'	0.280915286	4.57E-04
'BCL2L11'	0.302075438	0.01544233
'BCL2L2'	0.293996657	0.036507613
'BMP4'	0.244440265	6.57E-05
'BSDC1'	0.28749007	0.008549224
'C15orf62'	1.393115473	0.037562495
'C1orf116'	0.458017677	0.002311367
'CA2'	-0.259811707	0.027569745
'CANX'	0.148667741	0.030080773
'CAVIN1'	-0.18293327	0.033923585
'CBWD6'	-0.445725595	0.043512638
'CCDC167'	-0.43510117	0.043799441
'CCN1'	-0.431959852	0.002283722
'CCT7'	-0.188645587	0.04269235
'CCT8'	-0.171465866	0.046774477
'CD109'	0.340355811	1.81E-05
'CD59'	0.20901797	0.035344569

'CDC25B'	0.173221347	0.017630192
'CDC37'	-0.227131241	0.030080773
'CDK6'	0.462390931	4.24E-11
'CDKN2A'	-0.406118895	0.010605781
'CEMIP'	0.261292339	3.02E-05
'CFL1'	-0.221991642	8.63E-04
'CHCHD10'	-0.443285323	5.51E-05
'CHTF8'	0.331246729	4.22E-04
'CKB'	-0.36877216	1.88E-05
'CLDN3'	-0.497183673	0.017531596
'CLDND1'	0.249912481	8.64E-05
'CMTM7'	0.236278462	0.036566147
'CNN2'	-0.175710622	0.029552021
'CORO1B'	-0.229704517	0.034828847
'COX4I1'	-0.217589588	0.005350686
'COX7B'	-0.350946096	0.016457693
'CPOX'	0.267182052	6.22E-06
'CRABP2'	-0.247784507	0.033923585
'CYP1A1'	0.914928516	0.0170163
'CYP1B1'	0.910076384	3.08E-04
'DBN1'	-0.208963461	0.008247823
'DCBLD2'	0.188092299	4.78E-04
'DCLK1'	-1.124412872	0.011549363
'DDB2'	-0.298775599	0.016075538
'DGUOK'	-0.295204558	0.016007593
'DHRS3'	0.724993693	1.76E-19
'DPY19L1'	0.193655343	0.017332702
'DRAP1'	-0.335636513	0.002324329
'DRAXIN'	-0.483611602	0.048842004
'DSC2'	0.237946878	0.017531596
'DUSP5'	-0.196532822	0.040828717
'DUSP6'	-0.25693213	1.72E-04
'EEF1D'	-0.246557535	0.00128299
'EFNB2'	0.38696882	3.18E-05
'EGFR'	0.251657415	0.012098551
'ELF3'	0.285859436	0.007787077
'ENC1'	-0.362760906	2.39E-07
'EPHA2'	-0.363777151	6.11E-10
'EREG'	0.463652823	3.03E-18
'ERRFI1'	0.257546769	3.48E-06
'ETNK1'	0.234148193	0.024440522
'FADS1'	0.20966205	0.035504533
'FGF9'	0.785495881	0.024440522
'FGFBP1'	0.644564951	0.002201684

'FLNA'	-0.22004198	1.28E-04
'FLNC'	-0.176489353	0.036507613
'FOXD4L3'	5.586370085	0.01035178
'FOXO3'	0.253059768	0.002205381
'FRRS1'	0.356260356	3.68E-04
'GALNT5'	0.665814394	2.33E-16
'GCHFR'	1.568093115	6.91E-04
'GCLC'	0.205489761	0.046774477
'GDA'	0.468710901	1.09E-12
'GDI1'	-0.188711067	0.033181193
'GOLGA8A'	0.435663695	0.028395742
'GPAT3'	0.586123318	3.86E-12
'GSTO2'	0.249475638	0.023919883
'H1-0'	0.173403182	0.042359522
'HAS3'	1.094318672	1.24E-89
'HIPK2'	0.372779743	1.63E-04
'HMGA2'	0.269598446	1.14E-04
'HMGB1'	-0.171954791	0.024588916
'HNRNPAB'	-0.288391213	3.99E-04
'HS3ST1'	0.310499178	0.004792998
'HSPB1'	-0.242898444	0.003630302
'HYPK'	-0.429314619	0.016457693
'ID2'	-0.252601454	0.043438392
'IDH3B'	-0.263559495	0.021002954
'ISY1-RAB43'	-0.347187634	0.002427221
'JMJD7-PLA2G4B'	-7.934213942	0.032812114
'KIAA0754'	-0.295115498	0.040273634
'KSR2'	0.42191944	0.004562292
'LHFPL2'	0.3241445	0.002311367
'LIF'	0.325419652	0.030520752
'LPCAT3'	0.482652919	1.35E-11
'LRRC58'	0.217748584	0.026649587
'LSM7'	-0.391366081	0.034828847
'MAGI1'	-0.243198856	0.045527417
'MAGOH'	-0.347093074	0.007484486
'MARCHF8'	0.297609182	0.009996625
'MAT2A'	0.177528934	0.033923585
'MDH2'	-0.213475861	0.029070869
'MEAK7'	0.248960553	0.035344569
'MESD'	0.209612843	0.041536765
'MID1'	0.279632228	3.91E-04
'MIDN'	-0.232685119	0.007619959
'MIF'	-0.415818246	6.71E-07
'MLF2'	-0.230356895	0.035047406

'MRPS34'	-0.307084143	0.023426447
'MTAP'	0.180998327	0.009649624
'MTUS1'	0.669794354	6.78E-28
'MYL12A'	-0.268053095	5.07E-05
'MYL6'	-0.24142711	0.025061433
'MYO19'	0.223933968	0.007798736
'MYO1D'	0.215293794	0.018952294
'MYOF'	0.236783125	0.009158321
'MYZAP'	0.501385542	0.04189064
'NAB2'	-0.28752969	5.51E-04
'NDUFB1'	-0.573327558	0.004792998
'NDUFB6'	-0.421584043	0.040511699
'NFE2L2'	0.217732601	0.025833041
'NHSL1'	0.318222032	0.011319012
'NME1-NME2'	-0.488265323	6.88E-06
'NOC2L'	-0.196280982	0.025224005
'NPAS2'	0.342547691	0.033923585
'NQO1'	0.208930572	0.023030685
'NR2F1'	-0.480199193	1.79E-09
'NR2F2'	0.352344507	6.73E-08
'NRIP1'	0.38982133	4.22E-04
'NT5E'	0.230249902	0.011898747
'OSBPL8'	0.185596609	0.0347242
'PARP1'	-0.179889129	0.04353367
'PERP'	0.18951542	0.023030685
'PHLDA2'	-0.284366138	0.015775326
'PIR'	0.382787317	0.001895601
'PLEKHH3'	-0.308369423	0.019936655
'PLXNB2'	-0.207988956	0.028483741
'POLR2C'	-0.242550408	0.004966909
'PPFIBP1'	-0.248508302	0.001142921
'PPP1R9B'	-0.191939271	0.039554216
'PPTC7'	0.201194606	0.046438357
'PRDX5'	-0.26534882	2.48E-04
'PSD4'	0.342765888	0.015811332
'PTMS'	-0.277466894	0.015117409
'REEP3'	0.262307186	0.023426447
'RETSAT'	0.205591118	0.04353367
'RGPD5'	0.703022495	9.92E-06
'RHOBTB3'	-0.243974567	0.006187305
'RHOC'	-0.2195084	0.044533996
'ROCK2'	0.223365274	0.040511699
'RPL11'	-0.192403217	0.024625037
'RPL17'	-0.174788892	0.049249867

'RPL24'	-0.215513768	0.006376019
'RPL27'	-0.349248485	5.88E-07
'RPL35A'	-0.223350474	0.02022535
'RPL36AL'	-0.332469579	0.017382386
'RPL41'	-0.243691153	0.003522554
'RPL5'	-0.211811917	1.80E-04
'RPS16'	-0.21269034	0.006549517
'RPS20'	-0.334897094	8.69E-08
'RPS26'	-0.260357342	1.49E-04
'RPS27A'	-0.229270043	1.48E-04
'RPS28'	-0.465558199	0.032206665
'RPS4X'	-0.30029315	1.47E-07
'RPS5'	-0.262843031	0.034828847
'RPS7'	-0.259404401	0.013500055
'RPS8'	-0.16609293	0.022762042
'S100A10'	-0.323168065	0.001618127
'SCD'	0.207444817	0.002427014
'SCRIB'	-0.257427207	0.016007593
'SECTM1'	0.747611717	1.48E-04
'SELENOW'	-0.26191895	0.006549517
'SERINC5'	0.292910528	5.46E-04
'SH2B3'	-0.218293535	0.011509174
'SH3PXD2A'	0.368762029	0.009649624
'SHMT2'	-0.220584536	0.011045752
'SIVA1'	-0.312895647	0.045243158
'SKAP2'	0.239946188	0.030080773
'SLC16A6'	0.608174743	0.003094
'SLC2A1'	0.210494651	0.009649624
'SLC37A2'	0.517895874	0.015568107
'SLC38A2'	0.213499314	3.99E-04
'SLC6A17'	0.292741351	0.001860421
'SLC7A11'	0.570041001	1.73E-08
'SLC7A5'	0.167683131	0.013166767
'SMAD3'	0.469704734	6.62E-16
'SNRNP70'	-0.354070385	1.57E-09
'SNRPD2'	-0.291438932	0.009649624
'SNRPF'	-0.436860686	3.06E-06
'SNRPG'	-0.314094716	0.01684625
'SOCS6'	0.251220926	0.014855631
'SOX4'	-0.398848381	9.92E-13
'SPPL2A'	0.26382329	0.012122242
'SSR3'	0.213449283	0.011881988
'STIMATE-MUSTN1'	-6.270802481	1.72E-04
'SUMO2'	-0.172159042	0.025228277

'TACC3'	-0.219112327	0.041238186
'TBC1D3'	-2.437495782	0.041238186
'TBC1D3L'	0.541817829	0.009649624
'TGFA'	0.201595829	0.034617387
'TINAGL1'	0.273566302	0.009407141
'TIPARP'	0.922500765	9.66E-51
'TMC5'	0.385670041	0.003775789
'TMEM123'	0.245792948	1.72E-04
'TMEM200A'	0.363491868	5.33E-04
'TMSB10'	-0.351933596	7.59E-04
'TNFRSF12A'	-0.253280278	0.020669858
'TNFSF12-TNFSF13'	-1.908702105	2.03E-04
'TPM2'	-0.299583187	0.004792998
'TRAM2'	0.232821009	0.017560469
'TRIM2'	0.435886465	2.45E-10
'TRIM47'	-0.230828985	0.04353367
'TSC22D1'	0.280689838	6.00E-06
'TTC19'	0.275037676	0.013500055
'TUBA1B'	-0.173343341	0.013208822
'TUBB3'	-0.26881698	0.008625825
'TUBB4A'	-0.348223721	7.96E-04
'TUBB6'	-0.241207002	0.043636718
'TWSG1'	0.21316063	0.030549387
'UBE2S'	-0.439049813	2.00E-06
'UNC13A'	0.333367466	3.80E-04
'UQCR11'	-0.34029863	0.015811332
'UQCRQ'	-0.210267517	0.040273634
'VARS1'	-0.164943355	0.035344569
'WNT16'	0.311190446	0.043090123
'YAP1'	0.219492321	0.005229016
'YBX1'	-0.248785111	2.49E-04
'ZBED2'	0.6092095	1.14E-07
'ZDHHC3'	0.214398203	0.039733434
'ZNF177'	-1.786625362	0.007634748
'ZNF827'	0.334039684	8.32E-04

Supplementary table 3. List of DEGs in HCT116 DMSO vs HCT116 Auxin.

List of DEGs in HeLa WT vs HeLa SPRTN -/-/+ (Top 100)		
Gene	**log2** (HeLa SPRTN -/-/+ / HeLa WT)	**Q- value** (HeLa SPRTN -/-/+ / HeLa WT)
'A2M'	-5.3659	4.70E-94
'AHRR'	-1.42763	6.38E-39
'AK4'	-1.34164	2.32E-57
'ALDOC'	-1.42953	1.54E-49
'ALPP'	-2.92064	1.67E-54
'ASS1'	-0.85731	1.18E-65
'ATXN10'	0.915392	3.93E-51
'BCHE'	1.876435	2.38E-69
'BMP2'	0.748711	3.57E-40
'CCN3'	2.26462	3.70E-97
'CD99'	0.857477	2.12E-41
'CDH10'	1.039809	1.59E-53
'CGA'	5.330957	0
'CNTN1'	4.996023	0
'COL5A2'	6.055922	0
'DLX3'	2.814607	1.75E-106
'DSEL'	1.101735	3.57E-42
'EDIL3'	-1.68372	1.84E-81
'EEF2'	-0.8216	2.03E-51
'EMP1'	1.134748	3.89E-43
'F11R'	-0.87007	5.32E-42
'FDPS'	-1.01974	9.33E-54
'FN1'	3.938452	2.76E-58
'GAGE12J'	1.402723	6.09E-45
'GAREM1'	3.894506	1.96E-62
'GDA'	0.910544	1.21E-47
'GLP2R'	1.086853	2.40E-56
'GLUL'	-1.05072	4.43E-41
'HEG1'	1.192192	3.65E-74
'HK1'	1.821396	4.30E-85
'HLA-B'	-1.66693	4.99E-89
'HPD'	-0.93298	5.60E-71
'HSPA2'	1.869801	7.55E-66
'HSPB7'	-1.66731	1.85E-119
'HTRA1'	0.857008	7.31E-42
'ID1'	1.281777	1.67E-101
'ID2'	2.789727	5.04E-99
'IFITM3'	-0.77515	3.57E-40
'IGFBP3'	2.018033	7.96E-124
'IGSF10'	2.972743	2.92E-61
'IL32'	-3.7515	1.34E-38
'INSIG1'	-1.22216	3.46E-69

'JAK1'	-0.78191	2.39E-39
'KRT7'	0.972287	9.28E-62
'KRT80'	2.672094	0
'LAMB3'	0.812944	1.69E-38
'LRRC15'	4.906124	1.72E-57
'MFAP5'	4.726998	0
'MFSD14A'	-0.89218	2.34E-38
'MME'	1.425865	9.98E-90
'MUC13'	2.924793	2.37E-120
'NEGR1'	-3.0855	9.17E-164
'NEURL1B'	1.436194	1.88E-36
'NR4A1'	-0.98242	5.80E-37
'NRP2'	-2.32158	2.83E-52
'NSMF'	-0.95231	1.09E-47
'NUP210'	-0.75591	3.17E-50
'OLFML1'	2.40292	7.13E-38
'OSBP2'	0.964174	1.41E-42
'OTULINL'	1.129621	8.97E-53
'P2RY6'	-1.32575	4.53E-52
'PCDH17'	1.074035	1.24E-49
'PDE2A'	1.635681	2.37E-120
'PGBD5'	3.544284	3.79E-105
'PGM1'	-0.89433	1.52E-61
'PIK3CD'	-3.79525	8.17E-94
'PLIN4'	-3.46888	1.15E-154
'PLXNA2'	1.617535	2.18E-58
'PPL'	1.317804	2.03E-57
'QPRT'	-6.00919	1.59E-77
'RBM3'	1.243446	2.44E-39
'RCN3'	-1.69702	1.17E-38
'RIMKLB'	1.551422	2.56E-63
'RNF150'	-5.70169	2.57E-51
'RYR2'	2.277765	3.72E-42
'SCARA5'	3.476783	1.16E-55
'SELENOF'	-0.92778	3.41E-58
'SERBP1'	-0.72943	2.72E-40
'SERPINB5'	1.181206	1.96E-77
'SESN3'	1.211199	2.68E-78
'SH3BGRL'	2.238103	7.31E-74
'SHISAL1'	1.364263	6.68E-43
'SLC12A3'	1.986277	3.94E-145
'SLC1A3'	-1.01122	2.39E-45
'SLC27A2'	0.884925	3.36E-39
'SLC9A3'	-1.05911	1.30E-40

'SMAD6'	1.368127	7.08E-64
'SPINK5'	4.27264	1.21E-39
'SPINK6'	6.50782	8.54E-129
'SPOCK1'	1.726262	4.99E-117
'SYTL4'	4.486134	1.33E-45
'TLR4'	3.326715	2.72E-148
'TNS1'	-1.16114	1.44E-42
'TSPAN8'	1.30827	2.49E-53
'TTYH3'	-1.26017	3.10E-50
'TUSC3'	2.579756	1.29E-47
'USP33'	-0.81562	6.67E-43
'VWA5A'	-2.42581	6.45E-62
'WFDC1'	-2.04417	1.01E-83
'WLS'	-0.75842	2.07E-41

Supplementary table 4. List of DEGs in HeLa WT vs HeLa SPRTN $^{-/-/+}$ (Top 100 over 4575)

List of DEGs present in both HCT116 Auxin and HeLa SPRTN -/-/+				
Gene	**log2** (HCT116 Auxin / HCT116 DMSO)	**Q- value** (HCT116 Auxin / HCT116 DMSO)	**log2** (HeLa SPRTN -/-/+ / HeLa WT)	**Q- value** (HeLa SPRTN -/-/+ / HeLa WT)
'ABCC3'	0.378645	0.040201698	0.85917	9.40E-33
'ACSL5'	0.355584	3.43E-09	-1.98106	3.58E-31
'ACTN4'	-0.17086	0.021002954	0.148852	0.022542
'AHRR'	1.424863	6.85E-38	-1.42763	6.38E-39
'ALDH1A3'	1.209544	2.02E-134	0.614456	1.19E-13
'ALDOA'	-0.18072	0.04269235	-0.43332	4.40E-13
'AMOTL2'	-0.30936	0.009649624	0.612919	8.86E-16
'ANXA1'	0.204465	0.005350686	0.591296	6.79E-15
'ARHGAP26'	0.327775	6.16E-05	0.383142	8.22E-04
'ARHGAP29'	0.217717	0.001860421	-0.51533	9.04E-07
'ARHGDIA'	-0.16942	0.018929182	0.162408	0.030475
'ATL3'	0.208832	0.042285329	0.259586	3.76E-04
'ATP5PF'	-0.27684	0.017952335	-0.22455	0.016805
'ATP9A'	0.280915	4.57E-04	0.186832	0.019398
'C1orf116'	0.458018	0.002311367	1.404571	0.047926
'CAVIN1'	-0.18293	0.033923585	0.39189	1.15E-11
'CCDC167'	-0.4351	0.043799441	0.519169	0.002138
'CCT8'	-0.17147	0.046774477	-0.20531	0.005127
'CD109'	0.340356	1.81E-05	0.321732	8.69E-08
'CD59'	0.209018	0.035344569	0.33271	4.94E-09
'CDC25B'	0.173221	0.017630192	0.231839	0.001547
'CDK6'	0.462391	4.24E-11	0.44619	4.12E-10
'CDKN2A'	-0.40612	0.010605781	-0.39438	1.02E-05
'CHCHD10'	-0.44329	5.51E-05	-0.44392	1.02E-05
'CKB'	-0.36877	1.88E-05	-0.61495	1.52E-15
'CORO1B'	-0.2297	0.034828847	-0.28608	0.00415
'COX4I1'	-0.21759	0.005350686	-0.16148	0.033289
'CRABP2'	-0.24778	0.033923585	-0.70254	6.23E-07
'CYP1B1'	0.910076	3.08E-04	-0.40629	5.96E-04
'DGUOK'	-0.2952	0.016007593	-0.26541	0.021386
'DPY19L1'	0.193655	0.017332702	0.581811	7.57E-15
'DRAP1'	-0.33564	0.002324329	-0.36597	0.002061
'DSC2'	0.237947	0.017531596	1.077595	6.56E-35
'DUSP6'	-0.25693	1.72E-04	0.515017	0.004916
'EGFR'	0.251657	0.012098551	0.395388	7.72E-11
'ELF3'	0.285859	0.007787077	-0.39499	9.78E-05
'ENC1'	-0.36276	2.39E-07	2.587246	0.049642
'FADS1'	0.209662	0.035504533	0.491023	1.86E-19
'FGFBP1'	0.644565	0.002201684	0.799181	0.001962
'FLNA'	-0.22004	1.28E-04	-0.12881	0.030716
'FOXO3'	0.25306	0.002205381	-0.17243	0.043726

'GDA'	0.468711	1.09E-12	0.910544	1.21E-47
'H1-0'	0.173403	0.042359522	0.376228	8.99E-06
'HIPK2'	0.37278	1.63E-04	3.430527	5.01E-20
'HSPB1'	-0.2429	0.003630302	-0.57562	5.38E-15
'ID2'	-0.2526	0.043438392	2.789727	5.04E-99
'IDH3B'	-0.26356	0.021002954	-0.41966	4.57E-05
'LRRC58'	0.217749	0.026649587	0.205271	0.016651
'LSM7'	-0.39137	0.034828847	-0.57355	6.44E-04
'MAGI1'	-0.2432	0.045527417	-0.47619	2.83E-06
'MAT2A'	0.177529	0.033923585	0.241209	0.003514
'MIDN'	-0.23269	0.007619959	-0.39157	1.74E-05
'MIF'	-0.41582	6.71E-07	-0.29641	4.85E-04
'MRPS34'	-0.30708	0.023426447	-0.29119	1.24E-04
'MTUS1'	0.669794	6.78E-28	-0.32369	1.77E-04
'MYL12A'	-0.26805	5.07E-05	0.283836	1.06E-04
'MYO19'	0.223934	0.007798736	-0.14738	0.023523
'MYO1D'	0.215294	0.018952294	4.706983	0.014117
'MYZAP'	0.501386	0.04189064	1.81759	1.24E-09
'NDUFB1'	-0.57333	0.004792998	0.351471	0.002242
'NFE2L2'	0.217733	0.025833041	0.199281	0.011646
'NME1-NME2'	-0.48827	6.88E-06	-0.60879	4.94E-07
'NOC2L'	-0.19628	0.025224005	-0.32406	1.09E-06
'NPAS2'	0.342548	0.033923585	0.353838	0.0135
'NRIP1'	0.389821	4.22E-04	0.218706	0.017007
'NT5E'	0.23025	0.011898747	0.349235	0.036168
'OSBPL8'	0.185597	0.0347242	-0.22336	0.044684
'PERP'	0.189515	0.023030685	0.330347	2.43E-06
'PHLDA2'	-0.28437	0.015775326	-0.58483	2.94E-10
'PIR'	0.382787	0.001895601	0.341702	1.84E-04
'PLEKHH3'	-0.30837	0.019936655	-0.24946	0.009692
'PPTC7'	0.201195	0.046438357	-0.19765	0.018954
'PSD4'	0.342766	0.015811332	0.698497	0.014357
'REEP3'	0.262307	0.023426447	0.209002	0.01106
'RHOBTB3'	-0.24397	0.006187305	0.633988	2.01E-23
'RHOC'	-0.21951	0.044533996	-0.44326	1.14E-06
'ROCK2'	0.223365	0.040511699	0.519594	3.98E-12
'RPL11'	-0.1924	0.024625037	-0.18035	0.005709
'RPL17'	-0.17479	0.049249867	-0.21032	0.007578
'RPL24'	-0.21551	0.006376019	-0.25708	3.57E-05
'RPL27'	-0.34925	5.88E-07	-0.22011	0.039078
'RPL35A'	-0.22335	0.02022535	-0.20792	0.001855
'RPL5'	-0.21181	1.80E-04	-0.47909	9.73E-21
'RPS16'	-0.21269	0.006549517	-0.20275	0.003444
'RPS26'	-0.26036	1.49E-04	-0.15599	0.040512

'RPS27A'	-0.22927	1.48E-04	-0.15935	0.01471
'RPS28'	-0.46556	0.032206665	-0.40323	1.07E-06
'RPS4X'	-0.30029	1.47E-07	-0.22687	0.001443
'RPS5'	-0.26284	0.034828847	-0.38913	9.60E-07
'RPS7'	-0.2594	0.013500055	-0.2744	0.002548
'S100A10'	-0.32317	0.001618127	0.553038	2.61E-21
'SCD'	0.207445	0.002427014	-0.44205	2.00E-16
'SCRIB'	-0.25743	0.016007593	-0.2012	0.019455
'SECTM1'	0.747612	1.48E-04	-0.59215	1.98E-05
'SH2B3'	-0.21829	0.011509174	-0.70448	1.28E-07
'SHMT2'	-0.22058	0.011045752	-0.4617	5.77E-15
'SLC16A6'	0.608175	0.003094	-0.51754	3.62E-08
'SLC2A1'	0.210495	0.009649624	-0.48666	1.82E-12
'SLC6A17'	0.292741	0.001860421	-4.08117	8.85E-11
'SLC7A11'	0.570041	1.73E-08	0.451141	0.002018
'SLC7A5'	0.167683	0.013166767	-0.26142	2.99E-05
'SMAD3'	0.469705	6.62E-16	-0.34556	7.09E-05
'SNRNP70'	-0.35407	1.57E-09	-0.17333	0.009314
'SOCS6'	0.251221	0.014855631	0.352749	0.001006
'SPPL2A'	0.263823	0.012122242	0.208341	0.011048
'SSR3'	0.213449	0.011881988	-0.19534	0.008031
'TGFA'	0.201596	0.034617387	1.082988	0.007358
'TINAGL1'	0.273566	0.009407141	-0.6275	0.00165
'TNFRSF12A'	-0.25328	0.020669858	-0.55259	4.03E-10
'TPM2'	-0.29958	0.004792998	-0.32067	5.50E-05
'TRAM2'	0.232821	0.017560469	0.436197	1.08E-06
'TRIM2'	0.435886	2.45E-10	0.537763	1.36E-06
'TRIM47'	-0.23083	0.04353367	-0.30242	6.06E-04
'TSC22D1'	0.28069	6.00E-06	-0.16154	0.007017
'TWSG1'	0.213161	0.030549387	0.394516	4.86E-04
'VARS1'	-0.16494	0.035344569	-0.21698	4.53E-04
'YAP1'	0.219492	0.005229016	0.205835	0.005303
'ZDHHC3'	0.214398	0.039733434	-0.18658	0.019437

Supplementary table 5. List of DEGs present in both HCT116 Auxin and Hela SPRTN $^{-/-/+}$

List of DEGs present in in HCT116 Auxin, but not in HeLa SPRTN -/-/+		
Gene	**log2** (HCT116 Auxin / HCT116 DMSO)	**Q- value** (HCT116 Auxin / HCT116 DMSO)
'ACTB'	-0.20316	0.010907
'ACTG1'	-0.33856	7.13E-08
'ADAM19'	0.232949	0.00151
'ADGRF1'	0.840094	1.01E-37
'ADORA2B'	0.371337	2.80E-04
'AGRN'	-0.27711	0.001803
'AKAP12'	0.341628	3.06E-06
'ANAPC11'	-0.34647	0.01756
'ANKRD12'	0.336246	0.035776
'ARHGAP12'	0.219584	0.003897
'ATF6B'	0.271437	0.022654
'ATP5MG'	-0.29053	3.55E-04
'BCL2L11'	0.302075	0.015442
'BCL2L2'	0.293997	0.036508
'BMP4'	0.24444	6.57E-05
'BSDC1'	0.28749	0.008549
'C15orf62'	1.393115	0.037562
'CA2'	-0.25981	0.02757
'CANX'	0.148668	0.030081
'CBWD6'	-0.44573	0.043513
'CCN1'	-0.43196	0.002284
'CCT7'	-0.18865	0.042692
'CDC37'	-0.22713	0.030081
'CEMIP'	0.261292	3.02E-05
'CFL1'	-0.22199	8.63E-04
'CHTF8'	0.331247	4.22E-04
'CLDN3'	-0.49718	0.017532
'CLDND1'	0.249912	8.64E-05
'CMTM7'	0.236278	0.036566
'CNN2'	-0.17571	0.029552
'COX7B'	-0.35095	0.016458
'CPOX'	0.267182	6.22E-06
'CYP1A1'	0.914929	0.017016
'DBN1'	-0.20896	0.008248
'DCBLD2'	0.188092	4.78E-04
'DCLK1'	-1.12441	0.011549
'DDB2'	-0.29878	0.016076
'DHRS3'	0.724994	1.76E-19
'DRAXIN'	-0.48361	0.048842
'DUSP5'	-0.19653	0.040829
'EEF1D'	-0.24656	0.001283
'EFNB2'	0.386969	3.18E-05

'EPHA2'	-0.36378	6.11E-10
'EREG'	0.463653	3.03E-18
'ERRFI1'	0.257547	3.48E-06
'ETNK1'	0.234148	0.024441
'FGF9'	0.785496	0.024441
'FLNC'	-0.17649	0.036508
FOXD4L3'	5.58637	0.010352
'FRRS1'	0.35626	3.68E-04
'GALNT5'	0.665814	2.33E-16
'GCHFR'	1.568093	6.91E-04
'GCLC'	0.20549	0.046774
'GDI1'	-0.18871	0.033181
'GOLGA8A'	0.435664	0.028396
'GPAT3'	0.586123	3.86E-12
'GSTO2'	0.249476	0.02392
'HAS3'	1.094319	1.24E-89
'HMGA2'	0.269598	1.14E-04
'HMGB1'	-0.17195	0.024589
'HNRNPAB'	-0.28839	3.99E-04
'HS3ST1'	0.310499	0.004793
'HYPK'	-0.42931	0.016458
'ISY1-RAB43'	-0.34719	0.002427
'JMJD7-PLA2G4B'	-7.93421	0.032812
'KIAA0754'	-0.29512	0.040274
'KSR2'	0.421919	0.004562
'LHFPL2'	0.324145	0.002311
'LIF'	0.32542	0.030521
'LPCAT3'	0.482653	1.35E-11
'MAGOH'	-0.34709	0.007484
'MARCHF8'	0.297609	0.009997
'MDH2'	-0.21348	0.029071
'MEAK7'	0.248961	0.035345
'MESD'	0.209613	0.041537
'MID1'	0.279632	3.91E-04
'MLF2'	-0.23036	0.035047
'MTAP'	0.180998	0.00965
'MYL6'	-0.24143	0.025061
'MYOF'	0.236783	0.009158
'NAB2'	-0.28753	5.51E-04
'NDUFB6'	-0.42158	0.040512
'NHSL1'	0.318222	0.011319
'NQO1'	0.208931	0.023031
'NR2F1'	-0.4802	1.79E-09
'NR2F2'	0.352345	6.73E-08

'PARP1'	-0.17989	0.043534
'PLXNB2'	-0.20799	0.028484
'POLR2C'	-0.24255	0.004967
'PPFIBP1'	-0.24851	0.001143
'PPP1R9B'	-0.19194	0.039554
'PRDX5'	-0.26535	2.48E-04
'PTMS'	-0.27747	0.015117
'RETSAT'	0.205591	0.043534
'RGPD5'	0.703022	9.92E-06
'RPL36AL'	-0.33247	0.017382
'RPL41'	-0.24369	0.003523
'RPS20'	-0.3349	8.69E-08
'RPS8'	-0.16609	0.022762
'SELENOW'	-0.26192	0.00655
'SERINC5'	0.292911	5.46E-04
'SH3PXD2A'	0.368762	0.00965
'SIVA1'	-0.3129	0.045243
'SKAP2'	0.239946	0.030081
'SLC37A2'	0.517896	0.015568
'SLC38A2'	0.213499	3.99E-04
'SNRPD2'	-0.29144	0.00965
'SNRPF'	-0.43686	3.06E-06
'SNRPG'	-0.31409	0.016846
'SOX4'	-0.39885	9.92E-13
'STIMATE-MUSTN1'	-6.2708	1.72E-04
'SUMO2'	-0.17216	0.025228
'TACC3'	-0.21911	0.041238
'TBC1D3'	-2.4375	0.041238
'TBC1D3L'	0.541818	0.00965
'TIPARP'	0.922501	9.66E-51
'TMC5'	0.38567	0.003776
'TMEM123'	0.245793	1.72E-04
'TMEM200A'	0.363492	5.33E-04
'TMSB10'	-0.35193	7.59E-04
'TNFSF12-TNFSF13'	-1.9087	2.03E-04
'TTC19'	0.275038	0.0135
'TUBA1B'	-0.17334	0.013209
'TUBB3'	-0.26882	0.008626
'TUBB4A'	-0.34822	7.96E-04
'TUBB6'	-0.24121	0.043637
'UBE2S'	-0.43905	2.00E-06
'UNC13A'	0.333367	3.80E-04
'UQCR11'	-0.3403	0.015811

'UQCRQ'	-0.21027	0.040274
'WNT16'	0.31119	0.04309
'YBX1'	-0.24879	2.49E-04
'ZBED2'	0.60921	1.14E-07
'ZNF177'	-1.78663	0.007635
'ZNF827'	0.33404	8.32E-04

Supplementary table 6. List of DEGS present in HCT116 Auxin but not in HeLa SPRTN $^{-/-/+}$

Patient ID	**Treatment**	**Clinical response** (12 weeks)	**Overall survival** (Status)	**Overall Survival** (Months)	**Progression Free Survival** (Status)	**Progression Free Survival** (Months)	**Percentage of SPRTN-positive cells in malignant epithelium** (%)
SC02337A	FOLFIRI	PR	1	4.44	1	3.22	95.54
SC02332A	FOLFIRI	PD	1	19.63	1	3.06	93.82
SC02339A	FOLFIRI	PD	1	12.69	1	2.86	93.23
SC02392A	FOLFIRI	PR	0	25.08	1	8.02	92.97
SC02342A	FOLFIRI	PD	1	4.90	1	3.85	92.95
SC02357A	FOLFIRI	PD	1	18.31	1	3.22	92.94
SC02328A	FOLFIRI	SD	1	12.95	1	6.51	92.34
SC02351A	FOLFIRI	SD	1	8.15	1	8.15	92.29
SC02336A	FOLFIRI	SD	1	10.92	1	8.55	90.44
SC02355A	FOLFIRI	PR	1	27.39	1	9.17	87.52
SC02350A	FOLFIRI	SD	1	4.27	1	4.27	83.12
SC02330A	FOLFIRI	SD	1	9.01	1	5.75	82.64
SC02348A	FOLFIRI	PD	1	17.56	1	3.78	81.92
SC02333A	FOLFIRI	SD	1	14.50	1	6.44	81.16
SC02399A	FOLFIRI	SD	1	8.45	1	5.56	80.43
SC02340A	FOLFIRI	PD	1	7.76	1	6.05	79.54
SC02346A	FOLFIRI	SD	1	14.83	1	10.29	78.93
SC02405A	FOLFIRI	PR	1	11.41	1	5.69	77.75
SC02353A	FOLFIRI	SD	1	30.38	1	10.59	76.26
SC02334A	FOLFIRI	PR	1	19.89	1	11.67	73.77
SC02390A	FOLFIRI	PR	1	21.17	1	11.15	72.74
SC02331A	FOLFIRI	SD	1	9.67	1	9.67	70.62
SC02397A	FOLFIRI	PR	1	16.90	1	5.06	69.59
SC02402A	FOLFIRI	PD	1	7.13	1	3.22	69.05
SC02398A	FOLFIRI	PR	1	13.84	1	8.12	67.76
SC02341A	FOLFIRI	SD	1	9.04	1	7.50	62.79
SC02395A	FOLFIRI	SD	1	13.97	1	9.07	62.59
SC02403A	FOLFIRI	PD	1	10.92	1	2.73	60.71
SC02396A	FOLFIRI	SD	1	8.55	1	8.55	59.82
SC02329A	FOLFIRI	SD	1	13.61	1	13.61	57.71
SC02394A	FOLFIRI	SD	1	15.81	1	9.07	57.39
SC02327A	FOLFIRI	PD	1	21.44	1	6.25	47.88
SC02389A	FOLFIRI	PR	0	48.36	1	42.35	43.95
SC02412A	FOLFIRI	SD	1	24.79	1	12.59	38.66
SC02349A	FOLFIRI	SD	0	3.62	0	3.62	38.41
SC02335A	FOLFIRI	PR	1	20.42	1	12.39	37.13
SC02407A	FOLFIRI	SD	1	13.28	1	8.75	22.49
SC02413A	FOLFIRI	SD	1	24.99	1	10.82	21.51
SC02393A	FOLFIRI	PR	1	15.91	1	9.63	15.30
SC02352A	FOLFIRI	SD	1	14.76	1	8.78	13.23

SC02400A	FOLFIRI	PR	1	12.99	1	8.52	11.02
SC02410A	FOLFIRI	PR	0	33.53	1	16.57	8.00
SC02411A	FOLFIRI	PR	1	11.84	1	9.86	7.62
SC02406A	FOLFIRI	PR	0	21.40	1	9.40	2.46
SC02401A	FOLFIRI	PR	1	27.02	1	6.02	2.06

Supplementary table 7. Clinical and IHC data of 45 CRC patients treated with FOLFIRI (Exploratory analysis).

Clinical and immunohistochemical data of the 45 colorectal cancer patients analysed in Figure 44. Clinical response at 12 weeks is indicated as complete response (CR), partial response (PR), stable disease (SD) or progressive disease (PD).

Patient ID	**Treatment**	**Clinical response** (12 weeks)	**Overall survival** (status)	**Overall survival** (months)	**Progression Free Survival** (Status)	**Progression Free Survival** (censored at 15 months)	**SPRTN Nuclear Optical Density in malignant epithelium** (mean -logged)	**SPRTN Tertile**
SC00275A	5-Fluorouracil	SD	1	9.80	1	7.82	-0.92	High
SC00492A	5-Fluorouracil	PD	0	23.01	1	2.53	-1.03	High
SC00529A	5-Fluorouracil	PR	1	15.98	1	7.59	-1.24	High
SC00331A	5-Fluorouracil	SD	1	15.32	1	11.44	-1.35	High
SC00356A	5-Fluorouracil	SD	1	9.34	1	5.29	-1.39	High
SC00370A	5-Fluorouracil	SD	1	7.82	1	4.34	-1.42	High
SC02362A	FOLFIRI	SD	1	14.47	1	10.85	-1.45	High
SC00222A	5-Fluorouracil	PR	1	14.73	1	10.13	-1.47	High
SC00433A	5-Fluorouracil	SD	1	7.96	1	7.96	-1.50	High
SC00379A	5-Fluorouracil	PD	1	3.85	1	3.32	-1.62	High
SC00176A	FOLFOX	PR	1	28.11	1	11.28	-1.65	High
SC00434A	5-Fluorouracil	PR	1	17.13	1	9.01	-1.65	High
SC00378A	5-Fluorouracil	SD	1	19.79	1	9.90	-1.66	High
SC00515A	5-Fluorouracil	PD	1	11.38	1	2.76	-1.68	High
SC02401A	FOLFIRI	PR	1	27.02	1	6.02	-1.68	High
SC00160A	5-Fluorouracil	PD	1	3.72	1	2.73	-1.70	High
SC00498A	5-Fluorouracil	SD	0	14.73	1	11.51	-1.73	High
SC00446A	5-Fluorouracil	SD	1	4.60	1	3.45	-1.74	High
SC00346A	5-Fluorouracil	PD	1	16.64	1	2.66	-1.76	High
SC00154A	5-Fluorouracil	SD	1	27.02	1	8.71	-1.80	High
SC00317A	5-Fluorouracil	CR	1	19.04	0	15.00	-1.85	High
SC00285A	FOLFOX	SD	1	9.21	1	8.71	-1.86	High
SC00350A	5-Fluorouracil	PD	1	16.57	1	2.43	-1.86	High
SC02388A	FOLFIRI	SD	1	19.89	1	9.17	-1.87	High
SC00279A	5-Fluorouracil	PR	1	11.08	1	5.52	-1.89	High
SC02402A	FOLFIRI	PD	1	7.13	1	3.22	-1.89	High
SC00240A	5-Fluorouracil	SD	1	16.08	1	6.08	-1.91	High
SC02351A	FOLFIRI	SD	1	8.15	1	8.15	-1.91	High
SC02400A	FOLFIRI	PR	1	12.99	1	8.52	-1.91	High
SC02405A	FOLFIRI	PR	1	11.41	1	5.69	-1.92	High
SC02413A	FOLFIRI	SD	1	24.99	1	10.82	-1.92	High
SC00328A	5-Fluorouracil	SD	1	11.15	1	6.05	-1.94	High
SC02383A	FOLFIRI	SD	1	27.81	1	9.40	-1.94	High
SC00494A	5-Fluorouracil	SD	1	23.41	1	14.10	-1.95	High
SC00277A	5-Fluorouracil	SD	1	14.50	1	14.50	-1.96	High
SC00136A	5-Fluorouracil	PR	1	29.88	1	8.45	-1.97	High
SC00428A	5-Fluorouracil	PD	1	6.38	1	2.83	-1.97	High
SC00139A	FOLFOX	PR	1	9.14	1	8.32	-1.97	High
SC00470A	5-Fluorouracil	PR	1	17.16	1	8.35	-1.98	High
SC00271A	5-Fluorouracil	PD	1	7.63	1	2.56	-1.98	High
SC00437A	5-Fluorouracil	SD	0	27.02	1	9.86	-1.99	High

SC00404A	5-Fluorouracil	PD	1	3.35	1	2.79	-1.99	High
SC00327A	5-Fluorouracil	PD	1	3.65	1	3.19	-1.99	High
SC00291A	5-Fluorouracil	PD	1	8.81	1	2.83	-2.00	High
SC02332A	FOLFIRI	PD	1	19.63	1	3.06	-2.00	High
SC00403A	FOLFOX	SD	0	21.53	0	15.00	-2.00	High
SC00534A	5-Fluorouracil	SD	1	14.86	1	5.75	-2.00	High
SC02350A	FOLFIRI	SD	1	4.27	1	4.27	-2.01	High
SC00525A	5-Fluorouracil	PR	0	22.59	0	15.00	-2.01	High
SC00200A	5-Fluorouracil	SD	1	8.58	1	3.25	-2.01	High
SC00531A	5-Fluorouracil	PD	1	2.20	1	1.71	-2.02	High
SC00384A	5-Fluorouracil	PD	1	3.06	1	2.40	-2.02	High
SC00309A	FOLFOX	PR	1	17.56	0	15.00	-2.03	High
SC00511A	5-Fluorouracil	SD	1	6.38	1	5.23	-2.03	High
SC00364A	5-Fluorouracil	SD	1	27.09	1	6.77	-2.04	High
SC00483A	FOLFOX	PR	1	11.67	1	10.26	-2.04	High
SC00369A	5-Fluorouracil	SD	0	26.33	1	11.44	-2.05	High
SC00181A	5-Fluorouracil	PD	1	13.48	1	2.76	-2.06	High
SC00408A	5-Fluorouracil	PR	1	15.39	1	9.24	-2.07	High
SC00166A	5-Fluorouracil	SD	0	30.38	0	15.00	-2.07	High
SC00274A	FOLFOX	PR	1	7.66	1	7.00	-2.07	High
SC00269A	5-Fluorouracil	PD	1	4.08	1	2.53	-2.08	High
SC00512A	5-Fluorouracil	PR	1	7.92	1	6.31	-2.08	High
SC00504A	5-Fluorouracil	PD	1	9.40	1	3.45	-2.09	High
SC00237A	5-Fluorouracil	SD	1	16.27	1	8.81	-2.10	High
SC00293A	5-Fluorouracil	PR	1	21.47	1	8.48	-2.10	High
SC00227A	5-Fluorouracil	PR	0	6.97	0	6.97	-2.11	High
SC00280A	5-Fluorouracil	SD	1	11.05	1	5.49	-2.12	High
SC00475A	FOLFOX	SD	1	9.01	1	6.94	-2.12	High
SC00474A	FOLFOX	SD	1	16.41	1	8.52	-2.14	High
SC00431A	FOLFOX	PR	0	35.18	0	15.00	-2.16	High
SC00376A	5-Fluorouracil	SD	1	24.56	1	9.40	-2.17	High
SC00310A	5-Fluorouracil	PR	1	21.73	1	9.53	-2.17	High
SC00528A	5-Fluorouracil	PR	1	18.12	1	9.07	-2.17	High
SC00365A	5-Fluorouracil	SD	1	15.09	1	10.36	-2.18	High
SC00267A	5-Fluorouracil	SD	1	9.60	1	8.55	-2.18	High
SC02410A	FOLFIRI	PR	0	33.53	0	15.00	-2.18	High
SC00443A	5-Fluorouracil	PD	1	5.46	1	2.56	-2.18	High
SC00204A	5-Fluorouracil	PD	1	9.96	1	2.04	-2.19	High
SC00148A	FOLFOX	PR	1	7.36	1	5.88	-2.19	High
SC00457A	5-Fluorouracil	SD	1	12.99	1	2.04	-2.19	High
SC00156A	5-Fluorouracil	PR	1	15.95	1	6.12	-2.20	High
SC00178A	FOLFOX	PR	1	21.47	0	15.00	-2.20	High
SC00284A	5-Fluorouracil	PR	1	13.02	1	8.25	-2.20	High
SC00445A	5-Fluorouracil	SD	0	27.65	1	6.05	-2.20	High

SC00287A	5-Fluorouracil	SD	0	21.24	0	15.00	-2.20	High
SC00373A	5-Fluorouracil	PD	1	16.34	1	2.93	-2.21	High
SC00522A	5-Fluorouracil	CR	1	26.10	1	4.93	-2.21	High
SC00490A	FOLFOX	CR	0	23.77	0	15.00	-2.22	High
SC02411A	FOLFIRI	PR	1	11.84	1	9.86	-2.23	High
SC00325A	5-Fluorouracil	PD	1	10.95	1	3.25	-2.24	High
SC00355A	5-Fluorouracil	SD	1	11.24	1	11.24	-2.24	High
SC00125A	5-Fluorouracil	SD	1	13.71	1	8.28	-2.24	High
SC00212A	5-Fluorouracil	PD	1	10.09	1	2.50	-2.24	High
SC02389A	FOLFIRI	PR	0	48.36	0	15.00	-2.24	High
SC00294A	5-Fluorouracil	PR	1	23.24	1	12.03	-2.27	High
SC00398A	5-Fluorouracil	PR	1	23.70	1	9.17	-2.28	High
SC00415A	5-Fluorouracil	SD	1	21.63	1	9.24	-2.29	High
SC00432A	5-Fluorouracil	PR	1	24.16	1	8.98	-2.29	High
SC00440A	5-Fluorouracil	PD	1	4.64	1	3.45	-2.29	High
SC00268A	FOLFOX	PR	1	17.23	1	9.99	-2.29	High
SC00169A	5-Fluorouracil	PD	1	8.52	1	2.66	-2.29	High
SC00343A	FOLFOX	PD	1	8.35	1	3.45	-2.29	High
SC02404A	FOLFIRI	SD	1	9.24	1	7.86	-2.30	High
SC02399A	FOLFIRI	SD	1	8.45	1	5.56	-2.31	High
SC02339A	FOLFIRI	PD	1	12.69	1	2.86	-2.31	High
SC00319A	5-Fluorouracil	SD	1	17.56	1	6.31	-2.31	High
SC00189A	5-Fluorouracil	PD	1	8.55	1	1.87	-2.32	High
SC02336A	FOLFIRI	SD	1	10.92	1	8.55	-2.32	High
SC00485A	5-Fluorouracil	PR	1	13.45	1	8.35	-2.32	High
SC00224A	5-Fluorouracil	PD	1	9.04	1	2.63	-2.34	High
SC00150A	5-Fluorouracil	PD	0	39.12	1	2.53	-2.34	High
SC00175A	5-Fluorouracil	PR	1	7.82	1	7.82	-2.34	High
SC02337A	FOLFIRI	PR	1	4.44	1	3.22	-2.35	High
SC00206A	5-Fluorouracil	PR	0	31.40	0	15.00	-2.35	High
SC02357A	FOLFIRI	PD	1	18.31	1	3.22	-2.35	High
SC00444A	FOLFOX	PR	0	24.10	0	15.00	-2.35	High
SC00140A	5-Fluorouracil	PR	1	18.21	1	5.88	-2.36	High
SC00326A	5-Fluorouracil	PR	1	11.34	1	8.88	-2.36	High
SC00190A	5-Fluorouracil	SD	1	13.48	1	5.10	-2.36	High
SC00201A	5-Fluorouracil	SD	1	20.12	1	6.84	-2.37	High
SC02333A	FOLFIRI	SD	1	14.50	1	6.44	-2.38	High
SC00371A	5-Fluorouracil	PR	1	13.94	1	5.56	-2.38	High
SC00297A	5-Fluorouracil	PD	1	13.08	1	2.96	-2.38	High
SC00152A	5-Fluorouracil	SD	1	5.13	1	4.57	-2.40	Medium
SC02387A	FOLFIRI	PR	1	26.96	0	15.00	-2.42	Medium
SC00243A	5-Fluorouracil	SD	0	55.86	0	15.00	-2.43	Medium
SC00405A	5-Fluorouracil	PR	1	15.39	1	5.98	-2.43	Medium
SC00496A	5-Fluorouracil	SD	1	5.82	1	5.59	-2.43	Medium

SC02378A	FOLFIRI	SD	1	21.83	0	15.00	-2.43	Medium
SC00338A	5-Fluorouracil	SD	1	24.00	1	12.56	-2.43	Medium
SC00349A	5-Fluorouracil	PR	1	23.93	1	11.54	-2.43	Medium
SC00395A	5-Fluorouracil	SD	1	11.34	1	5.59	-2.44	Medium
SC00342A	5-Fluorouracil	SD	1	9.14	1	8.28	-2.44	Medium
SC00354A	FOLFOX	PR	1	25.12	1	14.37	-2.44	Medium
SC00482A	5-Fluorouracil	PR	1	14.93	1	5.59	-2.44	Medium
SC00353A	FOLFOX	PR	1	10.39	1	8.58	-2.44	Medium
SC00202A	5-Fluorouracil	SD	1	21.14	1	10.39	-2.45	Medium
SC00329A	5-Fluorouracil	SD	1	12.82	1	5.69	-2.46	Medium
SC00505A	FOLFOX	SD	1	8.38	1	8.38	-2.46	Medium
SC02408A	FOLFIRI	SD	1	10.65	1	5.88	-2.47	Medium
SC00429A	5-Fluorouracil	SD	1	7.79	1	5.56	-2.47	Medium
SC00225A	5-Fluorouracil	SD	1	17.92	1	9.11	-2.47	Medium
SC00301A	5-Fluorouracil	PD	1	13.12	1	3.02	-2.48	Medium
SC00486A	5-Fluorouracil	PD	1	4.04	1	3.25	-2.48	Medium
SC00416A	5-Fluorouracil	SD	1	7.00	1	5.56	-2.48	Medium
SC00524A	FOLFOX	PR	1	14.83	1	7.33	-2.48	Medium
SC00288A	5-Fluorouracil	PR	0	17.56	1	12.26	-2.49	Medium
SC02403A	FOLFIRI	PD	1	10.92	1	2.73	-2.49	Medium
SC00308A	5-Fluorouracil	PD	1	4.37	1	2.86	-2.49	Medium
SC00314A	5-Fluorouracil	PR	1	22.39	1	8.32	-2.49	Medium
SC00155A	5-Fluorouracil	PD	1	4.11	1	2.89	-2.49	Medium
SC00397A	5-Fluorouracil	SD	1	31.66	1	8.48	-2.49	Medium
SC00472A	5-Fluorouracil	SD	1	10.32	1	8.05	-2.49	Medium
SC00392A	5-Fluorouracil	PD	1	6.54	1	2.93	-2.50	Medium
SC00479A	5-Fluorouracil	PR	0	49.97	0	15.00	-2.50	Medium
SC02334A	FOLFIRI	PR	1	19.89	1	11.67	-2.50	Medium
SC00341A	FOLFOX	PR	1	8.71	1	6.18	-2.50	Medium
SC02385A	FOLFIRI	PR	1	10.95	1	6.02	-2.51	Medium
SC00196A	5-Fluorouracil	PR	1	23.28	1	8.02	-2.51	Medium
SC00262A	5-Fluorouracil	SD	1	18.87	1	11.21	-2.51	Medium
SC00488A	FOLFOX	PD	1	9.11	1	2.89	-2.52	Medium
SC00500A	5-Fluorouracil	PD	1	7.30	1	2.83	-2.52	Medium
SC00170A	FOLFOX	PR	1	5.92	1	5.92	-2.53	Medium
SC00246A	5-Fluorouracil	CR	1	16.34	1	14.47	-2.53	Medium
SC00439A	5-Fluorouracil	SD	0	19.07	1	9.53	-2.53	Medium
SC00455A	5-Fluorouracil	SD	1	10.22	1	5.56	-2.53	Medium
SC00179A	5-Fluorouracil	PR	1	16.14	1	9.14	-2.53	Medium
SC00422A	5-Fluorouracil	SD	1	8.58	1	5.39	-2.54	Medium
SC00141A	FOLFOX	PR	1	15.52	1	10.52	-2.54	Medium
SC00276A	FOLFOX	PR	1	14.70	1	14.70	-2.54	Medium
SC00146A	FOLFOX	PR	1	14.47	1	11.01	-2.55	Medium
SC02346A	FOLFIRI	SD	1	14.83	1	10.29	-2.55	Medium

SC02366A	FOLFIRI	PR	1	29.10	1	9.86	-2.55	Medium
SC00420A	5-Fluorouracil	PR	1	12.26	1	6.41	-2.56	Medium
SC00282A	5-Fluorouracil	SD	1	9.11	1	9.11	-2.57	Medium
SC00247A	5-Fluorouracil	SD	1	14.99	1	12.03	-2.57	Medium
SC00510A	5-Fluorouracil	PR	1	5.06	1	4.67	-2.57	Medium
SC00377A	5-Fluorouracil	PR	1	17.23	1	8.55	-2.58	Medium
SC00520A	5-Fluorouracil	SD	1	16.60	1	5.88	-2.58	Medium
SC02393A	FOLFIRI	PR	1	15.91	1	9.63	-2.58	Medium
SC00360A	5-Fluorouracil	PR	0	44.84	0	15.00	-2.59	Medium
SC02349A	FOLFIRI	SD	0	3.62	0	3.62	-2.59	Medium
SC00499A	5-Fluorouracil	PD	1	4.93	1	2.83	-2.59	Medium
SC00303A	5-Fluorouracil	SD	1	11.74	1	5.72	-2.60	Medium
SC00450A	FOLFOX	PR	1	11.77	1	6.77	-2.60	Medium
SC00312A	5-Fluorouracil	PD	1	3.52	1	2.50	-2.60	Medium
SC00299A	5-Fluorouracil	PR	1	55.40	0	15.00	-2.60	Medium
SC00173A	5-Fluorouracil	PR	1	41.36	1	14.60	-2.60	Medium
SC00205A	FOLFOX	PR	1	20.88	1	13.12	-2.61	Medium
SC00234A	5-Fluorouracil	PR	1	26.63	0	15.00	-2.61	Medium
SC00295A	5-Fluorouracil	SD	1	18.18	1	8.75	-2.61	Medium
SC02360A	FOLFIRI	PR	1	16.54	1	8.38	-2.61	Medium
SC00192A	5-Fluorouracil	PR	1	37.68	1	11.24	-2.61	Medium
SC00501A	5-Fluorouracil	SD	1	11.84	1	9.40	-2.62	Medium
SC00417A	5-Fluorouracil	SD	1	15.72	0	15.00	-2.62	Medium
SC00447A	5-Fluorouracil	SD	1	18.71	1	9.40	-2.62	Medium
SC00131A	5-Fluorouracil	PR	1	19.92	1	8.28	-2.64	Medium
SC00493A	5-Fluorouracil	PR	1	20.91	1	8.15	-2.64	Medium
SC00149A	5-Fluorouracil	PR	1	22.95	1	5.75	-2.65	Medium
SC00330A	5-Fluorouracil	SD	1	26.01	1	11.67	-2.65	Medium
SC02331A	FOLFIRI	SD	1	9.67	1	9.67	-2.65	Medium
SC00362A	5-Fluorouracil	SD	1	16.90	0	15.00	-2.65	Medium
SC00419A	5-Fluorouracil	PD	1	3.35	1	1.64	-2.66	Medium
SC00259A	FOLFOX	PD	1	8.55	1	2.93	-2.66	Medium
SC00407A	5-Fluorouracil	PR	1	27.98	1	9.27	-2.66	Medium
SC02390A	FOLFIRI	PR	1	21.17	1	11.15	-2.68	Medium
SC00518A	5-Fluorouracil	PD	1	12.76	1	3.85	-2.68	Medium
SC00358A	5-Fluorouracil	SD	1	20.68	1	9.07	-2.69	Medium
SC00263A	FOLFOX	PR	1	19.69	1	9.17	-2.70	Medium
SC00465A	5-Fluorouracil	SD	1	18.08	1	10.62	-2.70	Medium
SC00191A	5-Fluorouracil	PD	1	4.21	1	2.73	-2.70	Medium
SC00184A	FOLFOX	PR	1	20.81	0	15.00	-2.70	Medium
SC02368A	FOLFIRI	PD	1	10.59	1	3.72	-2.70	Medium
SC00389A	5-Fluorouracil	SD	1	14.93	1	6.41	-2.70	Medium
SC00448A	FOLFOX		1	20.15	1	11.24	-2.70	Medium
SC00473A	FOLFOX	SD	1	10.06	1	8.84	-2.71	Medium

SC00281A	FOLFOX	PR	1	19.27	1	6.90	-2.72	Medium
SC00266A	5-Fluorouracil	PR	1	16.77	1	12.72	-2.72	Medium
SC00260A	5-Fluorouracil	PR	1	12.76	1	8.61	-2.73	Medium
SC00406A	FOLFOX	PR	1	23.67	1	13.48	-2.74	Medium
SC00257A	FOLFOX	PD	1	8.65	1	2.86	-2.74	Medium
SC00321A	5-Fluorouracil	SD	1	6.08	1	6.08	-2.76	Medium
SC00252A	FOLFOX	PD	1	3.95	1	2.66	-2.77	Medium
SC00469A	5-Fluorouracil	PR	1	33.37	1	9.93	-2.77	Medium
SC00519A	5-Fluorouracil	PR	0	25.08	1	13.35	-2.77	Medium
SC00532A	5-Fluorouracil	PD	1	7.07	1	1.15	-2.77	Medium
SC00313A	5-Fluorouracil	SD	1	8.98	1	6.41	-2.77	Medium
SC00396A	5-Fluorouracil	PR	1	28.01	1	8.52	-2.78	Medium
SC02398A	FOLFIRI	PR	1	13.84	1	8.12	-2.78	Medium
SC02371A	FOLFIRI	PR	1	30.94	1	12.89	-2.79	Medium
SC02355A	FOLFIRI	PR	1	27.39	1	9.17	-2.80	Medium
SC00361A	5-Fluorouracil	SD	1	22.06	1	8.75	-2.80	Medium
SC00213A	5-Fluorouracil	SD	0	34.16	0	15.00	-2.80	Medium
SC00366A	5-Fluorouracil	PD	1	8.81	1	3.02	-2.80	Medium
SC02363A	FOLFIRI	SD	1	26.47	1	8.81	-2.81	Medium
SC00180A	5-Fluorouracil	PR	1	16.64	1	7.79	-2.81	Medium
SC00124A	5-Fluorouracil	SD	1	2.30	1	2.30	-2.82	Medium
SC00393A	5-Fluorouracil	PR	0	36.13	0	15.00	-2.82	Medium
SC02373A	FOLFIRI	PR	1	14.73	1	7.56	-2.83	Medium
SC00394A	5-Fluorouracil	SD	1	14.14	1	5.19	-2.83	Medium
SC00332A	5-Fluorouracil	PD	1	2.73	1	1.81	-2.84	Medium
SC00467A	5-Fluorouracil	PR	1	7.56	1	7.56	-2.84	Medium
SC00159A	5-Fluorouracil	PD	1	10.42	1	2.56	-2.85	Medium
SC00340A	5-Fluorouracil	PR	1	36.92	0	15.00	-2.85	Medium
SC00509A	5-Fluorouracil	PR	0	20.68	1	6.51	-2.86	Medium
SC02407A	FOLFIRI	SD	1	13.28	1	8.75	-2.86	Medium
SC00418A	FOLFOX	SD	1	12.76	1	0.00	-2.86	Medium
SC00195A	5-Fluorouracil	PR	1	38.43	0	15.00	-2.87	Medium
SC00413A	5-Fluorouracil	PR	1	23.44	1	13.02	-2.89	Medium
SC00480A	5-Fluorouracil	PD	1	11.61	1	3.45	-2.89	Medium
SC00459A	5-Fluorouracil	PD	1	19.33	1	2.96	-2.90	Low
SC00272A	5-Fluorouracil	PR	1	45.17	1	8.65	-2.90	Low
SC00527A	5-Fluorouracil	PR	0	22.03	1	8.88	-2.90	Low
SC00386A	5-Fluorouracil	SD	1	17.06	1	5.88	-2.90	Low
SC02367A	FOLFIRI	SD	1	12.23	1	8.35	-2.90	Low
SC00468A	5-Fluorouracil	PR	1	21.63	1	8.52	-2.91	Low
SC00226A	5-Fluorouracil	PR	1	21.76	0	15.00	-2.91	Low
SC02397A	FOLFIRI	PR	1	16.90	1	5.06	-2.92	Low
SC00315A	5-Fluorouracil	SD	1	31.79	1	5.23	-2.92	Low
SC00254A	5-Fluorouracil	PD	1	20.88	1	2.40	-2.92	Low

SC00177A	5-Fluorouracil	PD	1	14.40	1	2.96	-2.92	Low
SC00236A	5-Fluorouracil	PR	1	26.07	1	12.59	-2.92	Low
SC00185A	5-Fluorouracil	SD	1	11.77	1	9.01	-2.93	Low
SC00412A	5-Fluorouracil	SD	1	16.54	1	10.68	-2.93	Low
SC00425A	5-Fluorouracil	PR	1	9.53	1	8.71	-2.94	Low
SC00214A	5-Fluorouracil	PD	1	7.79	1	2.53	-2.94	Low
SC00129A	FOLFOX	PR	1	15.75	1	8.32	-2.94	Low
SC00264A	5-Fluorouracil	SD	1	15.55	1	9.96	-2.95	Low
SC02348A	FOLFIRI	PD	1	17.56	1	3.78	-2.95	Low
SC00130A	5-Fluorouracil	SD	1	41.79	1	8.25	-2.95	Low
SC00278A	5-Fluorouracil	SD	1	18.35	1	11.74	-2.95	Low
SC00320A	5-Fluorouracil	PD	1	7.63	1	3.02	-2.96	Low
SC00380A	FOLFOX	PR	0	22.13	1	10.36	-2.96	Low
SC00357A	FOLFOX	PR	1	7.66	1	7.66	-2.96	Low
SC00241A	5-Fluorouracil	SD	1	11.84	1	11.84	-2.98	Low
SC00142A	5-Fluorouracil	SD	1	25.78	1	12.36	-2.98	Low
SC00322A	5-Fluorouracil	PR	1	11.97	1	8.78	-2.98	Low
SC00507A	5-Fluorouracil	PD	1	12.85	1	3.42	-2.99	Low
SC00290A	5-Fluorouracil	SD	1	9.27	1	5.52	-2.99	Low
SC00402A	5-Fluorouracil	PD	1	4.57	1	2.53	-2.99	Low
SC00273A	FOLFOX	PR	1	9.24	1	6.67	-2.99	Low
SC00228A	5-Fluorouracil	PD	1	5.06	1	2.50	-2.99	Low
SC00244A	5-Fluorouracil	PR	1	30.35	1	9.21	-2.99	Low
SC00286A	5-Fluorouracil	PD	1	8.38	1	2.89	-2.99	Low
SC00381A	5-Fluorouracil	PR	1	16.21	1	7.27	-3.00	Low
SC00449A	5-Fluorouracil	PD	1	5.36	1	2.86	-3.00	Low
SC00333A	5-Fluorouracil	SD	1	11.93	1	5.29	-3.01	Low
SC02395A	FOLFIRI	SD	1	13.97	1	9.07	-3.01	Low
SC02380A	FOLFIRI	PR	1	10.85	1	9.30	-3.01	Low
SC00508A	5-Fluorouracil	PR	1	28.04	1	14.24	-3.01	Low
SC00305A	5-Fluorouracil	PR	1	28.87	1	9.67	-3.02	Low
SC00232A	5-Fluorouracil	PD	1	13.15	1	2.47	-3.02	Low
SC00463A	5-Fluorouracil	SD	1	25.61	1	6.74	-3.02	Low
SC00270A	5-Fluorouracil	PR	1	8.19	1	6.81	-3.02	Low
SC00401A	5-Fluorouracil	PD	0	26.53	1	2.37	-3.03	Low
SC02326A	FOLFIRI	PR	1	43.23	1	12.36	-3.04	Low
SC02376A	FOLFIRI	PR	1	8.32	1	7.46	-3.05	Low
SC00461A	5-Fluorouracil	SD	1	6.67	1	5.03	-3.05	Low
SC00253A	5-Fluorouracil	SD	1	18.64	1	7.63	-3.06	Low
SC00197A	5-Fluorouracil	PD	1	23.18	1	2.76	-3.06	Low
SC00162A	FOLFOX	PR	1	41.19	1	6.38	-3.06	Low
SC02370A	FOLFIRI	SD	1	23.70	0	15.00	-3.06	Low
SC00221A	5-Fluorouracil	SD	0	32.81	1	8.55	-3.07	Low
SC00375A	5-Fluorouracil	PR	1	12.49	1	10.26	-3.08	Low

SC02372A	FOLFIRI	PD	1	5.85	1	2.01	-3.09	Low
SC00491A	5-Fluorouracil	PD	1	12.33	1	3.06	-3.09	Low
SC00174A	5-Fluorouracil	SD	1	21.73	1	5.95	-3.10	Low
SC00382A	5-Fluorouracil	SD	0	62.76	1	13.18	-3.11	Low
SC00385A	5-Fluorouracil	SD	1	22.78	1	7.69	-3.13	Low
SC00368A	5-Fluorouracil	PR	1	20.32	0	15.00	-3.13	Low
SC00487A	5-Fluorouracil	PD	0	27.62	1	2.93	-3.13	Low
SC00516A	5-Fluorouracil	PD	1	7.30	1	1.78	-3.13	Low
SC00311A	5-Fluorouracil	PD	1	9.44	1	2.66	-3.13	Low
SC00388A	5-Fluorouracil	PD	1	10.03	1	3.12	-3.13	Low
SC00229A	5-Fluorouracil	SD	1	10.29	1	5.06	-3.14	Low
SC00163A	5-Fluorouracil	SD	1	24.89	0	15.00	-3.15	Low
SC00387A	5-Fluorouracil	SD	0	54.41	1	14.07	-3.15	Low
SC02347A	FOLFIRI	SD	1	18.21	1	9.80	-3.15	Low
SC00533A	5-Fluorouracil	SD	1	23.24	1	9.01	-3.15	Low
SC00481A	FOLFOX	SD	1	20.98	1	12.20	-3.15	Low
SC00283A	5-Fluorouracil	PR	0	31.07	0	15.00	-3.17	Low
SC00171A	5-Fluorouracil	PD	1	5.23	1	2.04	-3.17	Low
SC02335A	FOLFIRI	PR	1	20.42	1	12.39	-3.18	Low
SC02353A	FOLFIRI	SD	1	30.38	1	10.59	-3.19	Low
SC00258A	5-Fluorouracil	SD	1	9.40	1	5.10	-3.20	Low
SC00438A	5-Fluorouracil	PR	1	9.99	1	6.90	-3.21	Low
SC00138A	FOLFOX	PR	1	19.20	1	9.17	-3.23	Low
SC00245A	5-Fluorouracil	SD	1	25.41	0	15.00	-3.23	Low
SC00132A	5-Fluorouracil	PR	0	44.45	0	15.00	-3.24	Low
SC02345A	FOLFIRI	PD	0	5.46	1	2.70	-3.24	Low
SC02341A	FOLFIRI	SD	1	9.04	1	7.50	-3.24	Low
SC00359A	5-Fluorouracil	PR	1	13.05	1	8.81	-3.25	Low
SC00208A	5-Fluorouracil	PD	1	8.61	1	2.63	-3.25	Low
SC00217A	FOLFOX	PR	1	13.12	1	13.12	-3.26	Low
SC00462A	5-Fluorouracil	SD	1	30.58	1	5.52	-3.26	Low
SC00242A	5-Fluorouracil	SD	1	18.58	1	8.71	-3.26	Low
SC00426A	FOLFOX	PD	1	10.98	1	2.99	-3.27	Low
SC00193A	5-Fluorouracil	SD	1	17.23	1	5.52	-3.28	Low
SC02412A	FOLFIRI	SD	1	24.79	1	12.59	-3.28	Low
SC00367A	5-Fluorouracil	PR	0	27.48	1	9.90	-3.29	Low
SC00345A	5-Fluorouracil	PR	1	64.27	0	15.00	-3.29	Low
SC00476A	FOLFOX	PR	1	8.09	1	5.79	-3.30	Low
SC00233A	5-Fluorouracil	SD	1	31.53	1	9.37	-3.30	Low
SC02375A	FOLFIRI	PR	1	44.15	1	13.25	-3.30	Low
SC00424A	FOLFOX	PR	1	27.52	1	10.06	-3.30	Low
SC02374A	FOLFIRI	PR	1	13.55	1	11.21	-3.31	Low
SC00151A	5-Fluorouracil	PD	1	3.42	1	2.79	-3.31	Low
SC02343A	FOLFIRI	PR	0	15.42	1	11.47	-3.32	Low

SC00194A	5-Fluorouracil	PR	1	4.14	1	4.14	-3.32	Low
SC00216A	5-Fluorouracil	SD	1	5.26	1	4.01	-3.33	Low
SC00133A	FOLFOX	PR	1	11.24	1	10.59	-3.34	Low
SC00220A	FOLFOX	PR	1	27.65	1	9.90	-3.36	Low
SC00255A	5-Fluorouracil	SD	1	11.61	1	6.02	-3.36	Low
SC00230A	5-Fluorouracil	SD	1	19.13	1	5.52	-3.37	Low
SC00265A	FOLFOX	PR	1	20.15	1	13.02	-3.37	Low
SC00164A	5-Fluorouracil	PD	1	9.01	1	2.60	-3.37	Low
SC00399A	5-Fluorouracil	PR	0	21.53	1	9.17	-3.37	Low
SC00239A	5-Fluorouracil	PD	1	29.79	1	3.16	-3.38	Low
SC00339A	5-Fluorouracil	PR	1	27.68	0	15.00	-3.40	Low
SC00318A	FOLFOX	PR	1	22.13	1	9.53	-3.42	Low
SC00456A	5-Fluorouracil	PD	1	8.94	1	3.02	-3.43	Low
SC00292A	FOLFOX	SD	1	25.55	1	14.56	-3.44	Low
SC00199A	5-Fluorouracil	PD	1	13.78	1	2.60	-3.45	Low
SC02340A	FOLFIRI	PD	1	7.76	1	6.05	-3.45	Low
SC00451A	5-Fluorouracil	SD	1	33.47	1	13.25	-3.45	Low
SC00128A	5-Fluorouracil	PR	1	17.49	1	9.83	-3.45	Low
SC02394A	FOLFIRI	SD	1	15.81	1	9.07	-3.47	Low
SC00144A	5-Fluorouracil	PR	1	14.20	1	8.68	-3.48	Low
SC00521A	5-Fluorouracil	SD	1	9.37	1	8.19	-3.49	Low
SC02354A	FOLFIRI	PR	0	46.95	1	14.30	-3.50	Low
SC00198A	5-Fluorouracil	SD	0	29.95	1	6.48	-3.50	Low
SC00158A	FOLFOX	SD	1	12.69	1	7.73	-3.50	Low
SC00497A	5-Fluorouracil	SD	0	9.27	0	9.27	-3.52	Low
SC00302A	FOLFOX	SD	1	13.35	1	8.28	-3.57	Low
SC00464A	5-Fluorouracil	PR	1	10.22	1	5.52	-3.57	Low
SC00289A	FOLFOX	SD	1	46.85	0	15.00	-3.60	Low
SC00351A	5-Fluorouracil	PR	1	10.98	1	9.67	-3.61	Low
SC00153A	5-Fluorouracil	PD	1	16.44	1	3.58	-3.62	Low
SC02396A	FOLFIRI	SD	1	8.55	1	8.55	-3.63	Low
SC00348A	5-Fluorouracil	SD	1	16.44	1	9.04	-3.64	Low
SC00147A	5-Fluorouracil	SD	1	22.55	1	5.52	-3.67	Low
SC00411A	FOLFOX	PR	1	28.08	1	9.11	-3.70	Low
SC00296A	FOLFOX	SD	1	13.12	1	8.94	-3.74	Low
SC02329A	FOLFIRI	SD	1	13.61	1	13.61	-3.75	Low
SC02342A	FOLFIRI	PD	1	4.90	1	3.85	-3.76	Low
SC00218A	5-Fluorouracil	PR	1	27.22	1	10.62	-3.79	Low
SC00526A	5-Fluorouracil	PR	0	15.72	1	9.44	-3.80	Low
SC02381A	FOLFIRI	PR	1	18.87	0	15.00	-3.81	Low
SC00383A	5-Fluorouracil	PR	1	19.30	1	6.35	-3.84	Low
SC00454A	5-Fluorouracil	PR	1	37.81	0	15.00	-3.86	Low
SC00410A	FOLFOX	SD	1	19.82	1	9.47	-3.91	Low
SC00423A	FOLFOX	PR	1	4.04	1	4.04	-3.96	Low

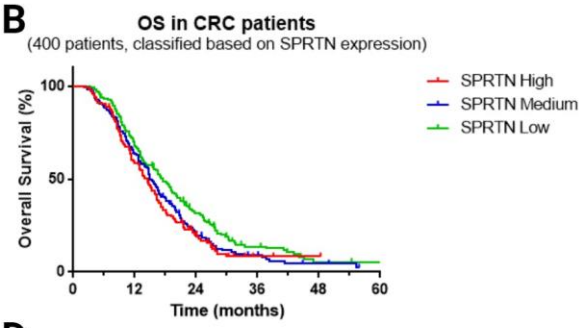
SC00250A	5-Fluorouracil	PD	1	11.21	1	2.86	-3.98	Low
SC00251A	5-Fluorouracil	SD	0	41.82	0	15.00	-3.98	Low
SC00187A	5-Fluorouracil	SD	0	22.72	1	6.31	-4.06	Low
SC02382A	FOLFIRI	PD	1	10.13	1	2.01	-4.25	Low
SC00316A	5-Fluorouracil	PR	1	26.83	1	8.15	-4.34	Low
SC00453A	5-Fluorouracil	PR	1	11.67	1	6.51	-4.56	Low
SC00137A	5-Fluorouracil	SD	0	36.66	1	8.48	-4.77	Low

Supplementary table 8. Clinical and IHC data of 400 CRC patients (Oxford Cohort). Clinical and immunohistochemical data of the 400 colorectal cancer patients analysed in Figure 45 and Figure 46. Clinical response at 12 weeks is indicated as complete response (CR), partial response (PR), stable disease (SD) or progressive disease (PD).

A

		Number of patients	Treatment	Median OS (months)
Overall Survival (Oxford Cohort)	SPRTN High	124	5-Fluorouracil, FOLFIRI or FOLFOX	14.47
	SPRTN Medium	127	5-Fluorouracil, FOLFIRI or FOLFOX	15.39
	SPRTN Low	149	5-Fluorouracil, FOLFIRI or FOLFOX	17.49

B



C

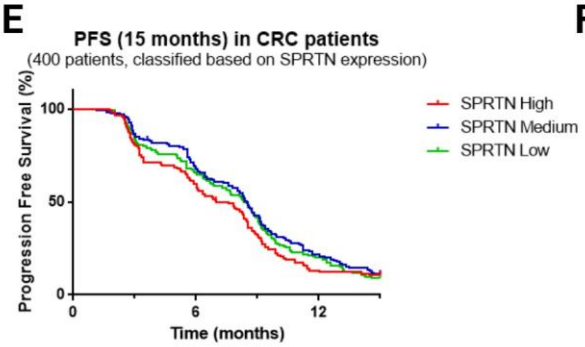
Statistical analysis

Comparison	Hazard Ratio (HR)	Confidence interval (95% CI)	p-value
SPRTN High vs SPRTN Low	1.392	1.062 - 1.823	**0.016**
SPRTN High vs SPRTN Medium	0.899	0.672 - 1.203	0.474
SPRTN Medium vs SPRTN Low	1.651	1.252 - 2.176	**<0.001**

D

		Number of patients	Treatment	Median PFS (months)
Progression Free Survival (Oxford Cohort)	SPRTN High	124	5-Fluorouracil, FOLFIRI or FOLFOX	7.00
	SPRTN Medium	127	5-Fluorouracil, FOLFIRI or FOLFOX	8.48
	SPRTN Low	149	5-Fluorouracil, FOLFIRI or FOLFOX	8.35

E



F

Statistical analysis

Comparison	Hazard Ratio (HR)	Confidence interval (95% CI)	p-value
SPRTN High vs SPRTN Low	1.149	0.885 - 1.491	0.298
SPRTN High vs SPRTN Medium	1.158	0.881 - 1.521	0.293
SPRTN Medium vs SPRTN Low	0.974	0.748 - 1.261	0.844

Supplementary Figure 1. SPRTN possible role as a prognostic biomarker in CRC patients (Oxford cohort – 400 patients).

(A) Overall survival summary of the 400 CRC patients in the Oxford cohort, divided according to SPRTN expression (B) Overall survival of the 400 patients enrolled in the Oxford cohort. Patients were divided according to SPRTN expression and their survival was plotted using Kaplan Meyer curves. (C) Cox regression analysis of the data presented in Supplementary Figure 1B. (D) Progression free survival summary of the 400 CRC patients in the Oxford cohort, divided according to SPRTN expression (E) Progression free survival of the 400 patients enrolled in the Oxford cohort. Patients were divided according to SPRTN expression and their survival was plotted using Kaplan Meyer curves. (F) Cox regression analysis of the data presented in Supplementary Figure 1E.

References

1. Hanahan, D. and R.A. Weinberg, The hallmarks of cancer. Cell, 2000. **100**(1): p. 57-70.
2. Hanahan, D. and R.A. Weinberg, Hallmarks of cancer: the next generation. Cell, 2011. **144**(5): p. 646-74.
3. Hanahan, D., Hallmarks of Cancer: New Dimensions. Cancer Discov, 2022. **12**(1): p. 31-46.
4. Bray, F., et al., Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin, 2018. **68**(6): p. 394-424.
5. Cree, I.A., Cancer biology. Methods Mol Biol, 2011. **731**: p. 1-11.
6. Stratton, M.R., P.J. Campbell, and P.A. Futreal, The cancer genome. Nature, 2009. **458**(7239): p. 719-24.
7. Lee, E.Y. and W.J. Muller, Oncogenes and tumor suppressor genes. Cold Spring Harb Perspect Biol, 2010. **2**(10): p. a003236.
8. Milholland, B., et al., Differences between germline and somatic mutation rates in humans and mice. Nat Commun, 2017. **8**: p. 15183.
9. Zhang, L. and J. Vijg, Somatic Mutagenesis in Mammals and Its Implications for Human Disease and Aging. Annu Rev Genet, 2018. **52**: p. 397-419.
10. Dekker, E., et al., Colorectal cancer. Lancet, 2019. **394**(10207): p. 1467-1480.
11. Tamas, K., et al., Rectal and colon cancer: Not just a different anatomic site. Cancer Treat Rev, 2015. **41**(8): p. 671-9.
12. Siegel, R., C. Desantis, and A. Jemal, Colorectal cancer statistics, 2014. CA Cancer J Clin, 2014. **64**(2): p. 104-17.
13. Cremolini, C., et al., First-line chemotherapy for mCRC-a review and evidence-based algorithm. Nat Rev Clin Oncol, 2015. **12**(10): p. 607-19.
14. Loupakis, F., et al., Initial therapy with FOLFOXIRI and bevacizumab for metastatic colorectal cancer. N Engl J Med, 2014. **371**(17): p. 1609-18.
15. Puppa, G., et al., TNM staging system of colorectal carcinoma: a critical appraisal of challenging issues. Arch Pathol Lab Med, 2010. **134**(6): p. 837-52.
16. Guinney, J., et al., The consensus molecular subtypes of colorectal cancer. Nat Med, 2015. **21**(11): p. 1350-6.
17. Isella, C., et al., Selective analysis of cancer-cell intrinsic transcriptional traits defines novel clinically relevant subtypes of colorectal cancer. Nat Commun, 2017. **8**: p. 15107.
18. Telloni, S.M., Tumor Staging and Grading: A Primer. Methods Mol Biol, 2017. **1606**: p. 1-17.
19. Weiser, M.R., AJCC 8th Edition: Colorectal Cancer. Ann Surg Oncol, 2018. **25**(6): p. 1454-1455.
20. Amin, M.B., et al., The Eighth Edition AJCC Cancer Staging Manual: Continuing to build a bridge from a population-based to a more "personalized" approach to cancer staging. CA Cancer J Clin, 2017. **67**(2): p. 93-99.
21. Compton, C.C. and F.L. Greene, The staging of colorectal cancer: 2004 and beyond. CA Cancer J Clin, 2004. **54**(6): p. 295-308.
22. Drouillard, A. and C. Lepage, [Colorectal cancer surgery with curative intent: what surveillance?]. Rev Prat, 2015. **65**(6): p. 798-801.
23. de Gramont, A., et al., Leucovorin and fluorouracil with or without oxaliplatin as first-line treatment in advanced colorectal cancer. J Clin Oncol, 2000. **18**(16): p. 2938-47.
24. Douillard, J.Y., et al., Irinotecan combined with fluorouracil compared with fluorouracil alone as first-line treatment for metastatic colorectal cancer: a multicentre randomised trial. Lancet, 2000. **355**(9209): p. 1041-7.
25. Colucci, G., et al., Phase III randomized trial of FOLFIRI versus FOLFOX4 in the treatment of advanced colorectal cancer: a multicenter study of the Gruppo Oncologico Dell'Italia Meridionale. J Clin Oncol, 2005. **23**(22): p. 4866-75.
26. Tournigand, C., et al., FOLFIRI followed by FOLFOX6 or the reverse sequence in advanced colorectal cancer: a randomized GERCOR study. J Clin Oncol, 2004. **22**(2): p. 229-37.
27. Golfinopoulos, V., et al., Survival and disease-progression benefits with treatment regimens for advanced colorectal cancer: a meta-analysis. Lancet Oncol, 2007. **8**(10): p. 898-911.

28. Hurwitz, H., et al., Bevacizumab plus irinotecan, fluorouracil, and leucovorin for metastatic colorectal cancer. N Engl J Med, 2004. **350**(23): p. 2335-42.
29. Van Cutsem, E., et al., Cetuximab and chemotherapy as initial treatment for metastatic colorectal cancer. N Engl J Med, 2009. **360**(14): p. 1408-17.
30. Douillard, J.Y., et al., Randomized, phase III trial of panitumumab with infusional fluorouracil, leucovorin, and oxaliplatin (FOLFOX4) versus FOLFOX4 alone as first-line treatment in patients with previously untreated metastatic colorectal cancer: the PRIME study. J Clin Oncol, 2010. **28**(31): p. 4697-705.
31. Diaz, L.A., Jr., et al., Pembrolizumab versus chemotherapy for microsatellite instability-high or mismatch repair-deficient metastatic colorectal cancer (KEYNOTE-177): final analysis of a randomised, open-label, phase 3 study. Lancet Oncol, 2022. **23**(5): p. 659-670.
32. Buchler, T., Microsatellite Instability and Metastatic Colorectal Cancer - A Clinical Perspective. Front Oncol, 2022. **12**: p. 888181.
33. Cercek, A., et al., PD-1 Blockade in Mismatch Repair-Deficient, Locally Advanced Rectal Cancer. N Engl J Med, 2022. **386**(25): p. 2363-2376.
34. Morris, V.K., et al., Treatment of Metastatic Colorectal Cancer: ASCO Guideline. J Clin Oncol, 2022: p. JCO2201690.
35. Chiorean, E.G., et al., Treatment of Patients With Late-Stage Colorectal Cancer: ASCO Resource-Stratified Guideline. JCO Glob Oncol, 2020. **6**: p. 414-438.
36. Pino, M.S. and D.C. Chung, The chromosomal instability pathway in colon cancer. Gastroenterology, 2010. **138**(6): p. 2059-72.
37. Cancer Genome Atlas, N., Comprehensive molecular characterization of human colon and rectal cancer. Nature, 2012. **487**(7407): p. 330-7.
38. Nusse, R. and H. Clevers, Wnt/beta-Catenin Signaling, Disease, and Emerging Therapeutic Modalities. Cell, 2017. **169**(6): p. 985-999.
39. Sawa, H., Control of cell polarity and asymmetric division in C. elegans. Curr Top Dev Biol, 2012. **101**: p. 55-76.
40. Komiya, Y. and R. Habas, Wnt signal transduction pathways. Organogenesis, 2008. **4**(2): p. 68-75.
41. Zhao, H., et al., Wnt signaling in colorectal cancer: pathogenic role and therapeutic target. Mol Cancer, 2022. **21**(1): p. 144.
42. Janda, C.Y., et al., Surrogate Wnt agonists that phenocopy canonical Wnt and beta-catenin signalling. Nature, 2017. **545**(7653): p. 234-237.
43. Kitagawa, M., et al., An F-box protein, FWD1, mediates ubiquitin-dependent proteolysis of beta-catenin. EMBO J, 1999. **18**(9): p. 2401-10.
44. Rowan, A.J., et al., APC mutations in sporadic colorectal tumors: A mutational "hotspot" and interdependence of the "two hits". Proc Natl Acad Sci U S A, 2000. **97**(7): p. 3352-7.
45. Rubinfeld, B., et al., Binding of GSK3beta to the APC-beta-catenin complex and regulation of complex assembly. Science, 1996. **272**(5264): p. 1023-6.
46. Mauri, G., et al., The Evolutionary Landscape of Treatment for BRAF(V600E) Mutant Metastatic Colorectal Cancer. Cancers (Basel), 2021. **13**(1).
47. Yarden, Y. and B.Z. Shilo, SnapShot: EGFR signaling pathway. Cell, 2007. **131**(5): p. 1018.
48. Simanshu, D.K., D.V. Nissley, and F. McCormick, RAS Proteins and Their Regulators in Human Disease. Cell, 2017. **170**(1): p. 17-33.
49. Downward, J., Targeting RAS signalling pathways in cancer therapy. Nat Rev Cancer, 2003. **3**(1): p. 11-22.
50. Lavoie, H., J. Gagnon, and M. Therrien, ERK signalling: a master regulator of cell behaviour, life and fate. Nat Rev Mol Cell Biol, 2020. **21**(10): p. 607-632.
51. Lievre, A., et al., KRAS mutations as an independent prognostic factor in patients with advanced colorectal cancer treated with cetuximab. J Clin Oncol, 2008. **26**(3): p. 374-9.
52. Dean, L. and M. Kane, Cetuximab Therapy and RAS and BRAF Genotype, in Medical Genetics Summaries, V.M. Pratt, et al., Editors. 2012: Bethesda (MD).
53. Dean, L. and M. Kane, Panitumumab Therapy and RAS and BRAF Genotype, in Medical Genetics Summaries, V.M. Pratt, et al., Editors. 2012: Bethesda (MD).
54. Soussi, T., The p53 tumor suppressor gene: from molecular biology to clinical investigation. Ann N Y Acad Sci, 2000. **910**: p. 121-37; discussion 137-9.
55. Levine, A.J., p53: 800 million years of evolution and 40 years of discovery. Nat Rev Cancer, 2020. **20**(8): p. 471-480.

56. Hassin, O., et al., Different hotspot p53 mutants exert distinct phenotypes and predict outcome of colorectal cancer patients. Nat Commun, 2022. **13**(1): p. 2800.
57. Blackford, A.N. and S.P. Jackson, ATM, ATR, and DNA-PK: The Trinity at the Heart of the DNA Damage Response. Mol Cell, 2017. **66**(6): p. 801-817.
58. Smith, J., et al., The ATM-Chk2 and ATR-Chk1 pathways in DNA damage signaling and cancer. Adv Cancer Res, 2010. **108**: p. 73-112.
59. Pabla, N., et al., ATR-Chk2 signaling in p53 activation and DNA damage response during cisplatin-induced apoptosis. J Biol Chem, 2008. **283**(10): p. 6572-83.
60. van Jaarsveld, M.T.M., et al., Tissue-Specific Chk1 Activation Determines Apoptosis by Regulating the Balance of p53 and p21. iScience, 2019. **12**: p. 27-40.
61. Tibbetts, R.S., et al., A role for ATR in the DNA damage-induced phosphorylation of p53. Genes Dev, 1999. **13**(2): p. 152-7.
62. Fischer, M., Census and evaluation of p53 target genes. Oncogene, 2017. **36**(28): p. 3943-3956.
63. Zhang, X. and T. Sjoblom, Targeting Loss of Heterozygosity: A Novel Paradigm for Cancer Therapy. Pharmaceuticals (Basel), 2021. **14**(1).
64. Jernvall, P., et al., Loss of heterozygosity at 18q21 is indicative of recurrence and therefore poor prognosis in a subset of colorectal cancers. Br J Cancer, 1999. **79**(5-6): p. 903-8.
65. Ogino, S., et al., Prognostic significance and molecular associations of 18q loss of heterozygosity: a cohort study of microsatellite stable colorectal cancers. J Clin Oncol, 2009. **27**(27): p. 4591-8.
66. Fearon, E.R., et al., Identification of a chromosome 18q gene that is altered in colorectal cancers. Science, 1990. **247**(4938): p. 49-56.
67. Hahn, S.A., et al., DPC4, a candidate tumor suppressor gene at human chromosome 18q21.1. Science, 1996. **271**(5247): p. 350-3.
68. Anmarkrud, J.A., et al., Microsatellite evolution: Mutations, sequence variation, and homoplasy in the hypervariable avian microsatellite locus HrU10. BMC Evol Biol, 2008. **8**: p. 138.
69. Schlotterer, C., Evolutionary dynamics of microsatellite DNA. Chromosoma, 2000. **109**(6): p. 365-71.
70. Liu, D., G. Keijzers, and L.J. Rasmussen, DNA mismatch repair and its many roles in eukaryotic cells. Mutat Res Rev Mutat Res, 2017. **773**: p. 174-187.
71. Harfe, B.D. and S. Jinks-Robertson, DNA mismatch repair and genetic instability. Annu Rev Genet, 2000. **34**: p. 359-399.
72. Plotz, G., et al., Analysis of the human MutLalpha.MutSalpha complex. Biochem Biophys Res Commun, 2006. **340**(3): p. 852-9.
73. Vilar, E. and S.B. Gruber, Microsatellite instability in colorectal cancer-the stable evidence. Nat Rev Clin Oncol, 2010. **7**(3): p. 153-62.
74. Zhang, Z., et al., Neoantigen: A New Breakthrough in Tumor Immunotherapy. Front Immunol, 2021. **12**: p. 672356.
75. Andre, T., et al., Pembrolizumab in Microsatellite-Instability-High Advanced Colorectal Cancer. N Engl J Med, 2020. **383**(23): p. 2207-2218.
76. Ogino, S., et al., CpG island methylator phenotype (CIMP) of colorectal cancer is best characterised by quantitative DNA methylation analysis and prospective cohort studies. Gut, 2006. **55**(7): p. 1000-6.
77. Lao, V.V. and W.M. Grady, Epigenetics and colorectal cancer. Nat Rev Gastroenterol Hepatol, 2011. **8**(12): p. 686-700.
78. Sanchez-Vega, F., et al., CpG island methylator phenotype in adenocarcinomas from the digestive tract: Methods, conclusions, and controversies. World J Gastrointest Oncol, 2017. **9**(3): p. 105-120.
79. Issa, J.P., Aging and epigenetic drift: a vicious cycle. J Clin Invest, 2014. **124**(1): p. 24-9.
80. Tao, Y., et al., Aging-like Spontaneous Epigenetic Silencing Facilitates Wnt Activation, Stemness, and Braf(V600E)-Induced Tumorigenesis. Cancer Cell, 2019. **35**(2): p. 315-328 e6.
81. Ferreira, H.J. and M. Esteller, CpG Islands in Cancer: Heads, Tails, and Sides. Methods Mol Biol, 2018. **1766**: p. 49-80.
82. Gowher, H. and A. Jeltsch, Mammalian DNA methyltransferases: new discoveries and open questions. Biochem Soc Trans, 2018. **46**(5): p. 1191-1202.

83. Gujar, H., D.J. Weisenberger, and G. Liang, The Roles of Human DNA Methyltransferases and Their Isoforms in Shaping the Epigenome. Genes (Basel), 2019. **10**(2).
84. Marisa, L., et al., Gene expression classification of colon cancer into molecular subtypes: characterization, validation, and prognostic value. PLoS Med, 2013. **10**(5): p. e1001453.
85. De Sousa, E.M.F., et al., Poor-prognosis colon cancer is defined by a molecularly distinct subtype and develops from serrated precursor lesions. Nat Med, 2013. **19**(5): p. 614-8.
86. Sadanandam, A., et al., A colorectal cancer classification system that associates cellular phenotype and responses to therapy. Nat Med, 2013. **19**(5): p. 619-25.
87. Schlicker, A., et al., Subtypes of primary colorectal tumors correlate with response to targeted treatment in colorectal cell lines. BMC Med Genomics, 2012. **5**: p. 66.
88. Budinska, E., et al., Gene expression patterns unveil a new level of molecular heterogeneity in colorectal cancer. J Pathol, 2013. **231**(1): p. 63-76.
89. Roepman, P., et al., Colorectal cancer intrinsic subtypes predict chemotherapy benefit, deficient mismatch repair and epithelial-to-mesenchymal transition. Int J Cancer, 2014. **134**(3): p. 552-62.
90. Mooi, J.K., et al., The prognostic impact of consensus molecular subtypes (CMS) and its predictive effects for bevacizumab benefit in metastatic colorectal cancer: molecular analysis of the AGITG MAX clinical trial. Ann Oncol, 2018. **29**(11): p. 2240-2246.
91. Buikhuisen, J.Y., A. Torang, and J.P. Medema, Exploring and modelling colon cancer inter-tumour heterogeneity: opportunities and challenges. Oncogenesis, 2020. **9**(7): p. 66.
92. Ezzeldin, H. and R. Diasio, Dihydropyrimidine dehydrogenase deficiency, a pharmacogenetic syndrome associated with potentially life-threatening toxicity following 5-fluorouracil administration. Clin Colorectal Cancer, 2004. **4**(3): p. 181-9.
93. Longley, D.B., D.P. Harkin, and P.G. Johnston, 5-fluorouracil: mechanisms of action and clinical strategies. Nat Rev Cancer, 2003. **3**(5): p. 330-8.
94. Pettersen, H.S., et al., UNG-initiated base excision repair is the major repair route for 5-fluorouracil in DNA, but 5-fluorouracil cytotoxicity depends mainly on RNA incorporation. Nucleic Acids Res, 2011. **39**(19): p. 8430-44.
95. Rosenberg, B., et al., Platinum compounds: a new class of potent antitumour agents. Nature, 1969. **222**(5191): p. 385-6.
96. Kidani, Y., et al., Antitumor activity of 1,2-diaminocyclohexane--platinum complexes against sarcoma-180 ascites form. J Med Chem, 1978. **21**(12): p. 1315-8.
97. Rixe, O., et al., Oxaliplatin, tetraplatin, cisplatin, and carboplatin: spectrum of activity in drug-resistant cell lines and in the cell lines of the National Cancer Institute's Anticancer Drug Screen panel. Biochem Pharmacol, 1996. **52**(12): p. 1855-65.
98. Wheate, N.J., et al., The status of platinum anticancer drugs in the clinic and in clinical trials. Dalton Trans, 2010. **39**(35): p. 8113-27.
99. Di Francesco, A.M., A. Ruggiero, and R. Riccardi, Cellular and molecular aspects of drugs of the future: oxaliplatin. Cell Mol Life Sci, 2002. **59**(11): p. 1914-27.
100. Woynarowski, J.M., et al., Sequence- and region-specificity of oxaliplatin adducts in naked and cellular DNA. Mol Pharmacol, 1998. **54**(5): p. 770-7.
101. Pendyala, L. and P.J. Creaven, In vitro cytotoxicity, protein binding, red blood cell partitioning, and biotransformation of oxaliplatin. Cancer Res, 1993. **53**(24): p. 5970-6.
102. Rabik, C.A. and M.E. Dolan, Molecular mechanisms of resistance and toxicity associated with platinating agents. Cancer Treat Rev, 2007. **33**(1): p. 9-23.
103. Deans, A.J. and S.C. West, DNA interstrand crosslink repair and cancer. Nat Rev Cancer, 2011. **11**(7): p. 467-80.
104. Reardon, J.T., et al., Efficient nucleotide excision repair of cisplatin, oxaliplatin, and Bis-aceto-ammine-dichloro-cyclohexylamine-platinum(IV) (JM216) platinum intrastrand DNA diadducts. Cancer Res, 1999. **59**(16): p. 3968-71.
105. Woynarowski, J.M., et al., Oxaliplatin-induced damage of cellular DNA. Mol Pharmacol, 2000. **58**(5): p. 920-7.
106. Basourakos, S.P., et al., Combination Platinum-based and DNA Damage Response-targeting Cancer Therapy: Evolution and Future Directions. Curr Med Chem, 2017. **24**(15): p. 1586-1606.
107. Rivory, L.P., et al., Conversion of irinotecan (CPT-11) to its active metabolite, 7-ethyl-10-hydroxycamptothecin (SN-38), by human liver carboxylesterase. Biochem Pharmacol, 1996. **52**(7): p. 1103-11.

108. Wu, M.H., et al., Irinotecan activation by human carboxylesterases in colorectal adenocarcinoma cells. Clin Cancer Res, 2002. **8**(8): p. 2696-700.
109. Xu, G., et al., Human carboxylesterase 2 is commonly expressed in tumor tissue and is correlated with activation of irinotecan. Clin Cancer Res, 2002. **8**(8): p. 2605-11.
110. Wiseman, L.R. and A. Markham, Irinotecan. A review of its pharmacological properties and clinical efficacy in the management of advanced colorectal cancer. Drugs, 1996. **52**(4): p. 606-23.
111. Stewart, L., et al., A model for the mechanism of human topoisomerase I. Science, 1998. **279**(5356): p. 1534-41.
112. Goldstein, M. and M.B. Kastan, The DNA damage response: implications for tumor responses to radiation and chemotherapy. Annu Rev Med, 2015. **66**: p. 129-43.
113. Chu, E., et al., Thymidylate synthase inhibitors as anticancer agents: from bench to bedside. Cancer Chemother Pharmacol, 2003. **52 Suppl 1**: p. S80-9.
114. Donne, R., et al., Replication stress triggered by nucleotide pool imbalance drives DNA damage and cGAS-STING pathway activation in NAFLD. Dev Cell, 2022. **57**(14): p. 1728-1741 e6.
115. Gralewska, P., et al., Participation of the ATR/CHK1 pathway in replicative stress targeted therapy of high-grade ovarian cancer. J Hematol Oncol, 2020. **13**(1): p. 39.
116. Krokan, H.E. and M. Bjoras, Base excision repair. Cold Spring Harb Perspect Biol, 2013. **5**(4): p. a012583.
117. Hegde, M.L., T.K. Hazra, and S. Mitra, Early steps in the DNA base excision/single-strand interruption repair pathway in mammalian cells. Cell Res, 2008. **18**(1): p. 27-47.
118. Beard, W.A. and S.H. Wilson, Structure and mechanism of DNA polymerase Beta. Chem Rev, 2006. **106**(2): p. 361-82.
119. Mojardin, L., et al., New insights into the RNA-based mechanism of action of the anticancer drug 5'-fluorouracil in eukaryotic cells. PLoS One, 2013. **8**(11): p. e78172.
120. Jung, Y. and S.J. Lippard, Direct cellular responses to platinum-induced DNA damage. Chem Rev, 2007. **107**(5): p. 1387-407.
121. Kartalou, M. and J.M. Essigmann, Recognition of cisplatin adducts by cellular proteins. Mutat Res, 2001. **478**(1-2): p. 1-21.
122. Jamieson, E.R. and S.J. Lippard, Structure, Recognition, and Processing of Cisplatin-DNA Adducts. Chem Rev, 1999. **99**(9): p. 2467-98.
123. Furuta, T., et al., Transcription-coupled nucleotide excision repair as a determinant of cisplatin sensitivity of human cells. Cancer Res, 2002. **62**(17): p. 4899-902.
124. Scharer, O.D., Nucleotide excision repair in eukaryotes. Cold Spring Harb Perspect Biol, 2013. **5**(10): p. a012609.
125. Su, X. and J. Huang, The Fanconi anemia pathway and DNA interstrand cross-link repair. Protein Cell, 2011. **2**(9): p. 704-11.
126. Chirnomas, D., et al., Chemosensitization to cisplatin by inhibitors of the Fanconi anemia/BRCA pathway. Mol Cancer Ther, 2006. **5**(4): p. 952-61.
127. Hashimoto, S., H. Anai, and K. Hanada, Mechanisms of interstrand DNA crosslink repair and human disorders. Genes Environ, 2016. **38**: p. 9.
128. Sarkar, S., et al., DNA interstrand crosslink repair during G1 involves nucleotide excision repair and DNA polymerase zeta. EMBO J, 2006. **25**(6): p. 1285-94.
129. Powers, K.T. and M.T. Washington, Eukaryotic translesion synthesis: Choosing the right tool for the job. DNA Repair (Amst), 2018. **71**: p. 127-134.
130. Knipscheer, P., et al., The Fanconi anemia pathway promotes replication-dependent DNA interstrand cross-link repair. Science, 2009. **326**(5960): p. 1698-701.
131. Long, D.T., et al., Mechanism of RAD51-dependent DNA interstrand cross-link repair. Science, 2011. **333**(6038): p. 84-7.
132. Walden, H. and A.J. Deans, The Fanconi anemia DNA repair pathway: structural and functional insights into a complex disorder. Annu Rev Biophys, 2014. **43**: p. 257-78.
133. Zhu, G., et al., Monofunctional platinum-DNA adducts are strong inhibitors of transcription and substrates for nucleotide excision repair in live mammalian cells. Cancer Res, 2012. **72**(3): p. 790-800.
134. Ming, X., et al., Mass Spectrometry Based Proteomics Study of Cisplatin-Induced DNA-Protein Cross-Linking in Human Fibrosarcoma (HT1080) Cells. Chem Res Toxicol, 2017. **30**(4): p. 980-995.
135. Saha, L.K., et al., Replication-dependent cytotoxicity and Spartan-mediated repair of trapped PARP1-DNA complexes. Nucleic Acids Res, 2021. **49**(18): p. 10493-10506.

136. Stingele, J., et al., Mechanism and Regulation of DNA-Protein Crosslink Repair by the DNA-Dependent Metalloprotease SPRTN. Mol Cell, 2016. **64**(4): p. 688-703.
137. Alagoz, M., O.S. Wells, and S.F. El-Khamisy, TDP1 deficiency sensitizes human cells to base damage via distinct topoisomerase I and PARP mechanisms with potential applications for cancer therapy. Nucleic Acids Res, 2014. **42**(5): p. 3089-103.
138. Meisenberg, C., et al., Clinical and cellular roles for TDP1 and TOP1 in modulating colorectal cancer response to irinotecan. Mol Cancer Ther, 2015. **14**(2): p. 575-85.
139. Lin, C.P., et al., A ubiquitin-proteasome pathway for the repair of topoisomerase I-DNA covalent complexes. J Biol Chem, 2008. **283**(30): p. 21074-83.
140. Fielden, J., et al., TEX264 coordinates p97- and SPRTN-mediated resolution of topoisomerase 1-DNA adducts. Nat Commun, 2020. **11**(1): p. 1274.
141. Vaz, B., M. Popovic, and K. Ramadan, DNA-Protein Crosslink Proteolysis Repair. Trends Biochem Sci, 2017. **42**(6): p. 483-495.
142. Stingele, J., R. Bellelli, and S.J. Boulton, Mechanisms of DNA-protein crosslink repair. Nat Rev Mol Cell Biol, 2017. **18**(9): p. 563-573.
143. Ruggiano, A. and K. Ramadan, DNA-protein crosslink proteases in genome stability. Commun Biol, 2021. **4**(1): p. 11.
144. Tretyakova, N.Y., A.t. Groehler, and S. Ji, DNA-Protein Cross-Links: Formation, Structural Identities, and Biological Outcomes. Acc Chem Res, 2015. **48**(6): p. 1631-44.
145. Hoa, N.N., et al., Mre11 Is Essential for the Removal of Lethal Topoisomerase 2 Covalent Cleavage Complexes. Mol Cell, 2016. **64**(3): p. 580-592.
146. Lee, K.C., et al., MRE11 facilitates the removal of human topoisomerase II complexes from genomic DNA. Biol Open, 2012. **1**(9): p. 863-73.
147. Cannavo, E. and P. Cejka, Sae2 promotes dsDNA endonuclease activity within Mre11-Rad50-Xrs2 to resect DNA breaks. Nature, 2014. **514**(7520): p. 122-5.
148. Deshpande, R.A., et al., Nbs1 Converts the Human Mre11/Rad50 Nuclease Complex into an Endo/Exonuclease Machine Specific for Protein-DNA Adducts. Mol Cell, 2016. **64**(3): p. 593-606.
149. Yang, S.W., et al., A eukaryotic enzyme that can disjoin dead-end covalent complexes between DNA and type I topoisomerases. Proc Natl Acad Sci U S A, 1996. **93**(21): p. 11534-9.
150. Pouliot, J.J., et al., Yeast gene for a Tyr-DNA phosphodiesterase that repairs topoisomerase I complexes. Science, 1999. **286**(5439): p. 552-5.
151. Debethune, L., et al., Processing of nucleopeptides mimicking the topoisomerase I-DNA covalent complex by tyrosyl-DNA phosphodiesterase. Nucleic Acids Res, 2002. **30**(5): p. 1198-204.
152. Cortes Ledesma, F., et al., A human 5'-tyrosyl DNA phosphodiesterase that repairs topoisomerase-mediated DNA damage. Nature, 2009. **461**(7264): p. 674-8.
153. Gao, R., et al., Proteolytic degradation of topoisomerase II (Top2) enables the processing of Top2.DNA and Top2.RNA covalent complexes by tyrosyl-DNA-phosphodiesterase 2 (TDP2). J Biol Chem, 2014. **289**(26): p. 17960-9.
154. Wang, J.C., Cellular roles of DNA topoisomerases: a molecular perspective. Nat Rev Mol Cell Biol, 2002. **3**(6): p. 430-40.
155. Kumar, S. and M.Y. Sherman, Resistance to TOP-1 Inhibitors: Good Old Drugs Still Can Surprise Us. Int J Mol Sci, 2023. **24**(8).
156. El-Khamisy, S.F., To live or to die: a matter of processing damaged DNA termini in neurons. EMBO Mol Med, 2011. **3**(2): p. 78-88.
157. Interthal, H. and J.J. Champoux, Effects of DNA and protein size on substrate cleavage by human tyrosyl-DNA phosphodiesterase 1. Biochem J, 2011. **436**(3): p. 559-66.
158. Vaz, B., et al., Metalloprotease SPRTN/DVC1 Orchestrates Replication-Coupled DNA-Protein Crosslink Repair. Mol Cell, 2016. **64**(4): p. 704-719.
159. Morocz, M., et al., DNA-dependent protease activity of human Spartan facilitates replication of DNA-protein crosslink-containing DNA. Nucleic Acids Res, 2017. **45**(6): p. 3172-3188.
160. Maskey, R.S., et al., Spartan deficiency causes accumulation of Topoisomerase 1 cleavage complexes and tumorigenesis. Nucleic Acids Res, 2017. **45**(8): p. 4564-4576.
161. Stingele, J., et al., A DNA-dependent protease involved in DNA-protein crosslink repair. Cell, 2014. **158**(2): p. 327-338.
162. Larsen, N.B., et al., Replication-Coupled DNA-Protein Crosslink Repair by SPRTN and the Proteasome in Xenopus Egg Extracts. Mol Cell, 2019. **73**(3): p. 574-588 e7.

163. Duxin, J.P., et al., Repair of a DNA-protein crosslink by replication-coupled proteolysis. Cell, 2014. **159**(2): p. 346-57.
164. Lopez-Mosqueda, J., et al., SPRTN is a mammalian DNA-binding metalloprotease that resolves DNA-protein crosslinks. Elife, 2016. **5**.
165. Lessel, D., et al., Mutations in SPRTN cause early onset hepatocellular carcinoma, genomic instability and progeroid features. Nat Genet, 2014. **46**(11): p. 1239-44.
166. Ruijs, M.W., et al., Atypical progeroid syndrome: an unknown helicase gene defect? Am J Med Genet A, 2003. **116A**(3): p. 295-9.
167. Nakano, T., et al., Translocation and stability of replicative DNA helicases upon encountering DNA-protein cross-links. J Biol Chem, 2013. **288**(7): p. 4649-58.
168. Baker, D.J., et al., Nucleotide excision repair eliminates unique DNA-protein cross-links from mammalian cells. J Biol Chem, 2007. **282**(31): p. 22592-604.
169. Quievryn, G. and A. Zhitkovich, Loss of DNA-protein crosslinks from formaldehyde-exposed cells occurs through spontaneous hydrolysis and an active repair process linked to proteosome function. Carcinogenesis, 2000. **21**(8): p. 1573-80.
170. Borgermann, N., et al., SUMOylation promotes protective responses to DNA-protein crosslinks. EMBO J, 2019. **38**(8).
171. Sun, Y., et al., A conserved SUMO pathway repairs topoisomerase DNA-protein cross-links by engaging ubiquitin-mediated proteasomal degradation. Sci Adv, 2020. **6**(46).
172. Ruggiano, A., et al., The protease SPRTN and SUMOylation coordinate DNA-protein crosslink repair to prevent genome instability. Cell Rep, 2021. **37**(10): p. 110080.
173. Huang, J., et al., Tandem Deubiquitination and Acetylation of SPRTN Promotes DNA-Protein Crosslink Repair and Protects against Aging. Mol Cell, 2020. **79**(5): p. 824-835 e5.
174. Reinking, H.K., et al., DNA Structure-Specific Cleavage of DNA-Protein Crosslinks by the SPRTN Protease. Mol Cell, 2020. **80**(1): p. 102-113 e6.
175. Perry, M. and G. Ghosal, Mechanisms and Regulation of DNA-Protein Crosslink Repair During DNA Replication by SPRTN Protease. Front Mol Biosci, 2022. **9**: p. 916697.
176. Srivastava, M., et al., Replisome Dynamics and Their Functional Relevance upon DNA Damage through the PCNA Interactome. Cell Rep, 2018. **25**(13): p. 3869-3883 e4.
177. Jones, M.L., et al., Structure of a human replisome shows the organisation and interactions of a DNA replication machine. EMBO J, 2021. **40**(23): p. e108819.
178. Kojima, Y., et al., FAM111A protects replication forks from protein obstacles via its trypsin-like domain. Nat Commun, 2020. **11**(1): p. 1318.
179. Hoffmann, S., et al., FAM111 protease activity undermines cellular fitness and is amplified by gain-of-function mutations in human disease. EMBO Rep, 2020. **21**(10): p. e50662.
180. Serbyn, N., et al., The Aspartic Protease Ddi1 Contributes to DNA-Protein Crosslink Repair in Yeast. Mol Cell, 2020. **77**(5): p. 1066-1079 e9.
181. Dokshin, G.A., et al., GCNA Interacts with Spartan and Topoisomerase II to Regulate Genome Stability. Dev Cell, 2020. **52**(1): p. 53-68 e6.
182. Cecile Otten, C.S.-P., Marin Kutnjak, Ivan Anticevic, Vanna Medved, Marta Popovic, ACRC/GCNA is an essential protease for the repair of DNA-protein crosslinks during vertebrate development. 2023.
183. Welter, A.L. and Y.J. Machida, Functions and evolution of FAM111 serine proteases. Front Mol Biosci, 2022. **9**: p. 1081166.
184. Arowolo, A., et al., Proposed Cellular Function of the Human FAM111B Protein and Dysregulation in Fibrosis and Cancer. Front Oncol, 2022. **12**: p. 932167.
185. Aviner, R., et al., Uncovering Hidden Layers of Cell Cycle Regulation through Integrative Multi-omic Analysis. PLoS Genet, 2015. **11**(10): p. e1005554.
186. Kawasaki, K., et al., FAM111B enhances proliferation of KRAS-driven lung adenocarcinoma by degrading p16. Cancer Sci, 2020. **111**(7): p. 2635-2646.
187. Yip, M.C.J., N.O. Bodnar, and T.A. Rapoport, Ddi1 is a ubiquitin-dependent protease. Proc Natl Acad Sci U S A, 2020. **117**(14): p. 7776-7781.
188. Leszczynska, K.B., et al., Hypoxia-induced p53 modulates both apoptosis and radiosensitivity via AKT. J Clin Invest, 2015. **125**(6): p. 2385-98.
189. Nakamura, K., et al., H4K20me0 recognition by BRCA1-BARD1 directs homologous recombination to sister chromatids. Nat Cell Biol, 2019. **21**(3): p. 311-318.
190. Molla Kazemiha, V., et al., Sensitivity of biochemical test in comparison with other methods for the detection of mycoplasma contamination in human and animal cell lines stored in the National Cell Bank of Iran. Cytotechnology, 2014. **66**(5): p. 861-73.

191. Birmingham, A., et al., A protocol for designing siRNAs with high functionality and specificity. Nat Protoc, 2007. **2**(9): p. 2068-78.
192. Franken, N.A., et al., Clonogenic assay of cells in vitro. Nat Protoc, 2006. **1**(5): p. 2315-9.
193. Wadhwa, R.R. and R. Marappa-Ganeshan, T Test, in StatPearls. 2023: Treasure Island (FL).
194. Walker, J.M., The bicinchoninic acid (BCA) assay for protein quantitation. Methods Mol Biol, 1994. **32**: p. 5-8.
195. Mahmood, T. and P.C. Yang, Western blot: technique, theory, and trouble shooting. N Am J Med Sci, 2012. **4**(9): p. 429-34.
196. Nolan, T., R.E. Hands, and S.A. Bustin, Quantification of mRNA using real-time RT-PCR. Nat Protoc, 2006. **1**(3): p. 1559-82.
197. Adhikari, B., A. Narain, and E. Wolf, Generation of auxin inducible degron (AID) knock-in cell lines for targeted protein degradation in mammalian cells. STAR Protoc, 2021. **2**(4): p. 100949.
198. Hume, S., et al., The NUCKS1-SKP2-p21/p27 axis controls S phase entry. Nat Commun, 2021. **12**(1): p. 6959.
199. Popp, H.D., et al., Immunofluorescence Microscopy of gammaH2AX and 53BP1 for Analyzing the Formation and Repair of DNA Double-strand Breaks. J Vis Exp, 2017(129).
200. Mishra, P., et al., Application of student's t-test, analysis of variance, and covariance. Ann Card Anaesth, 2019. **22**(4): p. 407-411.
201. Wang, Z., M. Gerstein, and M. Snyder, RNA-Seq: a revolutionary tool for transcriptomics. Nat Rev Genet, 2009. **10**(1): p. 57-63.
202. Koch, C.M., et al., A Beginner's Guide to Analysis of RNA Sequencing Data. Am J Respir Cell Mol Biol, 2018. **59**(2): p. 145-157.
203. Hrdlickova, R., M. Toloue, and B. Tian, RNA-Seq methods for transcriptome analysis. Wiley Interdiscip Rev RNA, 2017. **8**(1).
204. Seymour, M.T., et al., Different strategies of sequential and combination chemotherapy for patients with poor prognosis advanced colorectal cancer (MRC FOCUS): a randomised controlled trial. Lancet, 2007. **370**(9582): p. 143-152.
205. Tsutsumi, Y., Pitfalls and Caveats in Applying Chromogenic Immunostaining to Histopathological Diagnosis. Cells, 2021. **10**(6).
206. Magaki, S., et al., An Introduction to the Performance of Immunohistochemistry. Methods Mol Biol, 2019. **1897**: p. 289-298.
207. Horai, Y., et al., Quantification of histopathological findings using a novel image analysis platform. J Toxicol Pathol, 2019. **32**(4): p. 319-327.
208. Niazi, M.K.K., A.V. Parwani, and M.N. Gurcan, Digital pathology and artificial intelligence. Lancet Oncol, 2019. **20**(5): p. e253-e261.
209. Therasse, P., et al., New guidelines to evaluate the response to treatment in solid tumors. European Organization for Research and Treatment of Cancer, National Cancer Institute of the United States, National Cancer Institute of Canada. J Natl Cancer Inst, 2000. **92**(3): p. 205-16.
210. Tsherniak, A., et al., Defining a Cancer Dependency Map. Cell, 2017. **170**(3): p. 564-576 e16.
211. Dempster, J.M., et al., Chronos: a cell population dynamics model of CRISPR experiments that improves inference of gene fitness effects. Genome Biol, 2021. **22**(1): p. 343.
212. Cancer Genome Atlas Research, N., et al., The Cancer Genome Atlas Pan-Cancer analysis project. Nat Genet, 2013. **45**(10): p. 1113-20.
213. Draghici, S., et al., Reliability and reproducibility issues in DNA microarray measurements. Trends Genet, 2006. **22**(2): p. 101-9.
214. Peach, M., N. Marsh, and D.J. Macphee, Protein solubilization: attend to the choice of lysis buffer. Methods Mol Biol, 2012. **869**: p. 37-47.
215. Pommier, Y., Topoisomerase I inhibitors: camptothecins and beyond. Nat Rev Cancer, 2006. **6**(10): p. 789-802.
216. Chvalova, K., V. Brabec, and J. Kasparkova, Mechanism of the formation of DNA-protein cross-links by antitumor cisplatin. Nucleic Acids Res, 2007. **35**(6): p. 1812-21.
217. Toscano, F., et al., p53 dependent and independent sensitivity to oxaliplatin of colon cancer cells. Biochem Pharmacol, 2007. **74**(3): p. 392-406.

218. Manic, S., et al., Mechanisms controlling sensitivity to platinum complexes: role of p53 and DNA mismatch repair. Curr Cancer Drug Targets, 2003. **3**(1): p. 21-9.
219. Qiu, S., C.M. Adema, and T. Lane, A computational study of off-target effects of RNA interference. Nucleic Acids Res, 2005. **33**(6): p. 1834-47.
220. Maskey, R.S., et al., Spartan deficiency causes genomic instability and progeroid phenotypes. Nat Commun, 2014. **5**: p. 5744.
221. Nishimura, K., et al., An auxin-based degron system for the rapid depletion of proteins in nonplant cells. Nat Methods, 2009. **6**(12): p. 917-22.
222. Bekes, M., D.R. Langley, and C.M. Crews, PROTAC targeted protein degraders: the past is prologue. Nat Rev Drug Discov, 2022. **21**(3): p. 181-200.
223. Quint, M. and W.M. Gray, Auxin signaling. Curr Opin Plant Biol, 2006. **9**(5): p. 448-53.
224. Shetty, A., N.I. Reim, and F. Winston, Auxin-Inducible Degron System for Depletion of Proteins in Saccharomyces cerevisiae. Curr Protoc Mol Biol, 2019. **128**(1): p. e104.
225. Natsume, T., et al., Rapid Protein Depletion in Human Cells by Auxin-Inducible Degron Tagging with Short Homology Donors. Cell Rep, 2016. **15**(1): p. 210-218.
226. Ran, F.A., et al., Genome engineering using the CRISPR-Cas9 system. Nat Protoc, 2013. **8**(11): p. 2281-2308.
227. Noordermeer, S.M., et al., The shieldin complex mediates 53BP1-dependent DNA repair. Nature, 2018. **560**(7716): p. 117-121.
228. Panier, S. and S.J. Boulton, Double-strand break repair: 53BP1 comes into focus. Nat Rev Mol Cell Biol, 2014. **15**(1): p. 7-18.
229. Sharma, A., K. Singh, and A. Almasan, Histone H2AX phosphorylation: a marker for DNA damage. Methods Mol Biol, 2012. **920**: p. 613-26.
230. Mah, L.J., A. El-Osta, and T.C. Karagiannis, gammaH2AX: a sensitive molecular marker of DNA damage and repair. Leukemia, 2010. **24**(4): p. 679-86.
231. Shibata, A. and P.A. Jeggo, Roles for 53BP1 in the repair of radiation-induced DNA double strand breaks. DNA Repair (Amst), 2020. **93**: p. 102915.
232. Scully, R., et al., DNA double-strand break repair-pathway choice in somatic mammalian cells. Nat Rev Mol Cell Biol, 2019. **20**(11): p. 698-714.
233. Krenning, L., J. van den Berg, and R.H. Medema, Life or Death after a Break: What Determines the Choice? Mol Cell, 2019. **76**(2): p. 346-358.
234. Halder, S., et al., SPRTN protease and checkpoint kinase 1 cross-activation loop safeguards DNA replication. Nat Commun, 2019. **10**(1): p. 3142.
235. Smith, H.L., et al., DNA damage checkpoint kinases in cancer. Expert Rev Mol Med, 2020. **22**: p. e2.
236. Vera, J., et al., Chk1 and Wee1 control genotoxic-stress induced G2-M arrest in melanoma cells. Cell Signal, 2015. **27**(5): p. 951-60.
237. Barnum, K.J. and M.J. O'Connell, Cell cycle regulation by checkpoints. Methods Mol Biol, 2014. **1170**: p. 29-40.
238. Mosbech, A., et al., DVC1 (C1orf124) is a DNA damage-targeting p97 adaptor that promotes ubiquitin-dependent responses to replication blocks. Nat Struct Mol Biol, 2012. **19**(11): p. 1084-92.
239. Li, M. and P. Zhang, The function of APC/CCdh1 in cell cycle and beyond. Cell Div, 2009. **4**: p. 2.
240. Kulahoglu, C. and A. Brautigam, Quantitative transcriptome analysis using RNA-seq. Methods Mol Biol, 2014. **1158**: p. 71-91.
241. Hong, M., et al., RNA sequencing: new technologies and applications in cancer research. J Hematol Oncol, 2020. **13**(1): p. 166.
242. Powell, D.N., et al., Indoles from the commensal microbiota act via the AHR and IL-10 to tune the cellular composition of the colonic epithelium during aging. Proc Natl Acad Sci U S A, 2020. **117**(35): p. 21519-21526.
243. Heath-Pagliuso, S., et al., Activation of the Ah receptor by tryptophan and tryptophan metabolites. Biochemistry, 1998. **37**(33): p. 11508-15.
244. MacPherson, L., et al., 2,3,7,8-Tetrachlorodibenzo-p-dioxin poly(ADP-ribose) polymerase (TiPARP, ARTD14) is a mono-ADP-ribosyltransferase and repressor of aryl hydrocarbon receptor transactivation. Nucleic Acids Res, 2013. **41**(3): p. 1604-21.
245. Gomez, A., et al., Characterization of TCDD-inducible poly-ADP-ribose polymerase (TIPARP/ARTD14) catalytic activity. Biochem J, 2018. **475**(23): p. 3827-3846.
246. Perez-Castro, L., et al., The AHR target gene scinderin activates the WNT pathway by facilitating the nuclear translocation of beta-catenin. J Cell Sci, 2022. **135**(20).

247. Goel, S., J.S. Bergholz, and J.J. Zhao, Targeting CDK4 and CDK6 in cancer. Nat Rev Cancer, 2022. **22**(6): p. 356-372.
248. Meyerson, M. and E. Harlow, Identification of G1 kinase activity for cdk6, a novel cyclin D partner. Mol Cell Biol, 1994. **14**(3): p. 2077-86.
249. Sellers, W.R., J.W. Rodgers, and W.G. Kaelin, Jr., A potent transrepression domain in the retinoblastoma protein induces a cell cycle arrest when bound to E2F sites. Proc Natl Acad Sci U S A, 1995. **92**(25): p. 11544-8.
250. Villacanas, O., J.J. Perez, and J. Rubio-Martinez, Structural analysis of the inhibition of Cdk4 and Cdk6 by p16(INK4a) through molecular dynamics simulations. J Biomol Struct Dyn, 2002. **20**(3): p. 347-58.
251. Bates, S., et al., p14ARF links the tumour suppressors RB and p53. Nature, 1998. **395**(6698): p. 124-5.
252. Pomerantz, J., et al., The Ink4a tumor suppressor gene product, p19Arf, interacts with MDM2 and neutralizes MDM2's inhibition of p53. Cell, 1998. **92**(6): p. 713-23.
253. Kastan, M.B., C.E. Canman, and C.J. Leonard, P53, cell cycle control and apoptosis: implications for cancer. Cancer Metastasis Rev, 1995. **14**(1): p. 3-15.
254. Bertoli, C., J.M. Skotheim, and R.A. de Bruin, Control of cell cycle transcription during G1 and S phases. Nat Rev Mol Cell Biol, 2013. **14**(8): p. 518-28.
255. Senturk, E. and J.J. Manfredi, p53 and cell cycle effects after DNA damage. Methods Mol Biol, 2013. **962**: p. 49-61.
256. Karlsson, C., et al., Cdc25B and Cdc25C differ markedly in their properties as initiators of mitosis. J Cell Biol, 1999. **146**(3): p. 573-84.
257. Santamaria, D., et al., Cdk1 is sufficient to drive the mammalian cell cycle. Nature, 2007. **448**(7155): p. 811-5.
258. Timofeev, O., et al., Cdc25 phosphatases are required for timely assembly of CDK1-cyclin B at the G2/M transition. J Biol Chem, 2010. **285**(22): p. 16978-90.
259. Pelletier, J., G. Thomas, and S. Volarevic, Ribosome biogenesis in cancer: new players and therapeutic avenues. Nat Rev Cancer, 2018. **18**(1): p. 51-63.
260. Petibon, C., et al., Regulation of ribosomal protein genes: An ordered anarchy. Wiley Interdiscip Rev RNA, 2021. **12**(3): p. e1632.
261. Kanehisa, M. and S. Goto, KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res, 2000. **28**(1): p. 27-30.
262. Kanehisa, M., et al., KEGG: new perspectives on genomes, pathways, diseases and drugs. Nucleic Acids Res, 2017. **45**(D1): p. D353-D361.
263. K. Ramadan, J.F., T. Maughan, US20220381783 - BIOMARKER, in https://patentscope.wipo.int/search/en/detail.jsf?docId=US380160472&_fid=WO2021023973. 2020, Oxford University Innovation Limited: United States of America.
264. Delgado, A. and A.K. Guddati, Clinical endpoints in oncology - a primer. Am J Cancer Res, 2021. **11**(4): p. 1121-1131.
265. Paschke, S., et al., Are Colon and Rectal Cancer Two Different Tumor Entities? A Proposal to Abandon the Term Colorectal Cancer. Int J Mol Sci, 2018. **19**(9).
266. Yamauchi, M., et al., Assessment of colorectal cancer molecular features along bowel subsites challenges the conception of distinct dichotomy of proximal versus distal colorectum. Gut, 2012. **61**(6): p. 847-54.
267. Hiom, K., SPRTN is a new player in an old story. Nat Genet, 2014. **46**(11): p. 1155-7.
268. Montecucco, A., F. Zanetta, and G. Biamonti, Molecular mechanisms of etoposide. EXCLI J, 2015. **14**: p. 95-108.
269. Okamoto, H., et al., Randomised phase III trial of carboplatin plus etoposide vs split doses of cisplatin plus etoposide in elderly or poor-risk patients with extensive disease small-cell lung cancer: JCOG 9702. Br J Cancer, 2007. **97**(2): p. 162-9.
270. Nebenfuehr, S., K. Kollmann, and V. Sexl, The role of CDK6 in cancer. Int J Cancer, 2020. **147**(11): p. 2988-2995.
271. Li, J., M.J. Poi, and M.D. Tsai, Regulatory mechanisms of tumor suppressor P16(INK4A) and their relevance to cancer. Biochemistry, 2011. **50**(25): p. 5566-82.
272. Ozenne, P., et al., The ARF tumor suppressor: structure, functions and status in cancer. Int J Cancer, 2010. **127**(10): p. 2239-47.
273. Nilsson, I. and I. Hoffmann, Cell cycle regulation by the Cdc25 phosphatase family. Prog Cell Cycle Res, 2000. **4**: p. 107-14.
274. Lammer, C., et al., The cdc25B phosphatase is essential for the G2/M phase transition in human cells. J Cell Sci, 1998. **111 (Pt 16)**: p. 2445-53.

275. Liao, H., F. Ji, and S. Ying, CDK1: beyond cell cycle regulation. Aging (Albany NY), 2017. **9**(12): p. 2465-2466.
276. Lee, J.H. and T.T. Paull, Activation and regulation of ATM kinase activity in response to DNA double-strand breaks. Oncogene, 2007. **26**(56): p. 7741-8.
277. Hartman, M.L. and M. Czyz, BCL-w: apoptotic and non-apoptotic role in health and disease. Cell Death Dis, 2020. **11**(4): p. 260.
278. Lu, C., Z. Xie, and Q. Peng, MiRNA-107 enhances chemosensitivity to paclitaxel by targeting antiapoptotic factor Bcl-w in non small cell lung cancer. Am J Cancer Res, 2017. **7**(9): p. 1863-1873.
279. Fang, L., et al., Cyramza induces apoptosis of HCC4006 cell by affecting the level of Bcl-w. Eur Rev Med Pharmacol Sci, 2017. **21**(13): p. 3069-3074.
280. Pan, X., et al., Induction of SOX4 by DNA damage is critical for p53 stabilization and function. Proc Natl Acad Sci U S A, 2009. **106**(10): p. 3788-93.
281. Py, B., et al., Siva-1 and an alternative splice form lacking the death domain, Siva-2, similarly induce apoptosis in T lymphocytes via a caspase-dependent mitochondrial pathway. J Immunol, 2004. **172**(7): p. 4008-17.
282. Xue, L., et al., Siva-1 binds to and inhibits BCL-X(L)-mediated protection against UV radiation-induced apoptosis. Proc Natl Acad Sci U S A, 2002. **99**(10): p. 6925-30.
283. Marusyk, A. and K. Polyak, Tumor heterogeneity: causes and consequences. Biochim Biophys Acta, 2010. **1805**(1): p. 105-17.
284. Turajlic, S., et al., Resolving genetic heterogeneity in cancer. Nat Rev Genet, 2019. **20**(7): p. 404-416.
285. Arneth, B., Tumor Microenvironment. Medicina (Kaunas), 2019. **56**(1).
286. Xiao, Y. and D. Yu, Tumor microenvironment as a therapeutic target in cancer. Pharmacol Ther, 2021. **221**: p. 107753.
287. Oldenhuis, C.N., et al., Prognostic versus predictive value of biomarkers in oncology. Eur J Cancer, 2008. **44**(7): p. 946-53.
288. Van Schaeybroeck, S., et al., Implementing prognostic and predictive biomarkers in CRC clinical trials. Nat Rev Clin Oncol, 2011. **8**(4): p. 222-32.
289. Zhang, L. and J. Yu, Role of apoptosis in colon cancer biology, therapy, and prevention. Curr Colorectal Cancer Rep, 2013. **9**(4).
290. Takayama, T., et al., Colorectal cancer: genetics of development and metastasis. J Gastroenterol, 2006. **41**(3): p. 185-92.
291. Aubrey, B.J., et al., How does p53 induce apoptosis and how does this relate to p53-mediated tumour suppression? Cell Death Differ, 2018. **25**(1): p. 104-113.
292. Ozaki, T. and A. Nakagawara, Role of p53 in Cell Death and Human Cancers. Cancers (Basel), 2011. **3**(1): p. 994-1013.
293. Yu, J., et al., PUMA induces the rapid apoptosis of colorectal cancer cells. Mol Cell, 2001. **7**(3): p. 673-82.
294. abcam. p53 Transcription Factor Assay Kit (Colorimetric) (ab207225).
295. Signaling, C. Bcl-w (31H4) Rabbit mAb #2724.
296. Gurtu, V., S.R. Kain, and G. Zhang, Fluorometric and colorimetric detection of caspase activity associated with apoptosis. Anal Biochem, 1997. **251**(1): p. 98-102.
297. McStay, G.P. and D.R. Green, Measuring apoptosis: caspase inhibitors and activity assays. Cold Spring Harb Protoc, 2014. **2014**(8): p. 799-806.
298. van Engeland, M., et al., Annexin V-affinity assay: a review on an apoptosis detection system based on phosphatidylserine exposure. Cytometry, 1998. **31**(1): p. 1-9.
299. O'Neil, D., H. Glowatz, and M. Schlumpberger, Ribosomal RNA depletion for efficient use of RNA-seq capacity. Curr Protoc Mol Biol, 2013. **Chapter 4**: p. Unit 4 19.
300. Klages-Mundt, N.L. and L. Li, Formation and repair of DNA-protein crosslink damage. Sci China Life Sci, 2017. **60**(10): p. 1065-1076.
301. Schofer, C. and K. Weipoltshammer, Nucleolus and chromatin. Histochem Cell Biol, 2018. **150**(3): p. 209-225.
302. Hernandez-Verdun, D., et al., The nucleolus: structure/function relationship in RNA metabolism. Wiley Interdiscip Rev RNA, 2010. **1**(3): p. 415-31.
303. Busch, H. and Y. Daskal, Methods for isolation of nuclei and nucleoli. Methods Cell Biol, 1977. **16**: p. 1-43.
304. Pendle, A.F. and P.J. Shaw, Isolation of Nuclei and Nucleoli. Methods Mol Biol, 2017. **1511**: p. 31-44.

305. Kiianitsa, K. and N. Maizels, A rapid and sensitive assay for DNA-protein covalent complexes in living cells. Nucleic Acids Res, 2013. **41**(9): p. e104.
306. Hu, Q., et al., The ARK Assay Is a Sensitive and Versatile Method for the Global Detection of DNA-Protein Crosslinks. Cell Rep, 2020. **30**(4): p. 1235-1245 e4.
307. Esposito, G., et al., Immunotherapy in Small Cell Lung Cancer. Cancers (Basel), 2020. **12**(9).
308. Jain, M., et al., Metabolite profiling identifies a key role for glycine in rapid cancer cell proliferation. Science, 2012. **336**(6084): p. 1040-4.
309. Cottu, P.H., et al., Inverse correlation between RER+ status and p53 mutation in colorectal cancer cell lines. Oncogene, 1996. **13**(12): p. 2727-30.
310. Fichtner, M., et al., Molecular subtype-specific responses of colon cancer cells to the SMAC mimetic Birinapant. Cell Death Dis, 2020. **11**(11): p. 1020.
311. Yang, M., et al., An integrative gene expression signature analysis identifies CMS4 KRAS-mutated colorectal cancers sensitive to combined MEK and SRC targeted therapy. BMC Cancer, 2022. **22**(1): p. 256.
312. Ikediobi, O.N., et al., Mutation analysis of 24 known cancer genes in the NCI-60 cell line set. Mol Cancer Ther, 2006. **5**(11): p. 2606-12.
313. Ilyas, M., et al., Beta-catenin mutations in cell lines established from human colorectal cancers. Proc Natl Acad Sci U S A, 1997. **94**(19): p. 10330-4.
314. Gayther, S.A., et al., Mutations truncating the EP300 acetylase in human cancers. Nat Genet, 2000. **24**(3): p. 300-3.
315. Janakiraman, M., et al., Genomic and biological characterization of exon 4 KRAS mutations in human cancer. Cancer Res, 2010. **70**(14): p. 5901-11.
316. Berg, K.C.G., et al., Multi-omics of 34 colorectal cancer cell lines - a resource for biomedical studies. Mol Cancer, 2017. **16**(1): p. 116.
317. Hempen, P.M., et al., Evidence of selection for clones having genetic inactivation of the activin A type II receptor (ACVR2) gene in gastrointestinal cancers. Cancer Res, 2003. **63**(5): p. 994-9.
318. Eshleman, J.R., et al., Chromosome number and structure both are markedly stable in RER colorectal cancers and are not destabilized by mutation of p53. Oncogene, 1998. **17**(6): p. 719-25.
319. Cowley, G.S., et al., Parallel genome-scale loss of function screens in 216 cancer cell lines for the identification of context-specific genetic dependencies. Sci Data, 2014. **1**: p. 140035.
320. Medico, E., et al., The molecular landscape of colorectal cancer cell lines unveils clinically actionable kinase targets. Nat Commun, 2015. **6**: p. 7002.
321. Linnekamp, J.F., et al., Consensus molecular subtypes of colorectal cancer are recapitulated in in vitro and in vivo models. Cell Death Differ, 2018. **25**(3): p. 616-633.
322. Mouradov, D., et al., Colorectal cancer cell lines are representative models of the main molecular subtypes of primary cancer. Cancer Res, 2014. **74**(12): p. 3238-47.
323. Liu, Y. and W.F. Bodmer, Analysis of P53 mutations and their expression in 56 colorectal cancer cell lines. Proc Natl Acad Sci U S A, 2006. **103**(4): p. 976-81.
324. Ahmed, D., et al., Epigenetic and genetic features of 24 colon cancer cell lines. Oncogenesis, 2013. **2**: p. e71.
325. Kalu, N.N., et al., Genomic characterization of human papillomavirus-positive and -negative human squamous cell cancer cell lines. Oncotarget, 2017. **8**(49): p. 86369-86383.
326. Seitz, S., et al., Identification of microsatellite instability and mismatch repair gene mutations in breast cancer cell lines. Genes Chromosomes Cancer, 2003. **37**(1): p. 29-35.
327. Macville, M., et al., Comprehensive and definitive molecular cytogenetic characterization of HeLa cells by spectral karyotyping. Cancer Res, 1999. **59**(1): p. 141-50.
328. Landry, J.J., et al., The genomic and transcriptomic landscape of a HeLa cell line. G3 (Bethesda), 2013. **3**(8): p. 1213-24.