Stratifiers for Oxaliplatin Outcome in Colorectal Cancer

Sylvana Hassanieh
University College
University of Oxford

Thesis submitted to the University of Oxford in partial fulfilment of the requirements for the degree of Doctor of Philosophy.

Michaelmas Term 2021

Supervisors

Prof. Tim Maughan (University of Oxford)
Dr. Enric Domingo (University of Oxford)

The work presented in this thesis was performed at the Department of Oncology, Old Road Campus Research Building (ORCRB), University of Oxford. This study was funded by Cancer Research UK prize studentship and Clarendon scholarship.

Declaration

I confirm that the work presented in this thesis is my own unless indicated. Whenever I have presented collaborated work and received technical help, I confirm that this has been credited in the thesis.

Dedicated to my amazing mother and partner,

as their support made me reach this place in life

Abstract

Stratifiers for Oxaliplatin Outcome in Colorectal Cancer

A thesis submitted to the University of Oxford
for the degree of Doctor of Philosophy

Sylvana Hassanieh, University College, Michaelmas Term 2021

Colorectal cancer (CRC) is the second most common cancer in terms of mortality in the world. Oxaliplatin is the one of the most common approved adjuvant chemotherapeutic drugs for advanced stage II and stage III. Despite the improvement in outcome, a high percentage of cases show lack of response and resistance. Thus, my aim for this thesis is to discover or validate a biomarker of oxaliplatin response with potential use in the clinic to guide therapy decisions. Response has been chosen as the primary end point because it presents a more direct measure of the drug effect. To this aim, I have used multiomic data from two clinical trials of the S:CORT consortium (Focus and Foxtrot) and generated transcriptome from another clinical trial (Coin) for deep bioinformatic interrogation of the clinical question. I first investigated 13 candidate biomarkers known for their prominent role in CRC biology. I have found that *KRAS* mutation is associated with poor response and survival to 5FUFA, but predicts benefit from the addition of oxaliplatin. Furthermore, DDRD signature positivity is predictive of no benefit of oxaliplatin compared to 5FUFA alone. Secondly, I have investigated 32 candidate biomarkers identified from a systematic literature review that I conducted for oxaliplatin resistance. High *CDC6* expression and DNA repair score favours no oxaliplatin and predicts lack of response, while high *NFE2L2* expression favours oxaliplatin response. Finally, I generated two signatures from RNA expression data of progressive versus non-progressors patients. From this study, I present predictive biomarkers of response for oxaliplatin, important in neoadjuvant clinical settings.

Acknowledgements

The S:CORT consortium was a stratified medicine consortium in colorectal cancer funded by the Medical Research Council and co-funded by Cancer Research UK. It is led by my supervisor, Prof. Tim Maughan whom I would like to thank first. Your support, guidance and supervision over the last years have been extremely valuable. I really appreciate you giving me the chance to be part of your group after a devastating first year experience at Oxford. You opened the door for me to re-believe in pursuing a PhD in a healthy, encouraging, positive environment and group. I learnt a lot from your rich experience and knowledge and I will always be so honoured to have been your student.

I owe a debt of gratitude to Dr. Enric Domingo, my co-supervisor and day to day mentor. I appreciate all the expert knowledge I was able to receive from you. Thank you for teaching me all the bio-informatic tools needed for me to become a computational scientist and thoroughly guiding me through the project. Thank you for being so accommodating with your time to meet me constantly with all my questions and concerns. I received impeccable guidance and support from you for which I express by deepest appreciation.

I would like to thank my scientific idol and mentor professor Raul Mostoslavsky, I owe my scientific success to you. You gave me the opportunity of a lifetime in 2015 when you believed in my skills and accepted me at your lab at Harvard, Massachusetts general hospital cancer center. You changed my life forever and I learnt a lot from your expert knowledge, impeccable guidance, and valuable supervision.

I also would like to thank Andrew Blake who has been a great supporter and encouraged me throughout the way. Also, from the S-CORT group, Dr. Philip Dune, you have been a valuable mentor who I referred to at several instances. Your knowledge, support and kindness have been invaluable to me. I also would like to thank Prof. Lou Brown and

David Fisher for their support. I would like to acknowledge the FOCUS trial group (CI Matt Seymour), the FoxTROT trial group (CI Dion Morton) and the COIN trial group (CI Tim Maughan) and the clinical trials units than ran those studies.

Last but not least, I owe all who I am to my mother, who brought me up on her own and sacrificed a lot for me to have a better life. Thank you momy for being the most wonderful mother in the world. You are the driving force for my existence. I do not need a father or a sibling as long as I have you. To my amazing partner and recently husband, thank you for supporting me in my PhD years and especially the last few months where you did it all just for me to focus on write up, I am so grateful. You are incredible and amazing. Thank you for being the best partner in the world. Thank you to my friend Leyla. Although being in the US, you managed to support me throughout the PhD years and be there for me to vent everyday about the pressure and stress endured. Thank you for being an awesome gem of a person.

Thank you all.

Table of Contents

List of Figures

List of Tables

Abbreviations

5FU	5-fluorouracil
ABC	ATP binding cassette
ABCC2	Adenosine triphosphate binding cassette subfamily C member 2
ABCC3	Adenosine triphosphate binding cassette subfamily C member 3
ACVR1B	Activin A receptor type 1B
ACVR2A	Activin A receptor type 2A
ANXA11	Annexin A11
AP	Abasic site
APC	Adenomatous polyposis coli
APE1	Apurinic/apyrimidinic endonuclease 1
AKT	Ak strain transforming/ Protein kinase B
ARID1A	AT rich interactive domain containing protein
ASCO	American society of clinical oncology
ATM	Ataxia telangiectasia mutated
ATP7B	ATPase copper transporting beta
AXIN	Axis inhibition protein 2
Bad	BCL-2 associated agonist of cell death
Bak	BCL-2 homologous antagonist killer
BARD1	BRCA1 associated ring domain 1
Bax	BCL-2 associated X protein
BCL-2	B-cell lymphoma 2
BCL-xl	B-cell lymphoma extra large
BER	Base excision repair
BMP4	Bone morphogenic protein 4

BRAF	V-raf murine sarcoma viral oncogene homolog B1
BRAF-V600E	Valine to glutamic acid substitution
BRCA1	Breast cancer 1
BRCA2	Breast cancer 2
CAPOX	Capecitabine and oxaliplatin
CDC6	Cell division cycle 6
CCND2	Cyclin D2
CDK5	Cyclin dependent kinase 5
CDKN1A	Cyclin dependent kinase inhibitor 1A
CDX2	Caudal type homeobox 2
CETN2	Centrin 2
ChAMP	Chip analysis methylation pipeline
CHEK1	Checkpoint kinase 1
Chk2/CHEK2	Checkpoint kinase 2
CIMP	CpG island methylator phenotype
CIN	Chromosomal instability
CK1α	Casein kinase 1 alpha
CMS	Colorectal cancer molecular subtypes
CMS-1	Colorectal cancer molecular subtype 1
CMS-2	Colorectal cancer molecular subtype 2
CMS-3	Colorectal cancer molecular subtype 3
CMS-4	Colorectal cancer molecular subtype 4
CRC	Colorectal cancer
CRIS	Colorectal cancer intrinsic subtypes
CRIS-A	Colorectal cancer intrinsic subtype A

CRIS-B	Colorectal cancer intrinsic subtype B
CRIS-C	Colorectal cancer intrinsic subtype C
CRIS-D	Colorectal cancer intrinsic subtype D
CRIS-E	Colorectal cancer intrinsic subtype E
CSA	Cockayne syndrome protein A
CSB	Cockayne syndrome protein B
CT	Computed tomography
CTC	Computed tomography colonography
CtIP	CtBP interacting protein
CTNNB1	Catenin beta 1
CR	Complete response
CRY2	Cryptochrome circadian regulator 2
CXCL8	CXC motif chemokine ligand 8
DACH	Diaminocyclohexane
DDIR	DNA damage immune response
DDR	DNA damage response
DDRD	DNA damage response deficiency
DESeq2	Differential gene expression analysis based on sequencing
DNA	Deoxyribonucleic acid
DCC	Deleted in colorectal cancer
DNMT	DNA methyl transferases
DPC	DNA protein cross links
DSB	Double stranded breaks
dsDNA	Double stranded DNA
dMMR	Deficient mismatch repair

E2F	E2F transcription factor 1
EGFR	Epidermal growth factor receptor
EIF4A2	Eukaryotic translation initiation factor 4A2
EMT	Epithelial to mesenchymal transition
EPCAM	Epithelial cellular adhesion molecule
ERBB	Erythroblastic oncogene B
ERBB2	Erb-B2 receptor tyrosine kinase 2
ERCC1	Excision repair cross complementation group 1
ERCC4	Excision repair cross complementation group 4
EZH2	Enhancer of zeste 2 polycomb repressive complex 2 subunit
FA	Fanconi anemia
FAAP24	Fanconi anemia core complex-associated protein 24
FAM123B	Family with sequence familiarity 123B
FANCA	Fanconi anemia complementation group A
FANCC	Fanconi anemia complementation group C
FANCCT	Fanconi anemia complementation group CT
FANCD2	Fanconi anemia complementation group D2
FANCG	Fanconi anemia complementation group G
FANCM	Fanconi anemia complementation group M
FBXW7	F-Box and WD repeat domain containing 7
FdUMP	Fluorodeoxyuridine monophosphate
FdUTP	Fluorodeoxyuridine triphosphate
FGF4	Fibroblast growth factor 4
FGF18	Fibroblast growth factor 18
FOCUS	Fluoxetine or control under supervision

FOLFOX	5FU, folinic acid, and oxaliplatin
FOLFOXIRI	Leucovorin, 5FU, oxaliplatin, and irinotecan
FOXC1	Forkhead box C1
FOxTROT	Fluorouracil and Oxaliplatin With or Without Panitumumab In Treating Patients with High-Risk Colon Cancer That Can Be Removed by Surgery
FRMD6	FERM domain containing protein 6
FUTP	Fluorouridine
Fz	Frizzled
G1	Gap 1 phase
G6PD	Glucose 6 phosphate dehydrogenase
GC	Guanine, cytosine
GG-NER	Glocal genome NER
GSK3B	Glycogen synthase kinase 3 beta
H2AX	Histone family member X
hCTR1	Human copper transporter 1
HER2	Human epidermal growth factor 2
HIF1α	Hypoxia inducible factor 1α
HMGB1	High mobility group box protein 1
HMGN1	High mobility group nucleosome binding domain containing protein 1
HNPCC	Hereditary non-polyposis colorectal cancer
HSPH11	Heat shock protein family B
HRAS	Harvey rat sarcoma
HTR2B	Hydroxytryptamine receptor 2B

ICL	Interstrand crosslink
ID1	Inhibition of differentiation 1
IDH1	Isocitrate dehydrogenase 1
IFL	Irinotecan, bolus 5FU and leucovorin
IL6	Interleukin 6
IGFR1	Insulin like growth factor receptor 1
IGF2	Insulin-like growth factor 2
IR	Ionising radiation
IROX	Irinotecan and oxaliplatin
IRS2	Insulin receptor substrate 2
IVT	In vitro transcribed
KRAS	Kirsten at sarcoma viral oncogene homolog
LIG1	DNA ligase 1
LIG3	DNA ligase 3
Limma	Linear models for microarray analysis
LPCAT2	Lysophosphatidylcholine acyltransferase 2
LOH	Loss of heterozygosity
LRP	Lipoprotein receptor related protein
MALAT1	Metastasis associated lung adenocarcinoma transcript 1
MAPK	Mitogen activated protein kinase
Mcl-1	Myeloid leukemia 1
MDM2	Mouse double minute 2 homolog
MEK	Mitogen activated protein kinase kinase
MHF1	Mph1-associated histone-fold 1
MHF2	Mph1-associated histone-fold 2

MLH1	MutL homolog 1
MMPs	Matrix metalloproteinases
MMR	Mismatch repair
MRE11	Meiotic recombination 11 homolog A
MRI	Magnetic resonance imaging
MRN	MRE11-RAD50-NBS1
MRP	Multi drug resistance associated proteins
MSH2	MutS homolog 2
MSH3	MutS homolog 3
MSH6	MutS homolog 6
MSI	Microsatellite instable/instability
MSS	Microsatellite stable
mTOR	Mammalian target of rapamycin
MUC5AC	Mucin 5AC
MUS81	Musculus 81
Mut	Mutated
NER	Nucleotide excision repair
NFE2L2	Nuclear factor erythroid 2 related factor 2
NF-κB	Nuclear factor kappa light chain enhancer of activated B cells
NFYB	Nuclear transcription factor Y subunit beta
NHEJ	Non-homologous end joining
NR1I2	Nuclear receptor subfamily 1 group member 2
NRAS	Neuroblastoma RAS viral (v-ras) oncogene homolog
OCT	Organic cation transporter
OCTN1	Organic cation transporter 1

OCTN2	Organic cation transporter 2
ORR	Objective response rate
OS	Overall survival
P21	Potent cyclin dependent kinase inhibitor 1
PARP1	Poly ADP-ribose polymerase 1
PCNA	Proliferating cell nuclear antigen
PD	Progressive disease
PD1	Programmed death 1
PDL1	Programmed death 1 ligand 1
PDXs	Patient derived xenograft
PFS	Progression free survival
PI3K	Phosphoinositol kinase
PIK3CA	Phosphatidylinositol-4,5-Bisphosphate 3-Kinase Catalytic Subunit Alpha
PIM1	Protooncogene serine threonine kinase
PLCε	Phospholipase C
pMMR	Proficient mismatch repair
PMS2	PMS1 homolog 2/ mismatch repair system component
POLB	Polymerase beta
POLE	DNA polymerase epsilon catalytic subunit
POLL	DNA polymerase lambda
POLZ	Polymerase zeta catalytic subunit
PR	Partial response
PTEN	Phosphatase tensin homolgue
RAC1	Rac family small GTPase 1

RAD23B	Radiation sensitive 23B
RAD50	Role of radiation protein 50
RAD51	Role of radiation protein 51
RAD51B	Role of radiation protein 51BRaf
	Rapidly accelerated fibrosarcoma
RALGDS	Ral guanine nucleotide dissociation stimulator
Ras	Rat sarcoma
RECIST	Response evaluation criteria in solid tumours
REV3	Reversionless 3
RFC	Replication factor C
RMA	Robust Multi-array average
RNA	Ribonucleic acid
RNAP2	RNA polymerase 2
RNA pol II	RNA polymerase II
RPA	Replication protein A
RPL11	Ribosomal protein L11
ROC	Receiving operating characteristic
ROCK	Rho/Rho-associated protein kinase
Ror	Receptor tyrosine kinase like orphan receptor
ROS	Reactive oxygen species
Ryk	Receptor tyrosine kinase
SCNAs	Somatic copy number alterations
S:CORT	Stratification for colorectal cancer consortium
SD	Stable disease
SLC9A9	Solute carrier family 9 member A9

SLC22A1	Solute carrier family 22 member 1
SLC22A2	Solute carrier family 22 member 2
SLC22A3	Solute carrier family 22 member 3
SLX4	Structure specific endonuclease subunit 4
SMAD2	SMAD family member 2
SMAD3	SMAD family member 3
SMAD4	SMAD family member 4
SOX9	Syr box transcription factor 9
SSB	Single strand breaks
ssDNA	Single strand DNA
STAT3	Signal transducer and activator of transcription 3
T-BET-21	T-box transcription factor 21
TCF	T cell family
TCF7L2	Transcription factor 7 like 2
TCGA	The cancer genome atlas
TC-NER	Transcription coupled NER
TGF- β	Transforming growth factor β
TGFBR1	Transforming growth factor-beta 1
TGFBR2	Transforming growth factor-beta 2
TIP60	Tat interactive protein 60
TIM3	T-cell immunoglobulin and mucin-domain containing-3
TIMPs	Tissue inhibitors of metalloproteinases
TLR4	Toll like receptor 4
TLS	Translesion synthesis
TNF-α	Tumour necrosis factor alpha

TNM	Tumour, nodes, metastasis
TOX3	Tox high mobility group box family member 3
TP53	Tumour protein 53
TSR	Tumour stroma ratio
UFT	Tegafur and uracil
USP7	Ubiquitin specific processing protease 7
UV	Ultraviolet
UVDDB	UV DNA damage binding protein
UVSSA	UV stimulated scaffold protein A
VEGF	Vascular endothelial growth factor
Wt	Wild type
XIAP	X-linked inhibitor of apoptosis
XPC	Xeroderma pigmentosum complementation group C
XPF	Xeroderma pigmentosum complementation group F
XPG	Xeroderma pigmentosum complementation group G
XRCC1	X-ray repair cross complementing protein 1
XRCC2	X-ray repair cross complementing protein 2
XRCC3	X-ray repair cross complementing protein 3
ZEB1	Zinc finger E-box binding homeobox 1

Chapter 1 Introduction

1.1 Colorectal Cancer

Colorectal cancer (CRC) is the third most common cancer in terms of incidence in the world representing 10% of all malignant tumours but it comes second in terms of mortality. In 2020, 935,000 people died as a result of colorectal cancer globally, making the disease the second-most common cause of cancer-related death worldwide after lung cancer (Colorectal cancer statistics, 2021). In a considerable percentage of patients (25% with colon cancer and 18% with rectal cancer), metastases are present upon diagnosis and the five-year survival rate in metastasized cases drops to 14% compared to 90% with localized cancers. The major metastasis localization is to the liver followed by the lungs (Kekelidze, 2013).

1.1.1 Diagnosis and staging

The gold standard of diagnosis for colorectal lesions is to perform a complete colonoscopy along with biopsies for histopathological assessment (Kaminski, 2010). In some cases, however, colonoscopy would not be successful due to patient intolerance, bowel obstruction or poor preparation, where computed tomography (CT) colonography (CTC) would be performed for diagnosis (Pickhardt, 2011). CRCs are classified using the American Joint Committee on Cancer (AJCC) TNM system (Edge & Compton, 2010; Weiser, 2018). The T stands for tumour size; how far has the tumour grown locally through the walls of the colon or rectum starting from the inner lining called mucosa from which all colorectal cancers originate (Figure 1.1). This layer includes a thin muscular layer called muscularis mucosa. Beneath this muscle layer is a fibro-vascular tissue called submucosa. Following this is muscularis propria which is a thick muscle layer. Finally,

the outermost thin connective tissue layer (subserosa and serosa) that covers most of the colon but not the rectum (Figure 1.1). N in TNM staging stands for the spread to nearby lymph **n**odes and finally the M stands for **m**etastasis if the tumour has spread to distant sites or organs. These stages are combined into an overall stage definition on which therapeutic decisions are based on. Stage 0 when the cancer is only on the mucosa (inner lining) of the rectum or colon; Stage I is when the primary tumour may involve but does not go through the muscularis propria and there is no nodal involvement neither metastases; Stage II is when the primary tumour has penetrated the full thickness of the muscularis propria and may (T4) or may not (T3) have grown through to serosal and peritoneal surface but there are no nodal or distant metastases. Stage III is when the primary tumour is of any size but it has spread to at least one loco-regional lymph node, and finally stage IV is when there are distant metastases, regardless of T/N stage. Staging of CRC is done mainly by using imaging modalities of chest, abdomen and pelvis with computed tomography (CT scan). The sensitivity of the latter for detecting liver metastasis is 74% - 84% with a specificity of 95% - 96% (Floriani, 2010). However, CT has been shown to be a poor tool for evaluating nodal involvement. Magnetic resonance imaging (MRI) is also used in diagnosis and assessments of CRC where it can evaluate liver lesions <1 cm in size with high specificity and sensitivity. Both imaging modalities are usually used in preoperative imaging of rectal cancer, MRI is recommended for local staging while whole body CT is used for detection of distant metastases (Mitry, 2010).

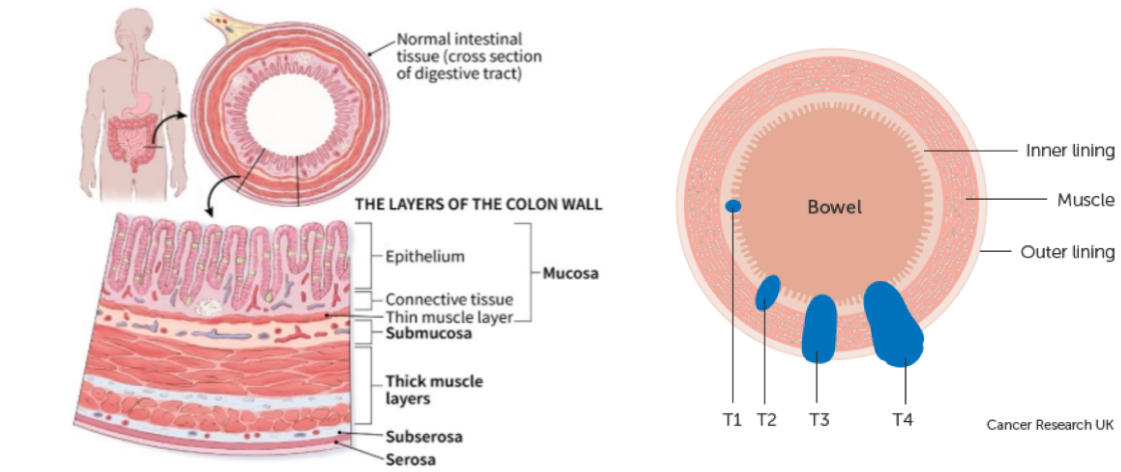


Figure 1.1 Human colon anatomy. The layers of the colon wall consist of the inner lining called mucosa from which all colorectal cancers originate. This layer includes a thin muscular layer called muscularis mucosa. Beneath this muscle layer is a fibro-vascular tissue called submucosa. Following this is muscularis propria which is a thick muscle layer. Finally, the outermost thin connective tissue layer (subserosa and serosa) that covers most of the colon but not the rectum. CRCs are classified using the TNM system. The T stands for tumour size; how far has the tumour grown locally. N in TNM staging stands for the spread to nearby lymph nodes and finally the M stands for metastasis if the tumour has spread to distant sites or organs. These stages are combined into an overall stage definition on which therapeutic decisions are based. Stage 0 when the cancer is only on the mucosa (inner lining) of the rectum or colon; Stage I is when the primary tumour may involve but does not go through the muscularis propria and there is no nodal involvement neither metastases; Stage II is when the primary tumour has penetrated the full thickness of the muscularis propria and may (T4) or may not (T3) have grown through to serosal and peritoneal surface but there are no nodal or distant metastases. Stage III is when the primary tumour is of any size but it has spread to at least one loco-regional lymph nodes, and finally stage IV is when there are distant metastases, regardless of T/N stage. Figure adapted from American Cancer Society, Colorectal cancer stages, 2019 and CRUK, TNM Staging, 2019 and is an open access under the Creative Commons Attribution License.

1.1.2 Management strategies

For stage 0 and I, the primary tumour is resectable, surgery is performed without further treatment. For stage II, surgery to remove the part of the colon that has cancer is performed. If the CRC is of high risk of recurrence, adjuvant (after surgery) chemotherapy is given. The standard of care for stage III patients is surgery, along with nearby lymph nodes resection, followed by adjuvant chemotherapy. Patients who are not fit for surgery are given radio and/or chemotherapy. Stage IV involves metastasis mainly to the liver but other organs can also be involved such as lungs, brain, peritoneum or distant lymph

nodes. Treatment is colonic surgery to remove the primary along with the metastatic site (if resectable) with chemotherapy before and after surgery. If the tumour is unresectable, chemotherapy and/or radiotherapy are the treatments of choice.

1.1.3 Conventional systemic therapies

The common chemotherapy regimens include: FOLFOX (folinic acid (leucovorin), 5-FU(5-fluorouracil), and oxaliplatin) and FOLFIRI (leucovorin, 5FU, and irinotecan), CAPEOX or CAPOX(capecitabine and oxaliplatin), FOLFOXIRI (leucovorin, 5-FU, oxaliplatin, and irinotecan) and several targeted therapies (bevacizumab targets VEGF, cetuximab or panitumumab target EGFR) (Poston, 2011). One of the most common approved chemotherapeutic regimens is the administration of 5-fluorouracil (5FU) combined with oxaliplatin as it improves outcome in the adjuvant and advanced disease settings compared to single agent 5FU (de Gramont, 2000; Pogue-Geile, 2019).

Despite the improvement in outcome seen in advanced disease when treated with oxaliplatin, a high percentage of cases show lack of response and resistance while bearing all the side effects of this chemotherapeutic drug. Thus, the identification of a biomarker for oxaliplatin outcome is crucial for better patient management.

This thesis will focus on discovering or validating a biomarker for oxaliplatin outcome in colon cancer metastatic setting. In order to provide the background for this work, this introductory chapter of the thesis will feature a discussion on oxaliplatin and its role in CRC, the current literature that explains its mechanisms of action, an overview of our current understanding of CRC biology, the different types of biomarkers and their implications on treatment decisions, the different endpoints for biomarkers testing, and

Stratification for Colorectal Cancer (S:CORT) primary hypothesis for oxaliplatin outcome.

1.2 Role of Oxaliplatin: benefits in adjuvant and metastatic colorectal cancer and toxicity

Almost half the patients with CRC develop metastasis of the disease, thus adjuvant treatment post resection is important for early-stage disease to reduce the risk of recurrence (Nicum S, 2000). In cases of liver metastasis, chemotherapy post-surgery prolongs survival (Rougier P, 2001). For the past decades, the main chemotherapeutic therapies for CRC have been based on the fluoropyrimidine, 5FU and was associated with poor response in metastatic setting which triggered its coupling with thymidine synthase modulators like leucovorin and levamisole to improve response but had no effect on progression free survival or overall survival (Schmoll, 2002). 5FU administration was also manipulated from bolus to infusion to study if there was any effect on response or survival, but the effect seen was minimal and thus did not lead to a unified regimen of administration and standard of care. In addition, new therapeutic options have been introduced like oral fluoropyrimidines such as uracil and tegafur (UFT) and capecitabine and thymidylate synthase inhibitors such as raltitrexed (Cunningham, 2002; McGavin & Goa, 2001) giving clinicians a broad spectrum of treatment options on selected patients. All of these agents cause cell death through several mechanisms: incorporation of fluorouridine (FUTP) into RNA, incorporation fluorodeoxyuridine triphosphate (FdUTP) into DNA and inhibition of thymidylate synthase, the enzyme that catalyzes the production of thymidine which is one of the nucleotides in DNA, by fluorodeoxyuridine monophosphate (FdUMP) which result in DNA damage leading to *TP53* activation and eventually cell death through apoptosis (Wilson, 2014) (Figure 1.2).

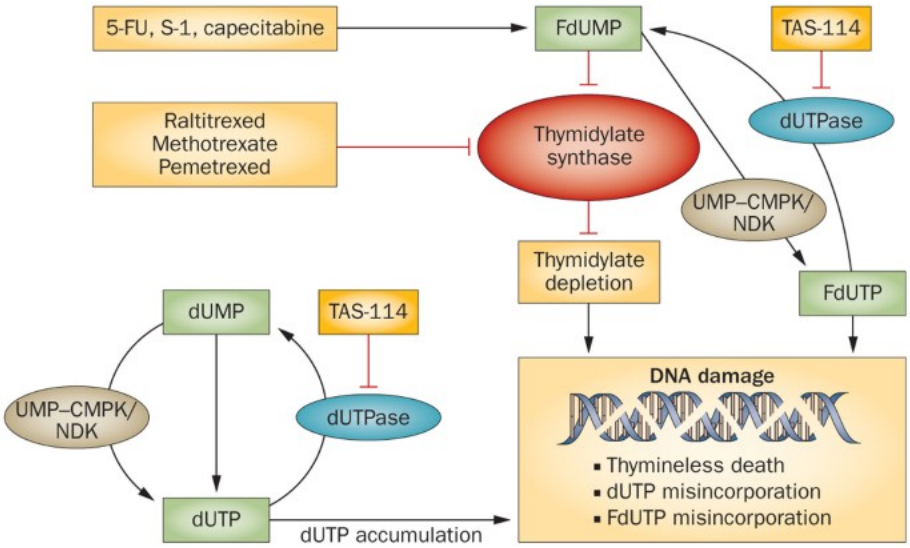


Figure 1.2 5-Fluorouracil mechanism of action. 5FU, capecitabine and raltitrexed cause cell death by several mechanisms. The first is incorporation of fluorouridine (FUTP) into RNA, another mechanism is incorporation fluorodeoxyuridine triphosphate (FdUTP) into DNA and finally inhibition of thymidylate synthase by fluorodeoxyuridine monophosphate (FdUMP). Thymidylate synthase is the enzyme that catalyzes the production of thymidine which is one of the nucleotides in DNA. All those mechanisms result in DNA damage leading to TP53 activation and eventually cell death through apoptosis. Figure taken from Standing the test of time: targeting thymidylate biosynthesis in cancer therapy (Wilson, 2014) and is an open access under the Creative Commons Attribution License.

Following this, in the mid 1990s, was the introduction of two new agents with mechanism of action unrelated to thymidylate synthase or 5FU which are irinotecan which is a topoisomerase I inhibitor and oxaliplatin, the only platinum compound with significant results in CRC (Grothey & Schmoll, 2001). Oxaliplatin monotherapy has shown modest results early on when it became available as an interesting agent in CRC setting (Bécouarn, 1998). Yet administered as a combination therapy with 5FU, oxaliplatin improved response, progression free survival, overall survival (Braun, 2003; de Gramont, 2000; Lévi, 1997) and showed synergism with the fluorpyrimidine partner . (DeBraud, 1998). Oxaliplatin in combination therapy is given under the FOLFOX regimen which has a two-week treatment cycle where oxaliplatin is administered at the beginning of each cycle. Different modifications to this regimen have been observed in

terms of dosing and schedule. FOLFOX1 was a feasibility study, followed by FOLFOX2 which is a two-week treatment cycle (De Gramont, 1997) of 100 mg/m^2 of oxaliplatin at start of the cycle then 500 mg/m^2 leucovorin by two hour infusion on days 1 and 2 followed by 1500–2000 mg/m^2 of fluorouracil continuous infusion for 24 hours on day 1 and 2, same as FOLFOX3 which only differed in oxaliplatin dosing of 85 instead of 100 mg/m^2. FOLFOX4 is 85 mg/m^2 of oxaliplatin on day 1 then 400 mg/m2 of leucovorin over two hours on days 1 and 2 before fluorouracil then fluorouracil 400 mg/m^2 bolus, then 600 mg/m^2 over 22 hours on days 1 and 2 (LV5-FU2), then a simplified version of LV5-FU2 was introduced as FOLFOX 6 (Maindrault-Goebel, 1999) with 100 mg/m2 of oxaliplatin then 400 mg/m^2 of leucovorin over two hours day 1 and fluorouracil 400 mg/m^2 bolus day 1, followed by 2400 to 3000 mg/m^2 continuous infusion over 48 hours (sLV5-FU2). Finally, FOLFOX7 which is the same as FOLFOX6 but 130 mg/m^2 of oxaliplatin. The latter is a well-tolerated regimen with high oxaliplatin dose and simplified LV5-FU2 (Maindrault-Gœbel, 2001). Higher dose of oxaliplatin (100mg/m^2 under FOLFOX2 proved to have higher response rates (46%) in metastatic disease compared to lower doses (85 mg/m^2) as seen in FOLFOX 3 and FOLFOX 4 (De Gramont, 1997; Maindrault-Goebel, 1999; Maindrault-Gœbel, 2001). Further analysis on the importance of oxaliplatin high dose showed that patients given below 85 mg/m^2 every two weeks had an objective response rate (ORR) of 19% compared to patients receiving 100 mg/m^2 who had an ORR of 39% (Maindrault-Gœbel, 2000). FOLFOX6 ORR was lower than FOLFOX2 despite the same dosage of oxaliplatin but sLV5-FU2 resulted in lower response, while FOLFOX7 has the highest dosage of oxaliplatin and resulted in again high ORR of 42% making it the best regimen in terms of ORR and fewer treatment cycles and less toxicity compared to FOLFOX2.

CRC has a high tendency to metastasize particularly to the liver where surgery is necessary as a line of treatment. Neoadjuvant chemotherapy prior to resection of a liver metastases increases resectability rate and improves survival (Daniel & Haller, 2001). The role of oxaliplatin has been shown as highly significant in liver metastatic cases where it not only improved the systemic diseases but also increased the resectability in cases where resection was not an option. Upon neoadjuvant therapy, 35% of patients has shown a shift from unresectable to resectable liver (Adam, 2001). In a study assigning patients to FOLFOX4, irinotecan and oxaliplatin (IROX) in one hand versus irinotecan plus bolus 5FU plus leucovorin (IFL), patients on IROX showed an increased median survival showing the importance of oxaliplatin and thus approving it as an important first line of treatment in metastatic CRC (Goldberg, 2004).

1.2.1 Side effects of oxaliplatin

Oxaliplatin causes adverse reactions in patients that narrow its therapeutic index. The first affected organ is the peripheral nerves. Peripheral neuropathy or neurotoxicity is very common and it manifests itself either in an acute and transient form that is observed during or after dose administration of oxaliplatin or the chronic neurotoxicity which is cumulative and appears after several doses of the drug. The acute neurotoxicity occurs within hours in all patients after infusion and is the result of transient impairment of the ion channels and nerve hyperexcitability due to voltage gated sodium channel activation (Sałat, 2020). The symptoms are usually tingling, burning or prickling felt in the hands, arms, legs or feet mainly, a condition called paresthesia. Other symptoms like jaw pain and eye pain, visual and voice changes have been reported. Another symptom is the dysesthesia where there is an abnormal sensation, like squeezing, in the legs or feet. Some of those symptoms are worsened by the exposure to cold where a few percentage

of patients (1% - 2%) will report transient pharyngolaryngeal dysesthesia manifested in breathing difficulty. Less frequently, motor symptoms are reported like muscle contractions and spasms and fasciculations which are unpredictable short contractions affecting a small number of muscle fibres causing a flicker of movement under the skin (Saif & Reardon, 2005; Wilson, 2002). The chronic neurotoxicity is also referred to as the dose limiting, cumulative neurotoxicity because it is reported after several doses of oxaliplatin. It is observed in a very high percentage of patients after a cumulative dose of 780 – 850 mg/m^2 oxaliplatin (Molassiotis, 2019). Similar symptoms of paresthesias and dysesthesia are seen, usually increasing and worsening with cumulative dosing. Loss of sensation and coordination known as sensory ataxia is also reported. In very advanced stages of neurotoxicity correlated with very high dosing of oxaliplatin (> 1000 mg/m^2), electric shock type sensation in the head is reported, known as Lhermitte's sign (De Gramont, 2000; Molassiotis, 2019; Saif & Reardon, 2005). Unlike cisplatin, oxaliplatin is not nephrotoxic, thus not damaging for the kidneys nor ototoxic (cause a hearing loss). In chronic cases, recovery of oxaliplatin induced peripheral neuropathy is usually incomplete even after chemotherapy termination. For this reason, management of those symptoms is very important and is usually done by dose modifications.

Mechanistically, several transporters have been studied to be the reason behind this neurotoxicity caused by oxaliplatin. OCTN1 and OCTN2 are organic cation transporters and are the main transporters that maintain intracellular concentrations of platinum compounds. OCTN1 is expressed within the dorsal root ganglia neurons and is attributed to be the main reason for oxaliplatin accumulation in the neurons (Jong, 2011). Another suggested reason causing the neurotoxicity is that oxaliplatin non-enzymatically breaks down into several metabolites, one of which is oxalate which is responsible for the cold induced acute neurotoxicity symptoms. Oxalate is a calcium chelator thus upon its

increase extracellularly, it chelates the calcium ions which leads to sodium increase leading to neuronal depolarization and hyperexcitation of the neurons (Deuis, 2013). Oxaliplatin has also been reported to cause myelotoxicity. Neutropenia with grades 3 or 4 is common and is usually proportional to the dosage of infusion (Tournigand, 2004). The incidence of hemolytic anemia and thrombocytopenia is less common and rare (Koutras, 2004). Finally, like any other chemotherapeutic drug, oxaliplatin has nonspecific effects on the rapidly dividing cells of the gastrointestinal tract resulting in nausea and vomiting.

1.3 Current understanding of oxaliplatin mechanism of action

Oxaliplatin is a third-generation platinum-based compound of chemical name trans-L-1,2-diamino-cyclohexane-oxalate-platinum. This group of chemotherapeutic agents includes cisplatin and carboplatin. Structurally, the latter two have two ammines carrier ligands, while in oxaliplatin those two amines have been replaced by diaminocyclohexane (DACH) ligand and oxalate as a leaving group. This structural difference confers a difference in the mechanism of action and cytotoxicity compared to its analogues. Once infused, oxaliplatin has a short initial phase of distribution but a long phase of clearance by the kidneys that usually takes 48 hours post drug administration (Graham, 2000). It has been evident for years that oxaliplatin, like other platinum compounds is passively incorporated into the cells, yet due to its high lipophilic nature of its DACH ligand, oxaliplatin passive uptake is higher compared to cisplatin or carboplatin (Dilruba & Kalayda, 2016). However, recent preclinical research has been suggesting that active transport of platinum compounds is also involved. The first oxaliplatin transporter is hCTR1 (human copper transporter 1) which is a copper influx and efflux transporter. Downregulation of CTR1 was reported with less oxaliplatin sensitivity (Howell, 2010).

Another copper transporter is ATP7B where low levels of it was correlated with higher oxaliplatin sensitivity (Martinez-Balibrea, 2015). Another group of transporters of oxaliplatin is the organic cation transporters (OCT) which are termed solute carriers and they have a role in the absorption and excretion of drugs into the intestines, liver and kidney. OCT group consists of OCT1 (SLC22A1), OCT2 (SLC22A2), OCT3 (SLC22A3) where research has suggested that OCT1 and OCT2 might transport oxaliplatin where their higher expression was correlated with higher oxaliplatin accumulation (Burger, 2011; Zhang, 2006). Another set of transporters that has a role in oxaliplatin are the ABC family which comprises multi-drug resistance associated proteins (MRP) and that has been shown to cause resistance to oxaliplatin in vitro in ovarian cancer. Once inside the cell, oxaliplatin mechanism of action has been studied and reported in several directions which will discussed in the sections that follow.

1.3.1 DNA lesions mechanism and arrest of DNA synthesis

As mentioned earlier, oxaliplatin has a DACH ligand which enters the nucleus, targets GC rich sites, form a bond with the nitrogen atom at position number 7 of guanine to form DNA transient monoadducts then stable di-adducts. Those adducts can be 1,2-intrastrand crosslinks, 1,2-interstand crosslinks or DNA-protein cross links (DPC) (Woynarowski, 2000) (Figure 1.3). Monoadducts cause no cell death and have no cytotoxicity. For oxaliplatin in particular, intra-strand crosslinks seem to be more dominant and are the cause of cell lesions (Faivre, 2003). Cisplatin forms similar adducts at the same site of DNA, but because of the DACH ligand, oxaliplatin cytotoxicity and effects differ. DACH-DNA links are bulkier than di-amine-DNA links of cisplatin, and thus they are more effective in causing cytolytic effects, if not repaired, by prohibiting DNA replication and initiating apoptosis. Moreover, cisplatin seems to form more inter-

strand cross links than intra-strand ones as the case of oxaliplatin (Woynarowski, 2000). Cisplatin DNA adducts are repaired by mismatch repair (MMR) system while oxaliplatin bulkier adducts are not. A common repair system that repairs adducts caused by both cisplatin and oxaliplatin is the nucleotide excision repair (NER) (Ahmad, 2010).

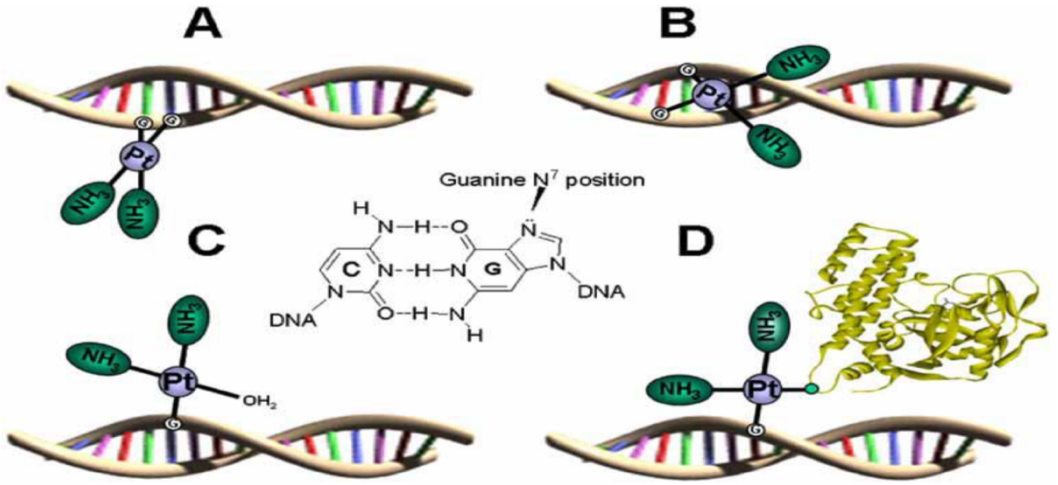


Figure 1.3 Oxaliplatin mechanism of action via DNA lesions and arrest of DNA synthesis. Oxaliplatin exerts DNA lesions by entering the nucleus, targeting GC rich sites, forming a bond with the nitrogen atom at position number 7 of guanine to form DNA transient monoadducts then stable di-adducts. Those adducts can be 1,2-intrastrand crosslinks (A), 1,2-interstand crosslinks (B), mono-adduct (C), or DNA-protein cross links (D). Figure taken from (Sehgal and Rippa, 2016) and is an open access under the Creative Commons Attribution License.

A major pathway of cytotoxicity of oxaliplatin is apoptosis. Tumour suppressor gene *TP53* detects DNA di-adducts and triggers programmed cell death. Mutations in *TP53* has been linked to oxaliplatin resistance in some references, while others did not find a clear correlation suggesting other genetic changes to be the reason for the observed resistance (Perona & Sánchez-Pérez, 2004). Those genetic changes can be attributed to proapoptotic members like (Bad, Bak, and Bax) and antiapoptotic members like (Bcl-2, Bcl-xl, Mcl-1). Down regulation of proapoptotic signaling has been correlated with higher resistance to oxaliplatin while loss of antiapoptotic signaling has been attributed to higher sensitivity to oxaliplatin (Gourdier, 2002; Hayward, 2004). Mechanistically, in colon cancer cell lines, the apoptotic cascade after oxaliplatin treatment starts with

activation of caspase 3, Bax translocates to the mitochondria while cytochrome C is released in the cytosol all signaling programmed cell death (Arango, 2004).

Other than apoptosis, oxaliplatin has been studied with 5FU to check for other mechanisms of cell death. It is usually administered with 5FU thus research into the synergistic mechanisms between the two agents has shown that oxaliplatin has a direct effect on thymidylate synthase. Thus, by oxaliplatin having an inhibitory effect on thymidylate synthase, it is causing nucleic acid synthesis arrest and effectively cell death (Fischel, 2002).

Another cytotoxic mechanism that oxaliplatin infers is inhibiting transcription. Transcription factors would have higher affinity to binding the platinum-DNA adducts rather than binding to the promoter sites to initiate transcription. Furthermore, platinum-DNA adducts blocked elongation by RNA polymerases where RNA pol II for example has been shown to be blocked by the platinum-DNA adduct (Alcindor & Beauger, 2011; Jung & Lippard, 2003).

1.3.2 Immunologic mechanism

Some research has shown that oxaliplatin can trigger an immunologic response that would result in cell death. Cells treated with oxaliplatin have been shown to emit signals that would trigger T cells to release interferon γ and activate dendritic cells with toll-like receptor 4. Clinically, metastatic patients carrying a mutation in *TLR4* gene were found to show a lesser response to oxaliplatin, along with poorer progression free survival (PFS) and overall survival (OS). Testing in mice and colon cancer cell lines, oxaliplatin was shown to trigger cells to release HMGB1 which is an extracellular inflammatory cytokine that has an effect on dendritic cells migration and T-cells activation (Tesniere, 2010).

1.3.3 Ribosomal biogenesis stress

Interesting recent results have shown in 40 different cell lines, loss of *XRCC2*, *XRCC3*, and *BRCA2* all important in homologous recombination (HR) are correlated with lack of sensitivity to oxaliplatin. Similarly, loss of *FANCC, FANCD2*, and *FANCG* genes all important in inter-strand cross link repair also confer low sensitivity to oxaliplatin. It has also been suggested that oxaliplatin causes the DNA adducts which are only toxic if the replication machinery is defective because when the replication bypass genes *POLZ* and *PCNA* were knocked out, oxaliplatin was more effective and cytotoxic. Furthermore, oxaliplatin has been shown to cause cell cycle arrest at G1 and p21 (CDKN1A) which is a cell cycle inhibitor is recruited upon its treatment. Chk2, the p53 activating kinase, is activated in response to double stranded breaks phosphorylating p53 on serine 20 overpassing MDM2 inhibition of p53. It has been shown that oxaliplatin results in signaling of γ-H_2AX and p53 activation but independent of Chk2. Thus, oxaliplatin is initiating apoptosis through p53 activation but independent of DNA damage signaling. Given the similarity between this type of apoptosis and ribosome biogenesis defects which cause rapid apoptosis independent of Chk2, ribosomal RNA synthesis has been studied upon oxaliplatin treatment. The phenotype observed, pre-RNA was highly upregulated upon oxaliplatin treatment, was similar to that observed by actinomycin D, a known ribosome biogenesis stress inducer. Ribosome biogenesis stress is when the ribosome is not produced and there is a defect or stress on the system causing pre-RNA to accumulate and excess subunits of RPL11 (component of the ribosome) that bind MDM2 releasing it from inhibiting p53 and inducing apoptosis. Moreover, cells treated with oxaliplatin had fewer polysomes which are grouped ribosomes that translate mRNA into proteins. Thus, oxaliplatin is causing a decrease in translation upon inducing a

ribosome biogenesis stress leading to apoptosis in a p53 dependent, DNA damage independent manner (Bruno, 2017).

1.4 Colorectal cancer biology

The next section will describe pathways and genes that play an important role in CRC biology, development and progression. This section will be the foundation for the first result chapter of this thesis.

1.4.1 Wnt signaling

The most dysfunctional and most mutated signaling pathway in sporadic CRC is Wnt/β-catenin. Under normal conditions, however, Wnt is a very important signaling cascade for stem and progenitor cells. Wnt can be activated non-canonically, in a β-catenin (a cadherin and transcription activator) independent manner, this signaling can happen through various receptors like receptor tyrosine kinase (Ryk) and receptor tyrosine kinase-like orphan receptor (Ror). The most studied pathway, however, in Wnt signaling is the canonical pathway that has β-catenin as a key player (Figure 1.4). Upon Wnt stimulus, Wnt protein ligand binds to its frizzled (Fz) receptor, a receptor complex composed of a seven-transmembrane receptor of the Frizzled family and co-receptor low density lipoprotein receptor-related protein (LRP) (Schweizer & Varmus, 2003). This inactivates GSK-3β kinase and stabilizes β-catenin which translocates to the nucleus, binds proteins of the T-cell family (TCF) and activates a range of target genes. Wnt signaling drives tumour proliferation and inhibition of differentiation as it ensures stem cells are maintained in their undifferentiated state at the crypts of the colon, it also allows the survival of cancer stem cell. Negative regulation of Wnt signaling happens when cytosolic β-catenin is degraded by a serine/threonine kinase casein kinase 1 alpha (CK1α)

and glycogen synthase kinase 3 (GSK3). This degradation is also regulated by axis inhibition protein (Axin) and the tumour suppressor gene adenomatous polyposis coli (*APC*). The latter blocks cell cycle progression from G1 to S phase. Thus, *APC*/axin/GSK-3β complex targets β-catenin for degradation in the absence of Wnt signal. Loss of function of *APC* through mutation therefore causes unregulated activation of β-catenin driving uncontrolled proliferation of the basal crypt cells opening the door to carcinogenesis.

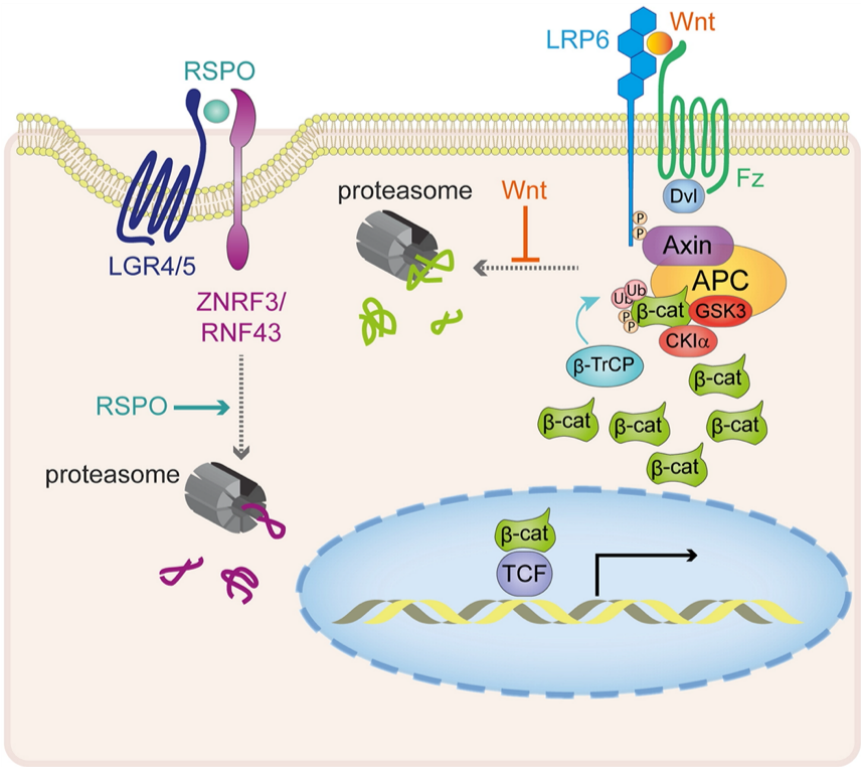


Figure 1.4 Wnt Signaling Pathway. Upon Wnt stimulus, Wnt protein ligand binds to its frizzled (Fz) receptor and co-receptor low density lipoprotein receptor-related protein (LRP). This inactivates GSK-3β kinase and stabilizes β-catenin which translocates to the nucleus, binds proteins of the T-cell family (TCF) and activates a range of target genes. Negative regulation of Wnt signaling happens when cytosolic β-catenin is degraded by a serine/threonine kinase casein kinase 1 alpha (CK1α) and glycogen synthase kinase 3 (GSK3). This degradation is also regulated by axis inhibition protein (Axin) and the tumour suppressor gene adenomatous polyposis coli (APC). Wnt agonists called R-Spondin (RSPO proteins) enhance Wnt circuit by stabilizing Fz and LRP proteins. RSPO forms a complex with LGR4/5 receptor and E3 ubiquitin ligase proteins ZNRF3/RNF43 which are antagonists of Wnt and thus RSPO proteins augment Wnt signaling. Figure taken from (Krausova & Korinek, 2014) with the permission from Elsevier, license number 4746071375060.

The development of CRC can be due to both genetic and environmental factors. The biggest percentage of cases are sporadic without evidence of a family history which accounts for 85% of cases (Kuipers, 2015). A subgroup of patients is affected by hereditary CRC syndromes, the most common one being Lynch syndrome and familial adenomatous polyposis which consequently develop into colorectal cancer. Loss of function in the *APC* gene through mutations cause familial adenomatous polyposis. This kind of mutation is also seen in the majority of sporadic colorectal cancer cases. *APC* is the first mutation to occur in a series of subsequent mutations that would drive clonal expansion. In addition, *APC* mutation would render β-catenin constitutively active activating targeted genes like *MYC* and *cyclin D1* which are cell cycle and transcription genes important for proliferation, vascular endothelial growth factor (VEGF) and fibroblast growth factor 4 and 18 (FGF4 and FGF18) which are important in signaling, E-cadherin which is important in adhesion, and other downstream genes important for tumour proliferation and growth (Klaus & Birchmeier, 2008). Finally, *APC* loss can trigger an inflammatory signal where it has been shown in mice that APC knockout and loss activate *RAC1* that produces reactive oxygen species (ROS) and activate NF-κB signaling. This inflammatory signal, where NF-κB is constitutively active, is shown to promote the survival of cancer cells by inhibiting apoptosis and stimulating cell proliferation (Hoesel & Schmid, 2013).

1.4.2 Adenomas to carcinomas progression

For epithelial cells in the colon to become adenocarcinoma, there is a progression of epigenetic, genetic and histological events over time. Cancer arises from a polyp beginning from an abnormal crypt then evolving into a tubular adenoma (<1cm in size) by accumulation of mutations and epigenetic changes to further develop into an adenoma

(>1cm in size) and finally into CRC. The genetic multistep model was first discussed in the 1990, where first colorectal cancer results from the activation of oncogenes and the inactivation of tumour suppressor genes. Second, there should be mutations in at least four or five genes for the cancer to arise. Third, the tumour biologic features and properties are determined by the number of mutations happening and the clonal expansion dictated by those mutations, rather than the sequence or order of the mutations (Fearon & Vogelstein, 1990). More recently, genomic instability in colorectal cancer has been focused on chromosomal instability pathway which is also termed adenoma to carcinoma sequence. The first genomic event in the majority of cases is an initial mutation in the *APC* gene where normal cells start changing into early adenoma by acquiring adhesive properties along with Wnt signaling. Several genomic events follow this initial event like activation of proto-oncogene *KRAS*, mutation in *SMAD4*, *TGFBR*, *PIK3CA*, loss of tumour suppressor *TP53*, and loss of heterozygosity for the long arm of chromosome 18 (18q LOH), all of which help early adenomas progress to late adenomas and consequently to adenocarcinomas by activating growth factors, angiogenesis, and inhibiting apoptosis (Figure 1.5) (Armaghany, 2012).

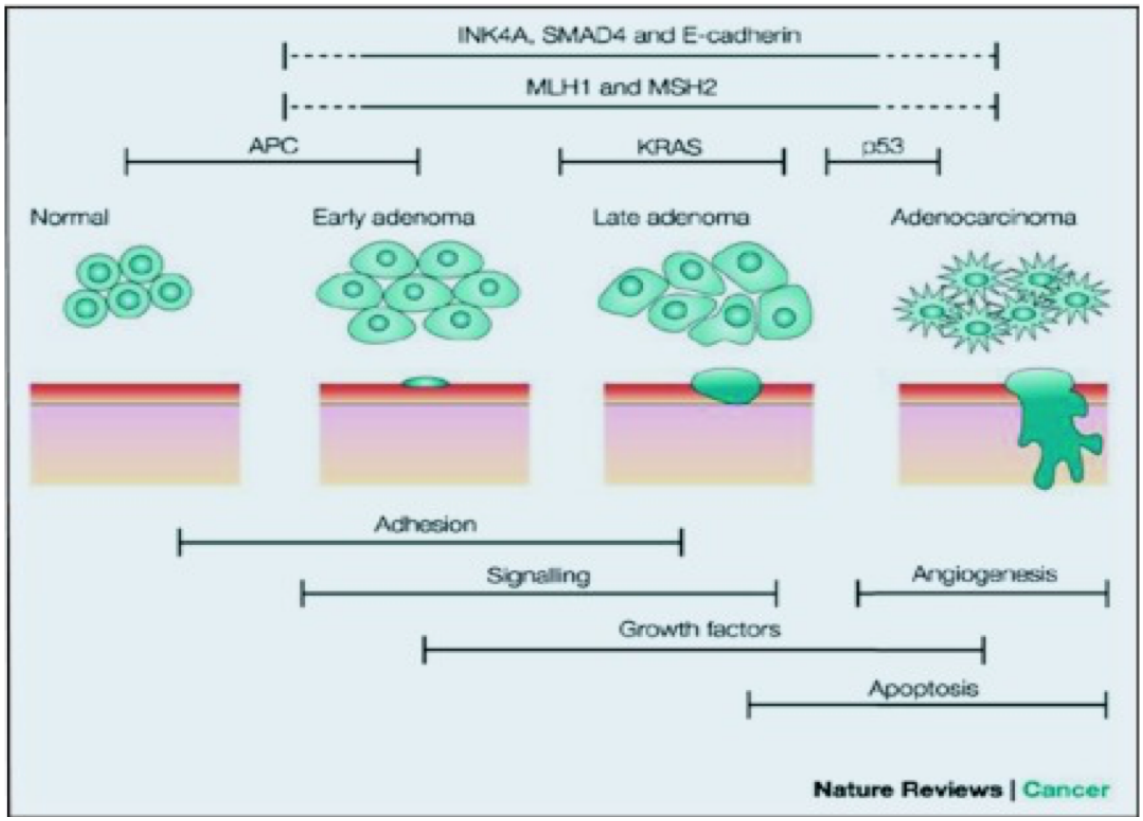


Figure 1.5 Adenomas to carcinomas model. The progression from adenoma to carcinoma is a result of several genetic and epigenetic events. The first genetic mutation in the progression from adenoma to carcinoma is mutation in the *APC* gene activating Wnt signaling leading to adhesion properties gain. This genetic event is followed by the activation of the proto-oncogene *KRAS* and finally inactivation of *TP53* inhibiting apoptosis and promoting angiogenesis. Figure taken from (Armaghany, 2012) and is an open access under the Creative Commons Attribution 4.0 International License.

KRAS and *BRAF* are part of the mitogen activated protein kinase (MAPK) pathway of signal transduction which affects cell proliferation, differentiation and cell death. *KRAS* is also a member of the RAS family of oncogenes which includes *NRAS* and *HRAS* but *KRAS* is the most frequently mutated in colorectal cancer. The signal transduction starts with the endothelial growth factor receptor (EGFR) where its ligand EGF binds and causes the receptor dimerization which consequently induces its phosphorylation. This phosphorylation signal activates a cascade of downstream effectors, the first of which is RAS which activates RAF which in turn activates MEK leading to cell proliferation and angiogenesis (Lièvre, 2008a). Simultaneously, there is a parallel signal from the EGF binding of the EGFR activating phosphoinositol kinase

(*PI3K*) which activates downstream effector *AKT* and in turn *mTOR* leading to inhibition of apoptosis. Those signals lead to cancer proliferation, invasion and metastasis (Samuels, 2004). *PTEN* is a negative regulator of the *PI3K/AKT/mTOR* pathway as it dephosphorylates *PIP3* and inactivates *AKT*. In many CRC cases, *PTEN* is also mutated and inactivated (Yin & Shen, 2008). *TP53* is a gene involved in cell cycle control at G1 prior to DNA replication at S phase, this control is ensured by DNA repair and if the latter is unsuccessful, *TP53* induces cell death through apoptosis. It is very commonly mutated in colorectal cancer (Vogelstein B., 1988). 18q LOH is loss of one of the alleles of a region on chromosome 18 long arm (18q21.3). The gene located at this region is a tumour suppressor gene and it was names Deleted in Colorectal Carcinoma (DCC). The ligand of this gene is called netrin-1 whose concentration in the colon crypts decreases as cells differentiate and move up the surface. Upon DCC loss, netrin-1 concentrations will remain high leading to abnormal cell survival. LOH is present in 70% of colorectal cancer cases (Armaghany, 2012). All those genetic instability events lead to adenoma to carcinoma progression.

1.4.3 Microsatellite instability model

Mismatch repair (MMR) is a repair system that proofreads DNA polymerase during cell replication. DNA polymerase assembles nucleotides in a 5' to 3' direction on the template strand, yet it has the capacity of going backwards, exerting its endogenous exonuclease activity and correcting for the errors in the assembled strands. MMR detects errors that are missed by the polymerase and repairs them. A deficient MMR (dMMR) system, or microsatellite instability (MSI) leaves the genome with microsatellites which are short tandem repeated sequences that are shorter or longer than the template strand leading to a hypermutable phenotype. The deficiency in MMR occur through a point

mutation in one of the MMR genes MLH1, MSH2, MSH6, PMS2, or epigenetic silencing by hypermethylation of the promoter region of MLH1, where most dMMR cases in colorectal cancer harbor a defect in MLH1 and MSH2 (Kawakami, 2015). The national cancer institute workshop specified 5 markers (two mononucleotide – BAT25/26 and three dinucleotide markers – D2S123, D5S346, and D17S250) for MSI detection on fresh frozen or paraffin embedded tissues using a PCR-based assay: MSI-high (MSI-H) is defined when the instability in MMR is seen in ≥30% of the MMR markers (two or more of the five markers) and MSI-low (MSI-L) is when the instability in MMR is seen in 10%–29% of the MMR markers (one out of the five markers). Microsatellite stable (MSS) is when there is no MMR deficiency or instability in the markers (Ogino, 2009). MSI status can also be determined by immunohistochemistry where reduced expression of *MLH1, MSH2, MSH6* and *PMS2* genes identifies tumours as MSI. MSI is a dominant marker in 95% of cases of hereditary non-polyposis colorectal cancer (HNPCC) or Lynch syndrome (which accounts to 3% - 4% of colorectal cancer cases) while in sporadic cases it is only present in 15% - 20% (Geiersbach & Samowitz, 2011). Patients with Lynch syndrome develop colorectal cancer at an early age and harbor *MLH1*, *MSH2*, *EPCAM*, and *PMS2* mutations. Importantly, Lynch syndrome patients who develop cancer do not carry neither a KRAS mutation nor BRAF V600E mutation (Kawakami, 2015). dMMR serves as a distinctive model or phenotype in colorectal cancer in comparison to proficient MMR (pMMR): dMMR is more common in stage II (20%) in comparison to stage III (12%) and only around 4% in metastatic setting. Moreover, dMMR have a proximal colon prevalence, mucinous histology and high immune infiltrated tumours (Roth, 2010). dMMR patients in stage II have a better prognosis compared with pMMR (Roth, 2012).

1.4.4 Genetic landscape of colorectal cancer

The Cancer Genome Atlas reported a key analysis of the genetic landscape of cancer. Tumour and matched normal pairs were analysed to assess the genetic landscape of colorectal cancer. Exome sequencing analysis was done on 224 matched pairs and the cases were separated with a mutation rate of <8.24 per 10^6 categorised as non-mutated tumours and those with a mutation rate of >12 per 10^6 basepairs categorised as hypermutated tumours. 77% of the hypermutated tumours were MSI-H with the majority having MLH1 mutation and CpG island methylator phenotype (CIMP). However, the top six highly mutated tumours were negative for MSI-H, CIMP or MLH1 mutation, but had mutations in other MMR genes or *POLE*. The gene mutation analysis showed 15 common mutations in the hypermutated tumours and 17 common gene mutations in the non-hypermutated tumours. *APC, TP53, KRAS* and *NRAS* in codons 12 and 13 or codon 61, *PIK3CA, FBXW7, SMAD4, TCF7L2, CTNNB1, SMAD2, FAM123B* (regulator of Wnt signaling) and *SOX9* (important in cell differentiation and intestinal stem cell signaling) were the most commonly mutated in the non-hypermutated tumour. On the other hand, the hypermutated tumours showed *ACVR2A, APC, TGFBR2, MSH3, MSH6, SLC9A9, TCF7L2* and *BRAF V600E.* Those genetic differences between non-hypermutated and hypermutated tumours suggest that tumour progression happens through different signaling and genetic pathways. Chromosomal and sub-chromosomal analyses on 257 tumours showed that hypermutated tumours had less somatic copy number alterations (CNA). The most significant deleted chromosome arms in colorectal cancer tumour were 18p and q that includes *SMAD4* and 17p and q that includes *TP53*. Other notable deletions span four tumour suppressor genes *SMAD4*, *APC*, *PTEN*, and *SMAD3* (all tumour suppressors). MSI and MSS tumours showed no difference in CNA.

Other chromosomal regions had noted amplification, one that contains *Myc* (8q24), and another that contains *ERBB2* (*HER2*) (17q21.1). Another focal amplification was detected

on chromosome 11 which harbours *IGF2* gene (11p15.5). Pathway analyses showed alterations in Wnt, MAPK, PI3K, TGF- β, and p53 pathways. Wnt signaling is affected in 93% of colorectal cancer tumours with loss of APC or gain mutations in CTNNB1 in the majority of cases. Other Wnt genes affected are *SOX9* (Wnt/β-catenin effector), *TCF7L2* (interacts with β-catenin to upregulate Wnt target genes), *AXIN2* (negative regulalor of Wnt), *FBXW7 (*Wnt/β-catenin effector), *ARID1A* (downstream effector of SOX9), and *FAM123B* (negative regulator of Wnt/β-catenin). MAPK and PI3K pathways are deregulated and altered in colorectal tumours due to alterations in *KRAS, NRAS, BRAF, PIK3R1, PIK3C, PTEN, IGF2,* and *IRS2* (encoding a protein linking IGF1R, the receptor for IGF2, with PI3K). Alterations in *TGFBR1, TGFBR2, ACVR2A, ACVR1B, SMAD2, SMAD3 and SMAD4* lead to alterations in TGF-β signaling seen in 27% of non-hypermutated and 87% in the hypermutated tumours. Mutations in *TP53* and *ATM* that activates *TP53* by phosphorylation affect p53 signaling pathway in colorectal tumours. Finally, taken altogether, gene expression, copy number methylation and pathway analyses, almost all cases of colorectal tumours show alterations in *Myc* and hypermutated show differential alterations than the non-hypermutated (Muzny, 2012).

1.4.5 Transcriptomic landscape of colorectal cancer

Based on transcriptomic data, colorectal cancer was classified into 4 different subtypes (called consensus molecular subtypes (CMS1-4)) by an international consortium which pooled and reanalysed over 4000 cases of CRC (Guinney, 2015). CMS1 samples are hypermutated, had low somatic copy number alterations (SCNAs), high MSI tumours, defective MMR, high *BRAF* mutation, increased expression of immune related genes, and strong activation of immune pathways. CMS2 samples represent the canonical pathway and have high chromosomal instability (CIN), high copy number gains in oncogenes and

losses in tumour suppressor genes, displays epithelial differentiation, upregulation of Wnt and *MYC*. CMS3 had fewer SCNAs, high CIMP, intermediate level of hypermethylation, high *KRAS* mutations, and enrichment for metabolism signatures that rendered the subtype as metabolic. CMS4 showed upregulation of genes that have a role in epithelial to mesenchymal transition (EMT), activated TGF-β signaling, angiogenesis, stromal infiltration profile, and activated complement inflammatory system. Thus, CMS1 is immune, CMS2 is canonical, CMS3 is metabolic and CMS4 is mesenchymal. CMS1 has been shown to be more frequent in females with right sided tumours and higher cancer grade, while CMS2 are mainly left sided. CMS4 are usually stages III and IV with worse overall survival. CMS2 have better survival after relapse compared to CMS1 which has poor prognosis with MSI and BRAF mutation (Guinney, 2015).

Whole genome RNA analysis (microarray or sequencing) is essential to call CMS, yet the procedure is costly in the clinical setting, thus an alternative is to use immunohistochemistry which has been utilized to call CMS. In several studies, high MSI were classified as CMS1 while the remaining cases were checked for epithelial component to be categorized as CMS2 or CMS3, and the mesenchymal feature for CMS4 was assessed based on four markers: *CDX2, FRMD6, HTR2B, and ZEB1* (Ten Hoorn, 2018; Trinh, 2017). CMS was also studied and its effect assessed in carcinogenesis and cancer progression. One study reported that adenomatous polyps were classified as CMS2 and were left sided, while hyperplasic polyps and serrated adenomas were classified as CMS1 and were frequently right sided. Low risk adenomas were classified into CMS3. CMS2 sporadic polyps may become CMS1 if they become deficient in MMR or they can become CMS3 if they acquired a *KRAS* mutation and activated MAPK pathway or they can become CMS4 if they activated TGF- β signaling. Metastatic colorectal cancers are grouped in CMS2 and CMS4 with very few cases in CMS1 and CMS3 (Chang, 2018;

Fontana, 2019). CMS classifications has also been linked to prognosis where CMS1 with mutated *TP53* and MSS status were associated with poor prognosis (Smeby, 2019). CMS1 with microsatellite stable (MSS) subtype and *BRAF* mutation is associated with worse overall survival (OS) compared with *BRAF* wild type (Smeby, 2018). *BRAF* mutation has also been studied in the CMS2 context. *BRAF* mutation in codons 597/601 and codons 594/596, thus not V600E, were likely to belong to CMS2 or CMS3 but survival was not affected when compared to *BRAF* wild type (Schirripa, 2019). CMS2/3 cases with *KRAS* mutations have poorer survival compared with wild type *KRAS* (Smeby, 2018). CMS2 has high frequency of CNAs and survival analysis of CMS2 stage I-III MSS cases with high focal amplifications in *ERBB2, MYC, TOX3, CCND2, and ANXA11* were all associated with poor survival (Berg, 2019). CMS2 stage II and III patients with CD8 positivity showed better prognosis (Allen, 2018). CMS3 with SMAD4 mutation had poorer OS compared with wild type SMAD4 (Sarshekeh, 2017). Gremlin1 expression is higher in CMS4 and is associated with poor prognosis (Dutton, 2019). Activation of NOTCH signaling in CMS4 setting drives metastasis and poor prognosis. CMS4 with macrophage positive cells activated NF-κB signaling and was associated with poor prognosis as well (Jackstadt, 2019). High expression of the hypoxia inducible factor 1A (HIF1A) and low expression of DNA repair proteins (RAD51, KU70) in CMS4 were associated with poor prognosis(Jongen, 2017).

Therapeutic effects of the common cytotoxic drugs in colorectal cancer have been tested in accordance with CMS. In CMS2, testing in cell lines showed higher apoptosis upon 5-FU with oxaliplatin treatment compared to CMS4. Similarly, in CMS2 PDXs, this treatment resulted in longer survival, but not in CMS4 (Sawayama, 2020). Adjuvant therapy improved the prognosis of stage II CMS2 cases and had longer survival in stage III CMS2 and CMS3 cases (Allen, 2018). The addition of oxaliplatin had a more

favorable prognosis on stage III CMS2 patients compared to standard chemotherapy (Song, 2016).

1.4.6 Importance of stroma

In colorectal cancer, it has become apparent that tumour stroma, that includes fibroblasts, inflammatory cells and endothelial cells wrapped in proteins of the extracellular matrix, plays a vital role in disease initiation, progression and prognosis. The tumour stroma supplies the tumour with growth factors, metabolites and cytokines. It has been shown that myofibroblasts, or cancer associated fibroblasts (CAFs), are present in adenomas and they have a role in tumour initiation (Cui, 2009). Furthermore, myofibroblasts have been associated with poor prognosis in colorectal cancer (Tsujino, 2007). This poor prognosis is the result of immature stroma and down regulation of the immune pathways (Ueno, 2004). In recent years, it has become evident that cells undergo transformation and differentiation into the tumour stroma with some suggestions that the myofibroblasts are derived from epithelial cells via EMT. This differentiation is induced by several cytokines most importantly TGF-β which mediates RAS and SMAD signaling to result in EMT (Pino, 2010; Safina, 2009). Furthermore, myofibroblasts has also been shown to activate Wnt signaling and dedifferentiate cells back to their stem like characteristics (Vermeulen, 2010). Myofibroblasts have been also shown to induce angiogenesis, invasion and promote metastasis. This is accomplished by their effect on metalloproteinases (MMPs) and their inhibitors (TIMPS) causing the remodeling of the extracellular matrix and the tumour microenvironment (Conti & Thomas, 2011). Cytokines interferon γ and TNF-α have been reported as factors inhibiting this trans-differentiation (Yamasaki, 2008). Some classes of proteases, αv integrins, have been shown to activate TGF- β signaling and

drive EMT and the formation of myofibroblasts. Those integrins have been shown to cause colorectal cancer progression and their blocking has suppressed chemotherapy resistance and decreased EMT and invasion (He, 2009; Monnier, 2008; Yang, 2008). Moreover, tumour-stroma ratio (TSR) is a prognostic marker in primary tumour that has been developed and assessed by immunohistochemistry via haematoxylin and eosin (H&E) stain that links patients with high ratio to worse prognosis (Huijbers, 2013; West, 2010). TSR has been shown to be a vital tool in identifying stage II patients with the risk of disease recurrence. Compared with the American society of clinical oncology (ASCO) criterion which selects those high-risk patients for adjuvant therapy (Benson, 2004), TSR improved this selection by 14%. Similarly, in metastatic setting, TSR has been found of importance where it has been shown as a prognostic factor irrespective of the N status. Furthermore, assessing TSR in the metastatic nodes of stage III patients has been shown of significance. The presence of stroma in the lymph nodes affected the prognosis, thus not only the number of metastatic lymph nodes would reflect the prognosis, but also the composition of the microenvironment with the metastatic lymph node is crucial for prognosis (Huijbers, 2013; van Pelt, 2016). Finally, the importance of stroma has further been consolidated by the CMS classification where CMS4 with high stromal content and high EMT is associated with the worse prognosis. (Isella, 2015).

1.4.7 Colorectal cancer intrinsic transcriptional traits

Using patient derived xenografts (PDXs), five different colorectal cancer intrinsic subtypes have been specified by a group in Torino. CRIS-A has a predominance for MSI and mutated *BRAF*, right colon, mucinous histology, glycolytic and enriched for *KRAS* mutations. CRIS-A has been associated with inflammatory traits. CRIS-B has an activated TGF-β signalling leading to high EMT and in general poor prognosis. Both

CRIS-A and CRIS-B have CIMP and hypermutation phenotypes. CRIS-C has activated EGFR and ERBB signalling and consequently receptive to EGFR inhibitors and is enriched for KRAS wild type. It also shows high focal amplification of 8q24.21 which contains *MYC*. CRIS-D has a stem-like phenotype with activated Wnt signalling along with *IGF2* overexpression and activation. Finally, CRIS E has also Wnt signalling with Paneth cell like phenotype and *TP53* mutations. CRIS-C-D-E show CIN and high CNA. Because CRIS is a classification focused on cancer cell intrinsic gene expression, it has been found to be more clinically useful in classifying patients samples successfully overcoming the stromal derived intra-tumour heterogeneity and clustering samples per patient rather than per region of origin (Dunne, 2017). Testing those subtypes with therapy sensitivity, CRIS-C is highly responsive to cetuximab which is an anti-EGFR antibody, while CRIS-A shows no response. This is explained by the lack of RAS mutations in CRIS-C group where RAS mutations has been shown to drive resistance to cetuximab. CRIS has also been assessed as an independent predictor of prognosis in colorectal cancer. CRIS-B has been associated with poor prognosis independent of tumour stage (Isella, 2017). CRIS-C shows low levels of CD8 tumour infiltrating cells and cluster into CMS2 with a better prognosis in stage II and III with adjuvant chemotherapy (Allen, 2018).

In this thesis, I have tested the impact of various of these fundamental colorectal cancer genes and transcriptional signatures to see if they impact oxaliplatin sensitivity.

1.5 DNA damage repair pathways

The cells are prone to several types of DNA damage both endogenous or exogenous leading to genome instability and increased mutagenesis. If not repaired or

corrected, these lesions have a direct effect on DNA replication and transcription. The first endogenous cause of DNA damage is physiological arising from DNA replication error. Despite, the DNA polymerase having an exonuclease activity where it corrects any misinserted nucleotide during DNA replication, yet replication error still accumulates (Kunkel, 2009). Another reason for replication errors is strand slippage which is the misalignment of the DNA strands when replicating repetitive sequences leading to change in the open reading frame and genetic rearrangements one of which is MSI (Castillo-Lizardo, 2014). Another endogenous cause is reactive oxygen species (ROS) which are by products of the cellular respiration process. One form of ROS is free radicals which are DNA damaging agents (Winterbourn, 2008). ROS and nitrogen compounds are also by products of the immune system in sites of inflammation and infections (Kawanishi, 2006). This can lead to the formation of single strand breaks (SSB) where in close proximity to another SSB can lead to double strand breaks (DSB) which are more toxic and difficult to repair. Hydrolytic DNA damage also occurs endogenously and involves the deamination and removal of the amine group from the DNA base or the complete removal of the base. This type of damage is the result of ROS or the chemical reaction of some metabolites in the body (Jackson & Bartek, 2009). Finally, DNA breaks also arise during aberrant DNA topoisomerase reactions which can occur due to a therapeutic cancer treatment or even spontaneously in cells (Ranjha, 2018).

Exogenous DNA damaging agents are numerous. The first example is ultraviolet (UV) radiation which is an environmental DNA damaging agent. Although the ozone layer absorbs most of the UV light, what escapes can cause thousands of lesions per cell per day. Another radiation is the ionising radiation (IR) which cause DNA breaks, particularly DSBs. Also, aromatic amines, which are alkylating agents produced from

tobacco products, coal, pesticides, and fuel, are chemical exogenous agents known to cause DNA damage (Chatterjee & Walker, 2017). Finally, there are some environmental factors that are linked to DNA damage, like extreme exposure to specific toxins like aflatoxins, hypoxia and oxidative stress (Gafter-Gvili, 2013). Another causative mechanism for DNA damage is oxidation of specific bases of the DNA known as oxidative DNA damage; one of the most common one is 8-hydroxydeocyguanosine which is often caused by chemical carcinogens. Finally, cancer chemotherapies are considered exogenous DNA damaging agents as they use chemicals that inflict different DNA lesions. Alkylating agents attach alkyl groups to DNA bases, while crosslinking agents like cisplatin, carboplatin and oxaliplatin induce crosslinks between DNA bases as their cytotoxic activity. Another group of chemical agents, like irinotecan, are topoisomerase enzyme inhibitors that induce SSBs and DSBs. Other chemical agents used to treat colorectal cancer are 5FU and capecitabine which are fluoropyrimidine analogs that incorporate fluorodeoxyuridine into DNA and inhibit thymidylate synthase, the enzyme that catalyzes the production of thymidine resulting in DNA damage (Ciccia & Elledge, 2010).

The cells are equipped with several DNA damage repair (DDR) mechanisms, cell cycle checkpoints and cell death pathways to tolerate DNA damage and reduce its effects. Mutations in DDR components are the predisposing cause to many cancers (Ciccia & Elledge, 2010). In this section, I will present DDR pathways that are important and play a role in colorectal cancer and/or oxaliplatin response, aside from MMR pathway which has been covered in section 1.4.3.

1.5.1 Base excision repair

The first DDR mechanism is base excision repair (BER) which corrects small base lesions that usually result from deamination, oxidation or methylating agents and do not cause DNA helix structure distortion. This repair mainly happens during G1 phase of the cell cycle and the lesion is detected and removed by a DNA glycosylase leaving an abasic site (AP-site) which is a location that is void of a base in the DNA sequence and can cause transient SSBs. This step is followed by Apurinic/apyrimidinic endonuclease 1 (APE1) activation which cleaves the sugar phosphate backbone, followed by DNA synthesis in a poly ADP-ribose polymerase 1 (PARP1) dependent manner to fill the gap, and finally ligation in polymerase beta (POLB) and DNA repair X-ray repair cross complementing protein 1 (XRCC1) dependent manner (Krokan & Bjørås, 2013; Slyskova, 2018). Several mutations in BER pathway have been linked to genome instability and cancer. Downregulation of OGG1 which is a DNA glycosylase has been associated with aging and cancer (Kovtun, 2007). Colorectal cancer specimens have been shown to harbour mutations in *POLB* affecting the efficiency of DNA repair and increasing the rate of mutations observed in colorectal cancer (Wang, 1992). *XRCC1* mutation is seen in several cancers including lung adenocarcinoma, colorectal cancer, melanoma and recently has been linked to the early pathogenesis of breast cancer (Ali, 2018; Sweeney, 2017). In addition, high BER genes expression has also been associated with poor clinical outcome in colorectal cancer (Leguisamo, 2017). BER pathway key players have also been studied in chemotherapy drug responses. Oxaliplatin resistant colon cancer cells have shown to have higher levels of POLB protein and faster repair of oxaliplatin adducts. POLB deficient fibroblasts were more sensitive to oxaliplatin compared to normal fibroblasts (Yang, 2010). Furthermore, *PARP1*, *XRCC1* and *POLB* inactivation in a CRISPR/Cas9 loss of function screen rendered cells more sensitive to oxaliplatin, particularly *XRCC1* result is seen in an

OGG1 dependent manner which is a key player of BER (Slyskova, 2018) indicating that cells use several DDR mechanisms, including BER, to repair lesions caused by oxaliplatin leading to resistance. Differential pathway analysis on patient derived organoids from oxaliplatin resistant gastric patients and oxaliplatin sensitive patients has shown enrichment pathways for drug resistant patients to include HR, DNA replication, BER and cell cycle regulation. PARP1 gene was upregulated in drug resistant organoids and its knockdown decreased resistance. PARP1 plays a vital role in BER and it enhances break repair caused by oxaliplatin. The PARP1 inhibitor olaparib combined with oxaliplatin significantly overcame the resistance to oxaliplatin.

1.5.2 Nucleotide excision repair

Nucleotide excision repair (NER) is a DDR mechanism that repairs bulky lesions from UV exposure and intrastrand crosslinks causing DNA double helix distortion. This helix distortion blocks replicative DNA polymerases progression causing replication fork stalling and DNA breaks formation (Budzowska & Kanaar, 2009). NER repair machinery is one of the mechanisms that decrease the consequences of replicative stress by repairing such lesions. It is sub-divided into two pathways: transcription coupled NER (TC-NER) which targets lesions that block transcription and global genome NER (GG-NER). In TC-NER, lesions are primarily detected by RNA polymerase 2 (RNAP2) which activates and recruits TC-NER proteins Cockayne syndrome protein A (CSA) and Cockayne syndrome protein B (CSB) which consequently activate several downstream TC-NER specific factors for repair like UV stimulated scaffold protein A (UVSSA), ubiquitin specific processing protease 7 (USP7), high mobility group nucleosome binding domain containing protein 1 (HMGN1) and others. In GG-NER, lesions are detected by UV DNA damage binding

protein (UVDDB) triggering the binding of Xeroderma pigmentosum complementation group C (XPC) complexed with UV excision repair protein Radiation sensitive 23B (RAD23B) and centrin 2 (CETN2). Both of the detection signals seen in TC-NER and GG-NER trigger excision of the damaged strand by ERCC1, Xeroderma pigmentosum complementation group F and G (XPF and XPG) endonucleases. Finally, gap filling synthesis is carried out by replication proteins like PCNA, replication factor C (RFC), and XRCC1, while ligation step is performed by ligation proteins like DNA ligase 1 and 3 (LIG1 and LIG3)(Chatterjee & Walker, 2017; Slyskova, 2018). Replication stress response pathway is associated with NER pathway. ATM and ATR are kinases that mediate replication stress response by stabilising and restarting replication at stalled forks avoiding DNA damage and genome instability. Deficiency in those proteins results in defects and inhibition of NER (Bélanger, 2016). Furthermore, ATM loss in colorectal cancer IHC samples has correlated with increased overall survival to first line oxaliplatin chemotherapy treatment, but not to irinotecan further confirming the role replication stress response and NER play in repairing oxaliplatin lesions (Sundar, 2018). Defects in any of the NER pathway factors lead to DNA repair impairment and thus increased genome instability and susceptibility to cancer. Knocking out CSA, which is unique for TC-NER, in colorectal cancer cell lines results in strong sensitivity to oxaliplatin. Similarly, TC-NER patient deficient fibroblasts show a stronger sensitivity to oxaliplatin compared to GG-NER deficient fibroblasts, indicating that the cells rely on TC-NER to repair oxaliplatin lesions (Slyskova, 2018). Thus, both BER and NER protect against oxaliplatin induced cytotoxicity. Since oxaliplatin induces mainly intrastrand crosslinks as its cytotoxic mechanism, they are mainly repaired by NER. Furthermore, human colorectal cancer cells treated with a photosensitizer (hypericin) and oxaliplatin show higher sensitivity to oxaliplatin where hypericin leads to

downregulation of two significant NER enzymes ERCC1 and XPF highlighting the role NER plays in oxaliplatin resistance (Lin, 2016).

1.5.3 Homologous recombination

This type of DDR mechanism repairs DSBs and interstrand crosslinks (ICL) by collaborating with NER and Fanconi anemia (FA) pathway. It is dominant in S/G2 phase of the cell cycle and requires the sister chromatid as template. HR is initiated by chromatin modification which triggers MRE11-RAD50-NBS1 (MRN). The latter has an endonuclease activity and ensures end resection at the break site which involves a 5′-to-3′ nucleolytic degradation to generate 3′ overhangs and exposes long tracts of single stranded DNA (ssDNA). This then activates Tat interactive protein 60 (TIP60) which activates ATM consequently phosphorylating histone family member X (H2AX). ATM also phosphorylates MDC1 which brings in E3 ubiquitin ligases RNF8 and RNF168. The latters ubiquitinate H2AX leading to BRCA1 recruitment activating downstream CtBP interacting protein (CtIP) which leads to activation of replication protein A (RPA) which coats the 3′ overhangs and is then displaced by role of radiation protein 51 (RAD51). The ssDNA then searches for a homologous double stranded DNA (dsDNA) sequence as a template for repair and DNA synthesis(Ranjha, 2018). HR has many modulators and factors playing role to repair the break and ensure genome stability (Chatterjee & Walker, 2017; Stracker & Petrini, 2011; Sun, 2005; Yu, 2006). The extended DNA end resection makes the break not suitable for ligation and promotes HR over non-homologous end joining (NHEJ). Mutations in *BRCA1* along with defective HR has been reported as a risk factor for colorectal cancer (Oh, 2018). Furthermore, there has been reported evidence of *ATM* and *PALB2* as colorectal cancer risk genes, highlighting the importance of the HR pathway in colorectal cancer (AlDubayan, 2018).

HR has also been linked to oxaliplatin where it has been reported that lesions caused by oxaliplatin are repaired through several DDR mechanisms including HR. Patients having mutated HR repair genes showed a better survival upon oxaliplatin treatment compared with wild type HR genes patients (Kondo, 2018). Compared with irinotecan-based treatment (FOLFIRI), patients with DDR mutated colorectal cancers (*BRCA1, BRCA2, ATM,* BRCA1 associated ring domain 1 (*BARD1), CHEK1, CHEK2, RAD51B, RAD51C, RAD51D* or *RAD54L*; all HR genes) receiving FOLFOX treatment have shown improved overall survival and higher response rate (Marks, 2021). Further studies on ATM deficient colorectal cancers treated with oxaliplatin chemotherapy have shown improved overall survival results confirming the role of HR in repairing interstrand crosslinks caused by oxaliplatin (Bakkenist, 2018). In the CRISPR/Cas9 loss of function screen reported previously, top hit genes fell into BER (*POLB,* DNA polymerase lambda *(POLL)* and *XRCC1*), NER (*CSA, XPF* and *ERCC1*), FA (*FANCA,* Fanconi anemia core complex-associated protein 24 *(FAAP24),* structure specific endonuclease subunit 4 (*SLX4), XPF* and *ERCC1*), HR (meiotic recombination 11 homolog A (*MRE11), RAD50, RAD51, BRCA1* and *BRCA2*) and translesion synthesis (TLS) (*REV3*) pathways, showing that oxaliplatin lesions are processed by multiple DDR pathways (Slyskova, 2018).

1.5.4 Fanconi anemia

The minor lesions generated by oxaliplatin are interstrand crosslinks, which are repaired by the coordinated action of Fanconi anemia (FA), NER, and HR proteins. FA pathway has 19 proteins (from FANCA to FANCT) and several other factors that play a role in other DDR pathways, particularly HR. Stalled replication forks due to an interstrand crosslink are primarily detected by FA pathway proteins (FANCM,

FAAP24, Mph1-associated histone-fold 1 (MHF1), MHF2 complex). Several FA proteins propagate the signaling cascade for unhooking of the interstrand crosslinks by DNA incisions by endonucleases (ERCC1/XPF, ERCC4, musculus 81 (MUS81)) which create a DSB. TLS fills the gap opposite to the unhooked crosslink which makes a perfect template for HR to repair the DSB created by the incisions. The unhooked crosslink is repaired by NER (Ceccaldi, 2016; Slyskova, 2018). The involvement of several DDR pathways in repairing oxaliplatin lesions justifies the increase of sensitivity of cells to oxaliplatin when knocking out FA (*FANCA, FAAP24, SLX4, XPF and ERCC1*) and TLS genes (reversionless 3 (*REV3*)) (Slyskova, 2018).

1.6 Prognostic and predictive biomarkers

With the advancement of therapeutics in the cancer field, response to treatment and survival has improved dramatically, yet there are still cases where patients do not benefit from therapy and have side effects and toxicity. This fact has elicited interest in biomarkers and their role in improving outcomes by selecting patients for different treatments. A biomarker is an indicator, that can be measured or assessed objectively, reflecting a normal biological activity, pathogenic process, or pharmacological response to a certain treatment. A prognostic biomarker is one that gives an indication of the patient's cancer outcome irrespective of any therapeutic regimen. The presence or absence of this prognostic biomarker can help clinicians to identify poor risk patients who might therefore be given treatment but does not indicate whether a certain treatment or any treatment might be effective. A predictive biomarker is an indicator of a certain therapeutic intervention effectiveness and in some cases a predictive biomarker can be a target for therapy (Oldenhuis, 2008). The standard of care for stage III colorectal cancer is surgery followed by oxaliplatin based adjuvant chemotherapy. 50% of patients in stage III will be cured by surgery, 20% will benefit from chemotherapy post-surgery and 30%

will relapse after few years. Thus, 80% of patients are being exposed to toxicity from oxaliplatin with no benefit (Auclin, 2017). Thus, the search for biomarkers in colorectal cancer has always been of high significance to improve patient selection and outcome.

The first prognostic biomarker in colorectal cancer is MSI. MSI is found in around 15% of all colorectal cases. 3% of this 15% are attributed to the hereditary cases caused by Lynch syndrome (HNPCC). The remaining 12% is seen in sporadic cases mainly due to the hypermethylation of the promoter of the *MLH1* gene, as discussed earlier. MSI status is highly dependent on stage, thus in stage II and III, MSI is 15% prevalent while it is only 4% - 5 % prevalent in stage IV. MSI high patients are frequent in the right colon, are associated with younger age and show high immune infiltrate with poor differentiation. This MSI high phenotype is associated with better prognosis in early stage disease (Lynch & De la Chapelle, 2003). Mismatch repair status testing has been crucial to stage II patients where there have been suggestions that MSI high stage II patients do not require chemotherapy as they have a better prognosis and no benefit from 5-FU treatments (Guastadisegni, 2010; Hutchins, 2011). Two large meta-analysis have had contradictory results regarding the prognostic value of MSI. One has cast doubts on MSI as a decision factor for adjuvant chemotherapy for stage II patients and found no significance between MSI and survival (Gkekas, 2017), while another had re-established the prognostic power of MSI in stage II patients (Petrelli, 2019). The second biomarker in colorectal cancer is KRAS mutational status. As discussed earlier, *KRAS* is mutated in 45% of metastatic cases and 15% - 37% in primary tumours and is more prevalent with MSS status. *KRAS* mutations have been associated with poor prognosis and worse survival in *BRAF* wild type MSS colorectal cancer stage III patients (Bokemeyer, 2012; Imamura, 2012). Similarly, data show that patients with mutated *KRAS*, particularly codon 13, and *BRAF* are associated with poorer survival in metastatic setting (Modest, 2016). *KRAS* wild type

metastatic colorectal cancer cases are put on anti-EGFR therapeutics (cetuximab, panitumumab) as it has shown to improve survival and disease progression, thus *KRAS* mutational status has proved to be an independent predictive biomarker for anti-EGFR treatment response (Lièvre, 2008). Another prognostic marker is *BRAF* mutation which occurs in around 10% of colorectal cancers. Most *BRAF* mutations are in codon 600 (V600E). It is now recommended that this mutation is tested in MSI patients that have MLH1 loss to rule out Lynch syndrome (Lynch & de la Chapelle, 2003). BRAF V600E mutation is more associated with the female gender, MSI status, high grade histology, the right colon, and lymph node involvement (Clancy, 2013). *BRAF* mutation in MSS stage III and IV is associated with worst prognosis and shorter survival (Bokemeyer, 2012). Furthermore, the presence of BRAF mutation along with KRAS wild type leads to lack of response with anti-EGFR therapy in a metastatic setting (de Roock, 2010; di Nicolantonio, 2008). BRAF mutations can be inhibited through selective BRAF inhibitors. These agents showed significant activity in melanoma, in which BRAF mutations are highly prevalent, but initial studies of BRAF inhibitor monotherapy in colorectal cancer were negative. The lack of response to anti-EGFR in the presence of BRAF mutation has been linked to feedback signaling network activating EGFR and consequently re-activating MAPK signaling. Thus, metastatic patients with V600E mutation have been treated with combination therapy of BRAF inhibitor (encorafenib), an anti-EGFR inhibitor (cetuximab), and a MEK inhibitor (binimetinib), a regimen called triplet therapy lacking chemotherapy. This regimen is a breakthrough in colorectal cancer metastatic settings where it is targeted therapy lacking chemo in a difficult subpopulation of patients (Kopetz, 2019). Thus, further biomarker research and identification is of high importance in colorectal cancer to improve patient outcome. The fact that 80% of stage III patients do not benefit from oxaliplatin based chemotherapy, and the fact that it is, to

date, the standard of care for this patients' group, highlights a crucial need for biomarker research to predict outcome and consequently stratify patients to this treatment. A randomised control clinical trial with an investigation arm having oxaliplatin and a control arm void of oxaliplatin, like the one explored in this thesis, is ideal to find a predictive biomarker for treatment because interaction analysis between both arms is feasible.

1.7 Endpoints for biomarker testing

An endpoint is an event or outcome that can be assessed objectively to check if an intervention or treatment has had any impact on the disease. In oncology, the true clinical endpoint is overall survival (OS) because it reflects the ultimate goal of the treatment which is extending patient life and it is completely unbiased since it is clear from patient survival status and date of death. However, this endpoint is not ideal for testing treatment effect or biomarker significance or its effect because this endpoint can be influenced by second line treatments and other confounders like age and noncancer deaths. OS is measured from date of patient inclusion to date of death irrespective of the cause of death which in some instances can be cancer unrelated. Another end point is objective response rate (ORR) which measures tumour regression upon treatment (Piedbois & Buyse, 2008). In colorectal cancer, tumour responses are evaluated after 8-12 weeks of treatment and follow the response evaluation criteria in solid tumours (RECIST). It specifies four categories: complete response (CR) when the target lesion completely disappears with any lymph node reduced to <10mm. Partial response (PR) is when the tumour has decreased by at least 30% in diameter in reference to the baseline

tumour/lesion size. A response is called as progressive disease (PD) if the tumour has increased by 20% (with an absolute value of at least 5mm) or if there is the appearance of one or more new lesions. Finally, a response to treatment is called stable disease (SD) if there is not sufficient reduction to call it PR neither enough increase to call it PD (Eisenhauer, 2009). Studies have shown that ORR did not correlate strongly with OS, and a good response to treatment does not translate all the time to a better overall survival and prolonged life because large differences in response are needed to predict survival benefit, making ORR not an ideal endpoint for reflecting OS (Johnson, 2006). A third end point is progression free survival (PFS) that is measured from patient inclusion until the date of first disease progression. This endpoint has been considered the primary endpoint in the majority of researches because PFS correlated strongly with OS. Delaying disease progression results in better survival because small differences in PFS are reflected in better survival (Tang, 2007).

The aim of this project is finding a marker predictive to oxaliplatin resistance. Thus, the rationale is to select a resistant subgroup showing a biology that could be used to predict such resistance in other cohorts. The primary end point that is selected therefore for this thesis is response. Response gives the most optimal view on oxaliplatin resistance since it gives a more direct measure of drug effect compared to survival endpoints where several confounders may result in background noise already in the variable that we want to predict. For example, early events shown in PFS data may not always be directly linked to lack of effects of the treatment, and conversely late events may not always be linked to a good effect uniquely due to the treatment. In addition, it is clear how the PD subset of patients are so different from other responders (Figure 1.6). They are clearly not being benefited at all by the treatment while the others show at least some benefit. PD patients

are an optimal endpoint, in particular, when the approach is assessing patients only from the oxaliplatin arm.

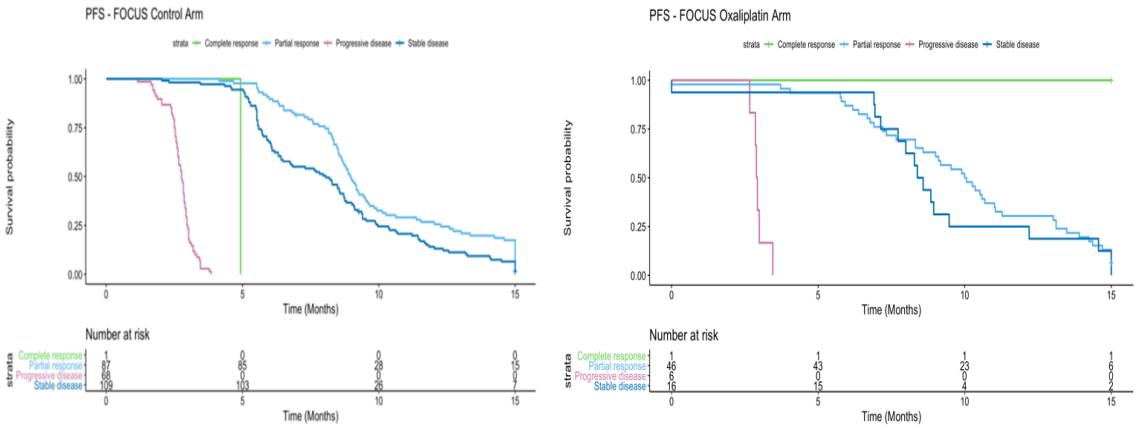


Figure 1.6 Progression free survival in the FOCUS trial. Progressive disease patients show no benefit from treatment regardless of the trial arm and treatment regimen. In the control arm of the FOCUS trial (discussed in next sections in details), the treatment regimen is 5 fluorouracil with folinic acid FUFA, PD patients show no benefit from treatment assessed by progression free survival. Similarly, in the oxaliplatin arm where FOLFOX is given as a treatment, PD patients show no benefit. Figure generated by Enric Domingo at S:CORT.

1.8 S:CORT original hypothesis

The S:CORT (Stratification for COloRecTal cancer) consortium is funded by MRC and CR-UK to assess DNA mutations, morphology, RNA expression, immunohistochemistry and DNA methylation profiles in 2000 colorectal cancer patients. S:CORT was established to identify new molecular methods and discovery of new molecular biomarkers for patient stratifications and prediction of colorectal cancer outcome to treatments. The most defective DNA damage pathway in colorectal cancer is MMR, thus, the primary hypothesis of S:CORT was (DNA damage immune response) DDIR signature would be predictive of oxaliplatin outcome in colorectal cancer as it has

been proven to be predictive to DNA damaging therapy in breast cancer. The DDIR signature, previously developed in breast cancer, is composed of a 44 gene expression signature related to DNA damage response pathways. It has been predictive of response to DNA damaging (anthracycline and cyclophosphamide) chemotherapy based on signaling associated with defective homologous recombination (Mulligan, 2014). In metastatic setting, there was no prognostic effect of DDIR status for patients treated with 5FUFA alone neither on OS nor on PFS. Interaction analysis between the oxaliplatin arm and the control arm showed DDIR positive patients did not benefit from oxaliplatin treatment as it would have been expected. In primary tumours, a similar result of insignificance of DDIR status in relation to oxaliplatin treatment has been reported. However, a trend has been noted in survival analysis both in early stage and in metastatic setting that DDIR negative status has appeared to respond more frequently to oxaliplatin based chemotherapy. In the metastatic setting, DDIR negative responded better to oxaliplatin with response as an endpoint (Malla, 2021). This is contrary to the original hypothesis proposed by S:CORT investigators, and to what has been shown in breast cancer, that DDIR positive is the group that responds to platinum-based chemotherapy. Since the original hypothesis of DDIR status being a predictor to oxaliplatin outcome did not validate in colorectal cancer, this thesis aims at identifying a biomarker that can determine which patients are resistant to oxaliplatin treatment. To that end, three clinical trials are explored: the FOCUS trial that encompasses patients with transcriptomic profiles randomised to oxaliplatin compared to standard chemotherapy (5-FU with folinic acid), FOxTROT that comprises patients treated with pre-operative oxaliplatin, and COIN that encompasses patients randomised into three arms: oxaliplatin-based chemotherapy (oxaliplatin plus capecitabine or oxaliplatin plus fluorouracil and folinic acid) or continuous chemotherapy plus cetuximab or intermittent chemotherapy.

1.9 Exclusion of MSI

The FOxTROT (Fluorouracil and Oxaliplatin With or Without Panitumumab In Treating Patients with High-Risk Colon Cancer That Can Be Removed by Surgery) trial is an international trial that encompasses more than 1,000 patients. The rationale behind the trial is to check whether six weeks of preoperative combination chemotherapy improves the prognosis of patients with high risk operable disease. It also assesses whether this prognosis can be further improved in KRAS/NRAS wild type patients by adding the anti-EGFR monoclonal antibody, panitumumab. The inclusion criteria encompass patients who had been diagnosed with an early stage CRC (T3/T4), with no metastasis detected and have their surgery for primary tumour resection planned. Patients are randomised into three treatment arms; Arm A: six weeks of oxaliplatin plus fluorouracil or oxaliplatin with capecitabine pre-surgery and then four to eight weeks after the surgical resection, patients are put on six or 18 weeks of chemotherapy again. Arm B, patients are assessed for *RAS* mutations and those who are *RAS* wildtype receive the same regimen as arm A but with the addition of panitumumab for the 6 weeks before surgery. Finally arm C, patients are not given chemotherapy before surgery, but they have 12 or 24 weeks of chemotherapy after their surgical resection (Seymour & Morton, 2019). In FOXTROT, the response criteria for histological sections followed Mandard Tumour Regression Grade where pathologists call complete regression if there is a complete absence of residual cancer, marked regression is defined by the presence of fibrosis with

rare residual cancer cells, moderate regression is defined by an increased presence of residual cancer cells but still high fibrosis, mild regression is when the residual cancer cells are higher than the fibrosis observed and finally no regression when there is no regression changes noted (Mandard, 1994).

In the FOxTROT trial it has been suggested that the MSI group which is immune rich and defined by loss of MMR machinery, fails to gain any benefit from oxaliplatin based chemotherapy (Figure 1.7). Prior data have shown that fluoropyrimidines alone are ineffective as adjuvant therapy for CRC with deficient MMR. However, an important individual patient meta-analysis from 12 adjuvant trials have shown that the addition of oxaliplatin to fluoropyrimidines in MSI stage III colorectal cancer prolongs overall survival and improves disease free survival (Cohen, 2021). Emerging data now show that MSI-H tumours respond better to checkpoint inhibitors and immune therapy, most reasonably due to the fact that they have high mutational load and high immune infiltration. A monoclonal anti-PD1 antibody (pembrolizumab) has been approved for MSI-H in all kind of cancer. Nivolumab and Ipilimumab are also used for stage IV patients of any cancer making MSI the first biomarker for therapy selection irrespective of cancer type (Bilgin, 2017). While it is clear that in stage III cancer oxaliplatin is effective in MSI patients, in this thesis, my analysis for identifying an oxaliplatin resistant biomarker has been conducted on MSS patients only in light of the changing paradigm in which MSI patients are treated increasingly with immunotherapy.

91% scored blind by central pathologist 9% by local pathologists	pMMR (MSS) n=592	dMMR (MSI) N=106	
Complete Response (TRG4)	3.3%	4.7%	p<0.0001 MH
Marked Regression (TRG3)	4.8%	0%	
Moderate Regression (TRG2)	14.5%	0%	
Little Regression (TRG1)	47.9%	21.7%	
No regression (TRG0)	26.6%	73.6%	

Figure 1.7 FOxTROT assessment of response to oxaliplatin in MSS versus MSI patients. The MSI group which is immune rich and defined by loss of MMR machinery, fails to gain any benefit from oxaliplatin based chemotherapy. Figure taken from (Seymour & Morton, 2019).

1.10 Aims of the project

Despite the improvement in outcome seen in advanced disease when treated with oxaliplatin, a high percentage of cases show lack of response and resistance while bearing all the side effects of this chemotherapeutic drug. Thus, my aim in my work for this thesis is to discover or validate a biomarker of oxaliplatin response with potential use in the clinic to guide therapy decisions. In order to investigate this, I will first test pre-specified biomarkers presented in section 1.4 of this introductory chapter and check their role in oxaliplatin outcome. I will test those candidate biomarkers in the FOCUS trial using logistic regression with binarised response to treatment as an endpoint. The FOCUS trial is a randomised control clinical trial with an investigation arm having oxaliplatin and a control arm void of oxaliplatin (5FUFA as treatment regimen). The pre-specified candidate biomarkers will be tested using regression in the control arm first to check if any of those candidate biomarkers are prognostic for 5FUFA, then in the oxaliplatin arm to check if any of the pre-specified candidate biomarkers are prognostic for oxaliplatin. Finally, I will perform interaction analysis between both arms to check if any of the candidates is a predictive biomarker to oxaliplatin. If any of the biomarkers is significant in the previous analyses, I will perform survival analysis using cox regression to check the significance translates to progression free survival.

Secondly, I will investigate the role of candidate biomarkers that I selected from a systematic literature review that I performed to find biomarkers of resistance to oxaliplatin that has been reported in the literature in the last five years. A similar analysis stated previously will be followed in this second aim of the project.

Finally, I will extract RNA from FFPE blocks for 3'RNASEQ from the COIN clinical trial where oxaliplatin based chemotherapy is given to patients. I will then

conduct differential gene expression analysis from COIN clinical trial using DESEQ2. The differentially expressed genes will be used to derive a transcriptomic signature that will be tested in the FOCUS and FOxTROT trials. Similarly, a second signature will be derived from the FOCUS microarray data using LIMMA and this second signature will be validated in the COIN and FOxTROT trials.

The data generated in this thesis will give an insight to potential biomarkers predictive of oxaliplatin response.

Chapter 2 Role of Pre-specified Biomarkers in Stratifying Patients to Oxaliplatin

This chapter aims at examining whether 13 prespecified biomarkers known for their prominent biology in CRC may be useful to stratify patients for oxaliplatin treatment in metastatic CRC. The selection of biomarkers is based on biology, statistical power, control for multiple testing and considering all types of data available (clinical, DNA sequencing, DNA methylation, RNA expression and protein expression). In order to investigate those candidate biomarkers, I will be using the FOCUS trial clinical and transcriptomic data. The analysis in this chapter and all the chapters in this thesis are conducted on samples excluding MSI-H cases.

2.1 Methods

2.1.1 FOCUS

FOCUS is a UK-based randomised controlled clinical trial that enrolled 2135 patients with advanced colorectal cancer between May 2000 and December 2003. The inclusion criteria encompassed patients with inoperable confirmed colorectal adenocarcinomas who had not received previous treatment for advanced disease. Patients were divided into three treatment regimens; strategy A was the control group, administering fluorouracil (5-FU) and folinic acid (FA) using the bolus and infusion regimen developed by de Gramont (Cheeseman, 2002) alone as first line of treatment until failure, then the second line of treatment was single agent irinotecan. Strategy B enrolled patients administering them on again on 5FUFA alone as first line of treatment but then patients fit to receive second line as combination therapy were divided further into two arms: B-ir where irinotecan was added to fluorouracil (the FOLFIRI regimen)

and B-Ox where oxaliplatin was added to fluorouracil (the FOLFOX regimen). Finally, strategy C administered first line of treatment as combination therapy and divided into two arms as well: C-ir where irinotecan is given with fluorouracil (FOLFIRI) and C-Ox giving oxaliplatin with fluorouracil (FOLFOX) (Seymour, 2007).

For the candidate analysis, formalin-fixed paraffin imbedded (FFPE) blocks from FOCUS patients with colonic primary tumours from a biobank of archival diagnostic tissue were selected from consenting patients in the relevant arms where a randomised comparison could be made between first-line 5FUFA alone (from arms A and B) or in combination with oxaliplatin from arm C-Ox (85 mg/m^2 two-weekly; Figure 1).

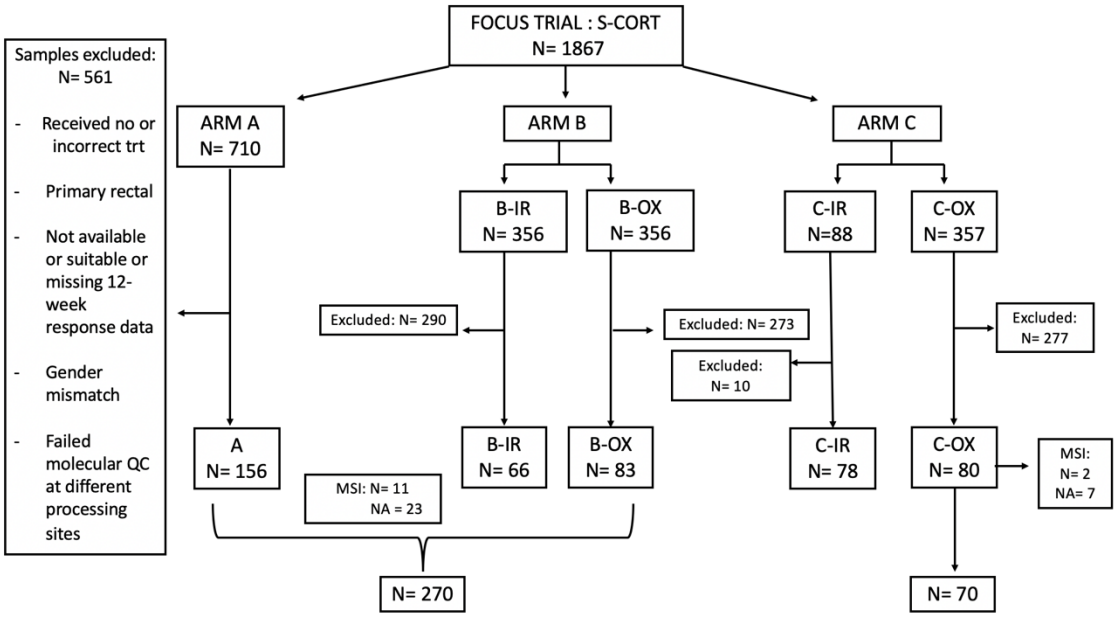


Figure 2.1 FOCUS consort diagram. The number of samples in the FOCUS trial in SCORT was originally 1867 divided into three arms. Arms A and B have 5FUFA as first line chemotherapy treatment and are considered the control arm in the analysis in this thesis. C-Ox is the treatment arm where combination therapy of FOLFOX is administered. The exclusion criteria are patients who did not give explicit consent to further bowel cancer research, those who did not receive treatment or received incorrect treatment, those with rectal primaries, those with missing response data, samples with gender discrepancy between clinical data and the molecular data for calling the gender, or finally samples that have failed molecular quality control. The whole analysis has been conducted on MSS samples and thus MSI or unavailable microsatellite status are both removed to end up with N=270 in the control 5FUFA arm and N= 70 in the oxaliplatin arm.

2.1.2 Response criteria in FOCUS

In the FOCUS trial the primary outcome was overall survival, with progression free survival and response as secondary endpoints. The response criteria used for CT

scans followed RECIST 1.1 version (The Response Evaluation Criteria for Solid Tumours). Complete response (CR) was called when the target lesion completely disappears with any lymph node reduced to <10mm. Partial response (PR) is when the tumour has decreased by at least 30% in diameter in reference to the baseline tumour/lesion size. A response is called as progressive disease (PD) if the tumour has increased by 20% (with an absolute value of at least 5mm) or if there is the appearance of one or more new lesions. A response to treatment was called as stable disease (SD) if there is not sufficient reduction to call it PR neither enough increase to call it PD (Eisenhauer, 2009).

2.1.3 RNA extraction

All wet-lab work from SCORT was processed using Standard Operation Procedures (SOPs) carefully developed in different UK laboratories with strong expertise. For the FOCUS trial, formalin fixed paraffin embedded (FFPE) blocks containing CRC were identified, sectioned and processed in University of Leeds pathology department (Prof Phil Quirke, Susan Richman). Presence of tumour in the sample was confirmed and the samples sections as illustrated below (Figure 2.2).

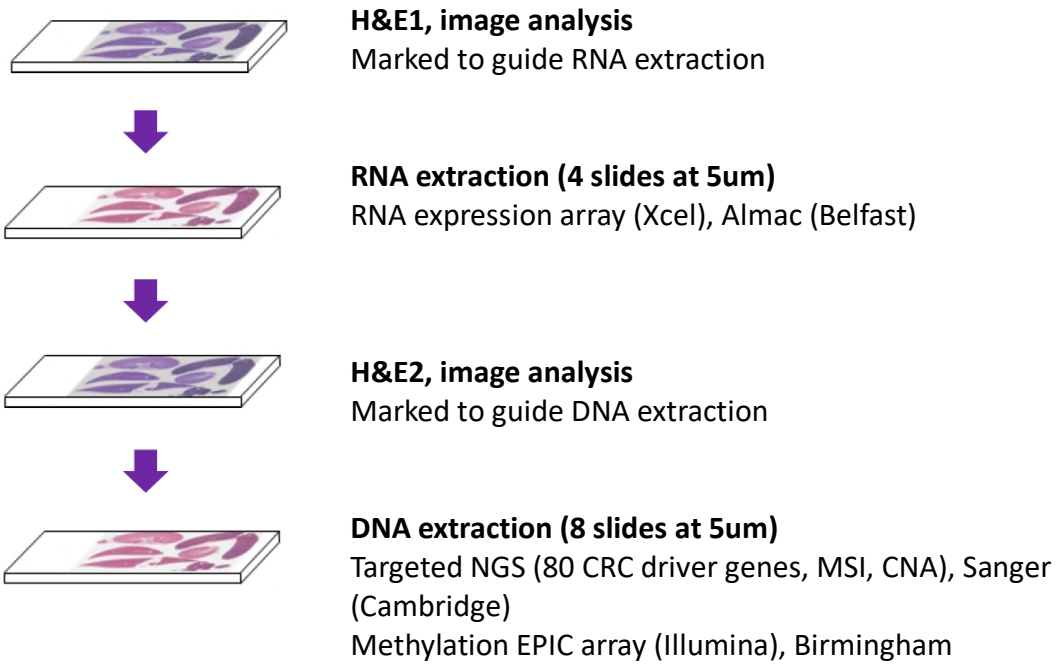


Figure 2.2 SCORT sample preparation. FFPE blocks containing tumours were identified, sectioned and stained with haematoxylin and eosin 1 (H&E1). 4 slides were prepared for RNA extraction each with 5 µm thickness. RNA expression profiling was done on Xcel array by Almac in Belfast. For DNA extractions, staining by H&E2 was done to prepare 8 slides at 5 µm for targeted next generation sequencing done at the Sanger institute at Cambridge and for methylation Epic array done on Illumina at Birmingham.

At Leeds University sample processing centre, 1 x 5 micron and 8 x 5 micron sections were prepared, by SCORT personnel, from all samples which had passed initial histology QC. The 5 µm section was H&E stained followed by annotation of the tumour region by a qualified pathologist. The H&E stained 1 x 5 micron and 2 x 10 micron sections were shipped at room temperature to Queens University, Belfast. Upon sample receipt, the material transfer agreement was completed and returned to Leeds, specimens were registered on a Queens SCORT processing database followed by registering all Human Tissue Act (HTA) relevant material on the Queens HTA database. Samples were dewaxed and macrodissected following SCORT-SOP-002-Macrodissection-and-RNA-Extraction. Dewaxing of samples was performed by an automated dewaxing process using the Tissue-Tek Prisma®/Film machine. This incorporated 3 xylene washes of 2, 2

and 1 min followed by 2 washes in 95% and 90% ethanol respectively. Macrodissections were performed within a designated macrodissection area treated with an RNase Decontamination Solution. The unstained section was overlaid on top of the H&E reference slide and using a clean scalpel blade the annotated area was scraped into a 1.5 ml RNase-free Eppendorf tubes containing the prepared Roche High Pure RNA Paraffin Kit tissue lysis reagent (catalogue 03270289001). The Roche High Pure RNA Paraffin Kit instructions for use (version 12) were followed for RNA extractions. Following RNA extraction, SCORT-SOP-003- Nucleic Acid Quantification was followed detailing quantification of the total RNA using the Nanodrop set to the RNA-40 parameter. Total RNA extraction, quantification and -80^{o}C storage location details were logged onto the SCORT Specimen Receipt Records spreadsheet. After RNA quantification, samples which had sufficient RNA concentration (≥ 10 ng/µl) were submitted for RNA profiling.

2.1.4 Gene expression profiling

FFPE is a well-established method for histopathological assessment but results in extensive RNA fragmentation and cross-linking, impacting on RNA profiling. Due to FFPE block processing, specific mRNA transcripts are more susceptible to degradation and the 5' end of the mRNA transcript is more sensitive to FFPE-dependent RNA degradation. To harness the power of large archived clinically annotated cohorts of FFPE tumour blocks, which is central to the SCORT proposal, FFPE-specific in vitro transcribed (IVT) reagents were utilised and the Almac Xcel Array. These arrays have probe designs focused on the extreme 3' of mRNA (Affymetrix Almac Xcel Array), as these regions are more stable in FFPE tissue. Furthermore, IVT reagents specifically designed for FFPE-extracted RNA have been developed including both oligo-dT primers (binding to polyA tails and specific for mRNA) and random hexamers (annealing

throughout the target molecule) for first strand cDNA synthesis. This enables amplification at the 3' end and throughout the whole transcript which is more applicable to partially degraded mRNA. The SCORT-SOP-004-3' IVT Pico Kit which details the GeneChip® 3' IVT Pico Reagent Kit (Part No. 703308; Version 1) instructions for use was followed. SCORT-SOP-005-Cartridge Array Hybridisation & Sample Registration was followed to hybridise the fragmented and labelled cDNA to the microarrays. Following a 16-hour overnight hybridisation the arrays were washed and stained following the SCORT-SOP-006- Fluidics_Start Up_ Wash_Stain_ShutDown. SCORT-SOP-007-Scanning of Array and Storage of CEL Files was performed to scan the hybridised arrays and to ensure the resulting image (CEL) files were securely backed up offsite to eliminate the risk of catastrophic data loss. The protocol entitled 'SCORT-SOP-008-Analysis of CEL Files to Determine QC Metrics was followed to assess QC metrics relating to monitor image quality, IVT, hybridisation to the array and RNA degradation prior to uploading to the Oxford server.

2.1.5 Next generation sequencing

DNA was extracted from FFPE samples, by SCORT personnel, at the University of Leeds and shipped to the Wellcome Trust Sanger institute where three quality control checks were performed: volume checks, picogreen quantification and fluidigm genotyping to assess sample integrity. After QC assessment, samples which passed minimum DNA concentration (total DNA >=100ng) and genotyping were submitted for library preparation and Illumina sequencing. Library preparations were performed using the Custom SureSelect Library Prep Kit from Agilent Technologies - Part Number 930075. Libraries were submitted for in-solution capture hybridisation using a custom SCORT colorectal V2 gene bait set, targeting all coding exons of 80 CRC driver genes

(Table 2.1), 51 regions of recurrent copy number gains/losses (with deeper coverage of 22 cancer genes present in these regions), and 120 MSI regions (mononucleotide repeats). DNA amplification was then performed for eight cycles before sequencing. Sequencing was performed on Illumina HiSeq2000 machines. The Illumina sequencing data BAM files were transferred to a dedicated variant calling pipeline developed by the Cancer Group Project (CGP) team in the Sanger Institute for sequence alignment and data analyses. Variant calling algorithms were first run on the data for identification of base insertion/deletions and substitutions in the data using Pindel as a bioinformatic tool to identify somatic indels (insertions and deletions) (https://github.com/genome/pindel) and Caveman as another tool to identify somatic single nucleotide variants (https://github.com/cancerit/CaVEMan) . The data was then flagged for variants present only in the list of the cancer genes of the custom bait design as well as for variants seen in germline, FFPE normal samples to remove sequencing artefacts. Further analysis of the data involved flagging for non-synonymous coding variants and driver mutations and also identification of microsatellite instability (MSI). Driver mutations were defined according to whether any given gene was dominant or recessive. For the former, identification was based on comparisons of the SCORT sample variants with recurrent mutations previously identified in a statistical analysis of 11,119 human tumours, spanning 41 cancer types (MT, 2016) as well as internally curated driver mutations from the Cancer Genome Project. For recessive genes, any truncating mutation (frameshift, nonsense and essential splice) was included as a putative driver mutation. The Wellcome Trust Sanger Institute provided the following data to the Oxford repository: BAM files (raw sequencing files aligned to the human reference genome), vcf files (raw calls for variants), flat file showing high quality variants with annotations (e.g. flagged for driver mutations), classification of tumours for microsatellite instability (MSI), and a summary

of coverage per tumour. These files were uploaded via the Globus secure file transfer system and stored on the secure SCORT fileserver. The original files were stored separately from the decrypted and uncompressed versions of the data to ensure data integrity at every stage of processing and the high-quality variants were loaded into the secure MYSQL database held at Oxford. Data for coverage and allelic frequency of each mutation have been added to the database from the vcf files with an R script. Each mutation call and associated individual reads can be visualised from the BAM files using a genome viewer such as IGV.

Genes_v2						
ACVR1B	BCL9L	ELF3	HLA-C	MET	PMS2	SOX9
ACVR2A	BMPR2	EP300	HNF4A	MLH1	POLE	TCF7L2
AKT1	BRAF	ERBB2	HRAS	MSH2	PPP2R1A	TGIF1
AMER1	BUB1B	ERBB3	IDH1	MSH3	PTEN	TP53
APC	CASP8	FAM123B	IGF2	MSH6	RAF1	UBR5
ARID1A	CD58	FBXW7	IKBKB	MYC	RBM10	WBP1
ATM	CDC27	FGFR3	KAT6A	NF1	RNF43	ZFP36L2
ATP1B4	CDK8	FLT3	IRS2	NRAS	SEMG2	ZNF781
ATR	CDKN2A	GNAS	KRAS	NUGGC	SMAD2	
AXIN2	CDX2	HDLBP	LIFR	PCBP1	SMAD3	
B2M	CREBBP	HLA-A	MAP2K4	PIK3CA	SMAD4	
BCL9	CTNNB1	HLA-B	MBD6	PIK3R1	SMARCA4	

Table 2.1 Sanger SCORT sequencing panel v2. DNA was extracted from FFPE samples for sequencing against a custom SCORT colorectal V2 gene bait set, targeting all coding exons of 80 CRC driver genes shown in this table.

2.1.6 DNA methylation

The Infinium HD FFPE methylation assay combines bisulfide conversion of genomic FFPE DNA and restoration followed by whole-genome amplification with direct, array-based capture. This results in the scoring of over 850,000 CpG loci on the Infinium HumanMethylation EPIC array. Signal intensity is recorded with the Illumina

iScan system generating β-values which give the degree of methylation, covering approximately 99% of RefSeq genes, 96% of UCSC CpG islands, and specific commonly methylated CpG sites in human cancers. One or two probes are used to interrogate a CpG locus, depending on the probe design for a particular CpG site. The Infinium I design has two probes per site and Infinium II has one probe per site. The 3' end of the probes is positioned directly adjacent to the CpG site (for Infinium I) or immediately adjacent to the site (for Infinium II). The Infinium HD methylation assay incorporates both of these chemistries with allele-specific single base extension of the probes utilizing a biotin labelled nucleotide (for C and G bases) or a dinitrophenyl labelled nucleotide (for A and T bases). Signal amplification of the incorporated label further improves the overall signal-to-noise ratio of the assay. DNA samples were sent to Birmingham University from Leeds. The DNA had been quantified by picogreen and normalized to 20 ng/µl. From this 17.5 µl (350 ng) of DNA was processed per sample. This is within the DNA input range Illumina recommend. To determine whether the FFPE DNA samples were suitable candidates for the Infinium HD FFPE methylation assay the quality was tested in duplicate by real-time PCR following the Illumina FFPE QC protocol (Part # 15020981 Rev. C). Amplification of the FFPE sample DNA was compared with the amplification of a Quality Control Template (QCT). The real-time PCR threshold cycle (Ct) was averaged and a DCt for each sample was calculated (CtFFPE - CtQCT). At least 50 ng of FFPE DNA was bisulfide converted using the EZ DNA Methylation Kit (cat# D5002, Zymo Research) following the manufacturer's instructions appropriate for the Illumina HD FFPE assay. Bisulphite conversion results in unmethylated cytosines converted to uracil, and methylated cytosines remaining unchanged. A total of 8 µl of DNA was eluted and taken through to the next stage FFPE DNA Restoration. The Infinium HD FFPE restore protocol, Part # 15014614 Rev C, restores degraded FFPE DNA to a state that is

amplifiable. All eluted restored DNA (approx. 8 µl) was taken through to the Infinium HD FFPE methylation assay. The protocol was carried out according to the Illumina recommendations for FFPE DNA (Part # 15027310 Rev. A). Post hybridization the signal intensity values were obtained by scanning on the Illumina iScan System to generate raw .idat files. All samples were processed and scanned for 12x1 Infinium HumanMethylation450 beadchips. Two quality elements were assessed prior to downstream bioinformatics applying the raw .idat files to Illumina Genome Studio software v2011.1. The first quality element checked assay and sample quality by assessing the signal intensity levels of a series of sample-independent and -dependent probes, the former using spike-in controls added prior to hybridization. Signal intensities for each need to fall within Illumina recommended ranges to determine whether a control probe has failed or not. The second quality element looked at the signal intensity for each of the CpG loci and from this the % detected per sample was calculated at the confidence level p=0.05. The cut-off value is ≥ 95% detection so a sample with less CpG sites detected would be flagged. The University of Birmingham provided raw Idat files to the data repository which were uploaded on the 13th January 2016 using the data upload portal. These files were stored on the secure file server at Oxford with the original files stored separately from the uncompressed version of the data. Quality control was performed using MethylAid by S:CORT personnel (https://bioconductor.org/packages/release/bioc/html/MethylAid.html). Bad quality samples can be detected using sample-dependent and sample-independent controls present on the array. Deep exploration of bad quality samples was performed using several interactive diagnostic plots of the quality control probes present on the array. Furthermore, the impact of any batch effects was also explored.

Plots for Median methylated vs Unmethylated, Overall Quality, Hybridisation, and detection P values were derived. The beta values for each CpG site in every sample was obtained using ChAMP (described in next section) ranging from 0 (unmethylated) to 1 (methylated). These data were further processed by calculating the median value from only island CpG sites within each gene resulting in a single call per gene across the dataset. Both sets of data were loaded into the secure MYSQL database in Oxford.

2.1.7 Immunohistochemistry

Immunohistochemistry was carried out at the University of Leeds and Queens University Belfast. Tissue microarrays (TMAs) for the SCORT FOCUS cohort were constructed using 0.6-mm cores taken in triplicate from epithelial-rich tumour regions in FFPE blocks for each patient using a Beecher manual arrayer (Beecher Instruments Inc., Sun Prairie, Wisconsin, USA). All work on the TMA sections was undertaken blinded to clinical outcomes in the Precision Medicine Centre of Excellence at Queen's University Belfast using standardised operating procedures for immunohistochemical staining, digital slide scanning and digital image analysis to reduce potential sources of bias in data collection. All procedures were reviewed and agreed by senior consultant pathologists (J.J., M.B.L. and M.S.T.). Immunohistochemistry was performed for adaptive immune (CD3, CD4, CD8, CD20 and FOXP3) and immune-checkpoint (ICOS, IDO-1 and PD-L1) biomarkers on either the Ventana BenchMark XT (Ventana Medical Systems, Oro Valley, Arizona, USA) or Leica BOND-MAX (Leica Biosystems, Wetzlar, Germany) automated immunostainers. Multiplex immunofluorescence for CD3, CD4 and CD8 was conducted using opal chemistry on the Leica BOND-MAX (Leica Biosystems, Wetzlar, Germany) automated immune-stainer. All immunostained slides were scanned using a Leica Aperio AT2 at ×40 magnification or an Akoya Vectra Polaris at ×20 magnification

using MOTIF scanning protocols if immunofluorescent stained. Immunohistochemistry for the hypoxia biomarker CAIX on the SCORT FOCUS TMAs was carried out by the Leeds Institute of Medical Research at St James's using the DAKO Autostainer Link 48 (Agilent Technologies, Santa Clara, California, USA). Assessment of all biomarkers was undertaken using open-source software QuPath version 0.2.0.m6. Full-face sections cut at 4 µm were annotated with the assistance of a senior consultant pathologist (M.S.T.). Assessment of the invasive margin on full-face CRC tissue sections was defined as a 500 µm border, taken from the outermost edge of the malignant glands. Multiplex findings were validated on full-face resection specimens of CRC tissue. Assessment of the immune biomarkers was taken as an average over the number of cores available. (adapted from (Craig, 2020)). CD8 IHC is tested in this chapter as a candidate biomarker for oxaliplatin outcome.

2.1.8 Software and regression model

For expression profiling, Almac and Queens provided RNA profiling data as raw CEL files were uploaded to Oxford using the data upload portal on the 17th February 2016. These files were stored on the secure file server at Oxford with the original files stored separately from the uncompressed version of the data. Quality control analysis was run on the samples using the R base AffyQC Module (https://github.com/BiGCAT-UM/affyQC_Module) by the S:CORT team in Oxford. This resulted in a variety of QC metrics and visualisations including box plots to show the distribution of the expression intensity across samples after normalization, PCA analysis, probe density histograms, chip hybridisation visualisation, normalised unscaled standard errors and RNA degradation plots. The files were processed using the R libraries limma (http://bioconductor.org/packages/release/bioc/html/limma.html) and affy

(http://bioconductor.org/packages/release/bioc/html/affy.html) to normalise the expression values using Robust Multi-array average (RMA) which a method for normalisation (https://felixfan.github.io/RMA-Normalization-Microarray/) and generate an expression matrix of probe intensities against the samples. This data was then used to generate an expression mean per gene data matrix giving a single expression value per gene, and both by probe and mean per gene data were then loaded into the MYSQL database at Oxford.

For DNA methylation, the University of Birmingham provided raw Idat files to the data repository which were uploaded on the 13th January 2016 using the data upload portal. These files were stored on the secure file server at Oxford with the original files stored separately from the uncompressed version of the data. Quality control was performed using MethylAid (https://bioconductor.org/packages/release/bioc/html/MethylAid.html). Bad quality samples can be detected using sample-dependent and sample-independent controls present on the array. Deep exploration of bad quality samples was performed using several interactive diagnostic plots of the quality control probes present on the array. Furthermore, the impact of any batch effects was also explored. Plots for Median methylated vs Unmethylated, Overall Quality, Hybridisation, and detection P values were derived. The beta values for each CpG site in every sample was obtained using ChAMP (https://www.bioconductor.org/packages/release/bioc/html/ChAMP.html) ranging from 0 (unmethylated) to 1 (methylated). This data was further processed by calculating the median value from only island CpG sites within each gene resulting in a single call per gene across the dataset. Both sets of data were loaded into the secure MYSQL database in Oxford by S:CORT personnel.

For analysis in this chapter, I used R version 3.5.1 for bioinformatic analyses, available for download at https://www.r-project.org/. The type of analysis that I first performed is regression. It is a statistical procedure assessing the relationship between one or more variables and an outcome, in this case our primary outcome is response to treatment. The type of regression performed depends on the outcome data, if the data is continuous, then the method to be used is linear regression, if the outcome data is binary, then logistic regression is performed, and finally if the outcome is time to event then cox proportional hazard is implemented. The regression analysis can be conducted as univariate regression analysis which is studying the relationship between the explanatory variable and the outcome of interest without taking into consideration any other variables. However, multivariate analysis is studying the relationship between several explanatory variables and the outcome of interest with the adjustment of the effect of the other variables on the analysis (Lewis, 2007). I conducted logistic regression analysis with binarised response to treatment (5FUFA arm versus FOLFOX treatment arm) as an endpoint using the function glm (https://cran.r-project.org/web/packages/cranly/vignettes/glms.html) from the stats R library. I scaled all tested variables from 0 to 1 to make them statistically comparable. I conducted cox regression in survival analysis (https://cran.r-project.org/web/packages/survival/survival.pdf) (https://cran.r-project.org/web/packages/survminer/index.html) with progression free survival as an endpoint and calculated statistical significance using cox proportional hazards progression (coxph). I also conducted multivariate analysis on significant candidates that came up from univariate interaction regression analysis to check if adding any of the significant variables to the model of the other one, would affect the significance of any of the significant variables.

2.1.9 Molecular profiling

As stated in the introductory chapter, DDRD is a 44-gene transcriptomic signature derived from breast cancer that is able to identify patients with loss of DNA damage response due to the FA/BRCA pathway deficiency. It has been renamed as DNA damage immune response (DDIR) signature to reflect greater understanding of its immune and IFN-related chemokine signalling (Malla, 2021). DDRD signature scores were calculated and a predefined threshold (0.1094) was assigned to call DDRD positive from DDRD negative groups. Almac diagnostics optimised this threshold in an independent 260 colorectal cancer samples study where this threshold was able to detect the DDRD positive subgroup while having the maximum sensitivity and specificity. Samples > 0.1094 are DDRD positive and samples ≤ 0.1094 are DDRD negative .

For the hypoxia signature, probesets were collapsed to single gene values by using the median of the probesets. Then, gene symbols from the hypoxia signature (Buffa, 2010)were identified and their median calculated in each sample.

The signature for TGF-β activation in fibroblast (Calon, 2012)was used and implemented as originally published. Firstly, probeset level data were collapsed at gene level by using the mean of all probesets linked to a gene. Secondly, gene symbols from the original signature were matched in transcriptome data with z-scores. Finally, the mean of the z-scores in the available signature genes was calculated in each sample.

Methylation array raw data analysed with the EPIC array (Illumina) was processed with the R package ChAMP (Tian, 2017). CIMP classification was generated with a recursively partitioned mixture model as previously done in CRC TCGA (Muzny, 2012) and Guinney et al (Guinney, 2015) with minor changes due to the higher number of probes.

The classification for the Consensus Molecular Subtype (CMS) has been derived with a script based on the R package CMSclassifier (https://github.com/Sage-Bionetworks/CMSclassifier) which has been kindly reviewed by one the original authors of the signature (Aurelien de Reynies). We obtained a sensible frequency of each subtype compared with the original CMS article (Guinney, 2015).

CRIS subtypes were derived using the CRISclassifier R package (https://rdrr.io/github/peterawe/CMScaller/man/templates.CRIS.html) released in the original publication (Isella, 2017). Samples showing BH.FDR>0.2 were called unclassified. Probeset data from the transcriptome had been collapsed at gene level by using the probeset with highest mean expression across the cohort as done on the original publication.

The targeted NGS panel applied to S:CORT cohorts contains probes spanning SNPs evenly distributed along the human genome (average of 1 SNP per 3 Mb) and also 51 chromosomal regions recurrently gained or lost in CRC. This design allows generation of copy number estimations from targeted NGS at low resolution, mindful it may not be comparable to techniques interrogating the whole genome. The quality of the analysis may be further enhanced by using CNVkit, a tool specifically designed to be used on targeted NGS data by analysis of both targeted and off-target reads (Talevich, 2016). Accordingly, CNVkit was applied adjusting by tumour purity assessed with digital pathology. Copy number segments with estimations ≥3 were classified as gain, 2 as neutral and ≤1 as loss. Whole Genome Instability Index (WGII) measuring the proportion of the genome with an aberrant copy number was calculated as the sum of the lengths of calls for either loss or gain divided by the whole length. Chromosomal Instability (CIN) calls were then performed with a threshold of WGII at >0.2 as previously defined (Burrell, 2013).

2.2 Selection of candidates

2.2.1 DNA damage repair deficiency (DDRD)

Interestingly, in breast cancer DDRD was found to be predictive of response to anthracycline and cyclophosphamide–based chemotherapy as patients with DDRD+ tumours received no benefit in an untreated cohort but showed strong better survival than DDRD- cases in a separated treated cohort (Mulligan, H, 2014). The signature was recently re-labelled as DNA damage immune response (DDIR) because of the strong presence of an inflammatory tumour microenvironment with high lymphocytic infiltration in these cases.

DDIR ability to predict response to oxaliplatin based chemotherapy was recently tested in two CRC cohorts as it was hypothesised that DDRD+ patients may benefit from addition of the DNA damage agent oxaliplatin, similarly to the observation reported in breast cancer. The first cohort was composed of patients who received 5FUFA (5-FU with folinic acid) with or without addition of oxaliplatin (FOLFOX) within the FOCUS trial and the second were patients who received neoadjuvant FOLFOX within FOxTROT trial. Contrary to the expected result, DDIR positive patients did not predict better outcome following oxaliplatin treatment in CRC while DDIR negative patients had a trend towards improved response for that treatment (Malla, 2021). Considering that DDRD was found to be an immune signature with enrichment for MSI in CRC, we added DDRD in our analyses where MSI cases have been excluded trying to gather further clarity on the unexpected outcome results. DDRD is a continuous variable that was binarised based on the threshold discussed above. DDRD+ was coded as 1 and DDRD- was coded as 0.

2.2.2 Sidedness

The colon is 150 cm long and is comprised of several parts: cecum, ascending colon, transverse colon, flexure, descending colon, sigmoid colon, rectum and anus. The right colon is generally considered to encompass the cecum, ascending and the transverse colon while the left colon encompasses the descending, sigmoid and the rectum (Gervaz, Bucher & Morel, 2004). Tumours occurring in right sided colon show different histological profiles compared with those detected in the left side. Right sided tumours may show serrated and mucinous adenocarcinomas, while left sided tumours have more polypoid morphologies and show tubular and typical adenocarcinomas (Marzouk & Schofield, 2011). Right sided CRC has more advanced, poorly differentiated tumours characterised by microsatellite instability-high (MSI-H) which harbours mutations or inactivation of the mismatch repair system. Conversely, most left sided CRCs show chromosomal instability (CIN) which is characterised by imbalances in chromosome number (aneuploidy) and loss of heterozygosity (LOH) (Pino & Chung, 2010). In addition, there are further distinct molecular features that differentiate right and left colon cancers. Right sided CRC is characterised by, in addition to MSI-H, CpG island methylator phenotype high (CIMP-H), mutations in *KRAS and BRAF-V600E,* transcriptomic consensus molecular subtype 1 (explained later in this chapter) and worse prognosis. Alternatively, left sided CRC is characterised by aneuploidy, mutations in *TP53,* and better prognosis (Domingo, 2013; Puccini, 2018). In addition, embryologically the right and the left colon are very different as they are derived from the midgut and the hindgut respectively. Finally, it has been suggested that clinical trials for advanced CRC should be stratified by sidedness(Kerr, 2016). Accordingly, sidedness was identified as a potential biomarker for prediction to oxaliplatin treatment in metastatic CRC. Sidedness was changed into numerical value for this analysis where left sided tumours were coded as 1 and right sided as 0.

2.2.3 *APC* mutation

The development of CRC can be due to both genetic and environmental factors. The biggest percentage of cases are sporadic while evidence of a family history accounts for 15-20% of cases (Kuipers, 2015). The most common hereditary CRC is Lynch syndrome, and the second is familial adenomatous polyposis arising from germline mutations in adenomatous polyposis coli (*APC*) gene that constitutively activates the Wnt signalling pathway (Vasen, Tomlinson & Castells, 2015). Most patients with this syndrome develop from hundreds to thousands of adenomas in early adulthood which may become malignant if not treated. *APC* is a tumour suppressor gene highly mutated in CRC in around 80% of sporadic CRC tumours (Fearnhead, Britton & Bodmer 2001). In normal conditions, APC forms a complex with Axin/Axin2 and GSK-3β that promotes the ubiquitination and subsequent proteasomal degradation of the oncogene β-catenin in the absence of Wnt signalling. Tumours tend to deactivate *APC* or *GSK3β* or stabilise *CTNNB1* (encoding β-catenin). Loss of *APC* function results in an accumulation of β-catenin, which translocates to the nucleus and engages the Tcf/Lef transcription factor complex to activate transcription of a large number of target genes like cyclin D1 and c-Myc. Wnt signalling allows tumour proliferation and inhibition of differentiation. Negative regulation of Wnt signalling happens when *CK1* and *APC/axin/GSK-3β* complex targets β-catenin for degradation (Noubissi, 2006). Due to its prominent role in CRC tumourigenesis, *APC* mutation was selected as one of the possible candidates for testing as a predictor for oxaliplatin response. *APC* mutation was coded as 1 and wild type as 0.

2.2.4 *TP53* mutation

In colorectal cancer, *TP53* is the second mutated driver gene after *APC*. It's the most frequently mutated gene in all tumour types. *TP53* regulates cell cycle arrest and cell death checkpoints and at normal conditions, the cells express low levels of *TP53* with its degradation tightly monitored by several ubiquitin ligases (Morikawa, 2012). As mentioned earlier, *TP53* mutations are more frequent in left sided than right sided colorectal tumours. These mutations are considered a late event in the progression of adenoma to carcinoma tumourigenesis and have been tightly associated with chromosomal instability in CRC and other tumour types as they probably result in tolerance to chromosome missegregation. In the past *TP53* mutations have been associated with both good (Lan, 2007; Noske, 2009; Soong, 1997) and bad prognosis (Auvinen, 1994; Elsaleh, 2001; Lanza, 1996; Morikawa, 2012; Pancione, 2010) in the literature. However, more recent results from large clinical trials suggest *TP53* mutations may result in poorer prognosis alone or in combination with Ras mutations (Domingo, 2018; Seligmann, 2021). Thus, *TP53* mutation was an obvious marker to be selected for testing its association with oxaliplatin response and survival given its eminent role in CRC biology and relevance in DNA damage response. *TP53* mutation was coded as 1 and wild type as 0.

2.2.5 *KRAS* mutation

Several signalling pathways that are relevant in CRC are controlled by Ras proteins, which are GTPases that work as a molecular switch. The most prominent downstream effector is Raf resulting in activation of the Ras/Raf/MAPK pathway, but several others have been identified such as PI3K/AKT, RALGDS or PLCε (Downward, 2003) . Several cell processes are regulated by these pathways such as inhibition of apoptosis, increased cell proliferation, promotion of cell growth, angiogenesis, and

differentiation (Jančík, 2010). The 3 Ras genes in human (*KRAS, NRAS* and *HRAS*) show different frequencies of mutation across tumour types by missense mutations, most of them in well-defined hotspots. In CRC, *KRAS* is frequently mutated at ~35-45%, while *NRAS* is much less common (~5%) and *HRAS* is not affected.

KRAS is located on chromosome 12 and encodes for a 21 kDa protein. *KRAS* driver mutations are linked to worse survival in late stages, usually within MSS CRCs (Domingo, 2018; Richman, 2009; Taieb, 2016). In addition, these mutations are predictive of treatment as such patients do not benefit from EGFR inhibition because Ras remains constitutively activated downstream (Douillard, 2013). Thus, due to its high frequency, biological relevance and predictive and prognostic effect, including cohorts treated with oxaliplatin, it was an interesting candidate to test for oxaliplatin stratification in metastatic disease. *KRAS* mutation was coded as 1 and wild type as 0.

2.2.6 *BRAF* V600E

Raf, as mentioned previously, is a downstream effector of RAS and plays a key role in the cascade of events to affect proliferation, differentiation and apoptosis. *BRAF* is the only Raf gene showing frequent mutations in some human cancers. In CRC occurs in ~10% of stage I-III (Domingo, 2013; Muzny, 2012) and ~5% of mCRCs (Roock, 2010). 95% of those mutations occur in exon 15 on codon 600 where there is conversion of T1799A resulting in amino acid glutamic acid (E) instead of valine (V) hence *BRAF*-V600E (Taieb, 2016). This is the most frequent amino acid change in CRC. Interestingly, this single mutation shows unequivocally a very strong biology as it is tightly associated with: MSI, low CIN, hypermethylation, right-sidedness and transcriptomic subtype CMS1 (Domingo, 2013; Guinney, 2015). In addition, it is also associated with APC wild type, it is clonal, and also frequently found in sessile serrated adenomas, which is the

most aggressive phenotype before malignancy. These last three points combined suggest it may be an initiating event in CRC. Similar to *KRAS* mutations, *BRAF*-V600E has a worse prognosis in late CRC within MSS cases (Domingo, 2018; Seligmann, 2015; Taieb, 2016), but with much worse hazard ratios. It has been suggested *BRAF*-V600E may be associated with resistance to standard chemotherapy in CRC as the worst prognosis seems to come from cohorts with such treatments. Accordingly, it could be hypothesised that this mutation could be a potential stratifier for oxaliplatin as one of the most common cytotoxic treatments in CRC. *BRAF V600E* mutation was coded as 1 and non-mutation as 0.

2.2.7 Hypoxia

Hypoxia is a suboptimal level of oxygen availability in tissue and it is known to have a key role in solid tumours. It has been linked with poor prognosis and drug resistance in cancer patients due to stem like features and increased metastasis (Man, 2018; Mi, 2020). In CRC, hypoxia has been correlated with poor outcomes, macrophage infiltration and activation of *RAS* signalling irrespective of *KRAS* mutation (Qi, 2020). Moreover, under hypoxic conditions, colon cells have EMT features and high expression of integrins to help them migrate and adhere to collagen and fibronectin (Hongo, 2013). In hypoxic CRC cells, oxaliplatin penetration, DNA damage levels and hypoxia-inducible factor-1 (HIF-1) processes were all contributors to oxaliplatin resistance (Roberts, 2009). Mechanistic studies have shown that HIF-1α repressed miR-338-5p conferring oxaliplatin resistance in CRC cell lines. Mir-338-5p has a direct target which is interleukin-6 (IL-6) which regulates STAT3 and Bcl2 activation, thus once HIF-1α is activated, miR-338-5p is repressed and thus IL-6 is not targeted by the miR causing drug resistance in CRC cell lines (Xu, 2019). For those reasons and due to its biological

relevance, a hypoxia signature (Buffa, 2010) showing reliable signals and commonly used in the field (Bhandari, 2020) was selected as a candidate biomarker to be tested against oxaliplatin treatment. Hypoxia signature is a continuous variable that has been scaled from 0 to 1 along with all the data.

2.2.8 Transforming growth factor beta (TGF-β)

TGF-β is a signalling pathway showing an interesting dual action in cancer. In normal cells and early stage tumours, TGF-β acts as a tumour suppressor by inducing apoptosis and inhibiting tumour growth. However, in advanced tumours TGF-β acts as a tumour promoter enhancing EMT and tumour cell migration (Seoane & Gomis, 2017). TGF-β signalling involves more than 30 components. TGF-β ligand is made up of two serine/threonine receptors, type one receptor is phosphorylated by type two receptor which then propagates the signal downstream to cytoplasmic effectors. The TGF-β pathway is divided into two paths: the activin path and the bone morphogenetic protein (BMP) path. Activin type one receptor phosphorylates SMAD 2/3 while BMP type one receptor phosphorylates SMAD 1/5/8. Those SMADs are termed R-SMAD. It takes two R-SMADs and one SMAD-4 to form the SMAD trimer which is crucial in TGF-β signalling in translocating to the nucleus and binding DNA. TGF-β ligands can also activate the pathway in SMAD independent manner affecting MAPK pathway, phosphoinositide 3-kinase (PI3K)/Akt and Rho/Rho-associated protein kinase (ROCK) pathways (Itatani, 2019).

TGF-β signalling has several clinical implications illustrated primarily by CMS. CMS4 which is underscored by activated TGF-β pathway and stromal infiltration shows poor prognosis and worst benefit from chemotherapy (Guinney, 2015; Okita, 2018). TGF-β in this setting activates EMT and results in an aggressive tumour state associated with

initiation of metastasis and poor prognosis (Calon, 2012). On the other hand, mutations and LOH in TGF-β tumour suppressor genes such as SMAD4, SMAD2 and SMAD3 occur in early CRC progression, usually in MSI- CRCs. These alterations result in non-functional proteins, disrupting the activation of the pathway. In MSI positive cases it is thought this is mediated by mutations in an exonic microsatellite in the *TGFBR2* gene. Due to its biological and clinical relevance, TGF-β was selected as a candidate to be tested against oxaliplatin treatment in mCRC. TGF-β is a continuous variable that has been scaled from 0 to 1 along with all the data.

2.2.9 CD8

This is a protein expressed on the surface of cytotoxic T-cells and some other immune cells. It is a transmembrane glycoprotein that dimerises and serves as a co-receptor for T cell receptor, mediating T cell signalling. High intra-tumour T-cell infiltrate is associated with less risk of recurrence and better outcome in CRC (Galon, 2006). Metastasis was also studied and compared with the tumour immune content where tumours without signs of early metastatic invasion had increased infiltrates of immune cells and increased mRNA for CD8, T-box transcription factor 21 (T-BET-21) and interferon-γ, all markers of type 1 helper T- cells. Those tumours also showed markers of T-cell migration, differentiation and activation (Pagès, 2009). An immune score methodology was created to quantify the intra-tumour immune infiltrate. It has been shown that patients with high immune score had significantly longer time to recurrence compared with patients with low immune score where they were mainly in high risk of recurrence (Galon, 2016). Increasing CD8 positive density was associated with reduced recurrence risk of CRC in more than 1800 samples of stage II and III CRC analysed. This association however was gradual across risk strata defined by tumour and nodal (T and

N) stage; not present in low risk cases, moderate in intermediate risk cases and strong in high risk cases. Thus, tumour and nodal strata affects the prognostic value of intra-tumoural CD8 positive infiltrates (Glaire, 2019). To assess the role of CD8 positive T cells on patients with liver metastasis following chemotherapy, T lymphocytes and HLA expression on tumour cells were studied. High CD8 positive intra-tumour infiltrates, high HLA-I expression and *RAS* wild type were associated with better overall survival where neoadjuvant oxaliplatin chemotherapy elicited CD8 positive cells recruitment and higher PD-L1 expression in *RAS* wild type patients increasing the chance for anti-EGFR therapy responsiveness (Ledys, 2018). Those observations had impact on therapy where it has been shown that FOLFOX treatment-controlled tumour burden in a CD8 T cells dependent manner. FOLFOX pushed tumour infiltrating lymphocytes into differentiation accessed by differentiation markers like PD-1 and TIM-3 and enhanced the effect of PD-1 checkpoint blockade (Guan, 2020). Moreover, it has been demonstrated that oxaliplatin treatment on mice showed that it reduces spleen size and cellularity while increasing the proportion of CD4 positive, CD8 positive and Treg cells. It was selectively enriching CD8 positive cells while depleting cytotoxically B cells. There was no effect on progenitor cells in the bone marrow compared to the mice control group (Stojanovska, 2019). From the above gathered literature, CD8 is an interesting marker to be tested as a predictor for oxaliplatin. CD8 is a continuous variable that has been scaled from 0 to 1 along with all the data.

2.2.10 CpG Island DNA Methylator Phenotype (CIMP)

DNA methylation is an epigenetic process involving the addition of a methyl group to a DNA strand, usually the cytosine of a CpG (cytosine-phosphate-guanine) site. Gene promoters harbour CpG islands which are regions enriched for such CpG

dinucleotides. DNA methyl transferases (DNMT) are enzymes that add a methyl group to the cytosine at position number 5 to become 5-methylcytosine. This enzymatic reaction regulates transcription and results in gene expression silencing. The primary target for DNMTs are CpG island rich regions (Hughes, 2012). In cancer, hypermethylation of CpG islands in promoter regions is linked with gene silencing while global hypomethylation is linked with gene activation (Esteller, 2007). In CRC, CpG island methylator phenotype (CIMP) is a subset of cancers showing concomitant hypermethylation in the promoter sites of several genes, leading to tumour suppressor genes silencing and oncogenes activation (Lao & Grady, 2011). It is an early event in the tumourigenic process although it is unclear if it is an initiating one or how it is triggered. Since BRAF-V600E mutations are strongly associated with CIMP, it has been hypothesised this may be the initiating alteration, although it has not been proven empirically. One gene commonly methylated in CIMP-H cases is the mismatch repair gene *MLH1*, giving rise to most cases of sporadic MSI and resulting in a strong association between both molecular phenotypes. In terms of prognosis, there are some conflicting reports so its involvement is unclear. This may be due to opposite prognostic links between *BRAF*-V600E and MSI, both strongly associated with CIMP. While originally CIMP was reported as a binary variable by running a handful of markers with PCR-related methods, subsequent analyses by methylation arrays showed CIMP was more granular than initially thought (Muzny, 2012).

Since CIMP is associated with concomitant silencing in several key CRC tumour suppressor genes with diverse functional roles, it was selected as a candidate for testing as a biomarker for oxaliplatin. CIMP is a continuous variable that has been scaled from 0 to 1 along with all the data.

2.2.11 Consensus Molecular Subtype

Based on transcriptomic data, a large international consortium developed a classifier showing four different CMS (Consensus Molecular Subtypes). CMS1 is the immune subtype and includes MSI, hypermutation, *BRAF*-V600E and has increased expression of genes associated with immune infiltrate. CMS2 is the canonical subtype which is usually CIN high, with pronounced *Myc* and Wnt activation and mainly of epithelial component. CMS3 is the metabolic subtype and it is also epithelial, showing frequently KRAS mutations and metabolic dysregulation. Finally, CMS4 is the mesenchymal subtype, showing high EMT, pronounced TGF-β activation, angiogenesis and stromal infiltration (Guinney, 2015). The biological features of CMS have been evaluated in relation to the clinical outcomes and it has been shown that in the metastatic setting, CMS2 has the best prognosis, while CMS1 has the worst prognosis upon standard cytotoxic chemotherapy but responds better to immune targeted therapy (Martini, 2020). There was also an attempt to check oxaliplatin benefit in CMS2 patients, yet no significant result was achieved in that direction (Pogue-Geile, 2019). Another hypothesis was generated to test whether the molecular subtypes of colon cancer were associated with prognosis and benefit from oxaliplatin therapy. Patients with stem-like characteristic tumours were associated with poor prognosis, suggesting lack of benefit upon oxaliplatin treatment, although no control, untreated group was available (Song, 2016). Thus, the CMS classification gives biological insights in CRC and the consensus was selected as a biomarker for oxaliplatin due to its importance in biology of CRC and possible links with treatment. CMS is a categorical variable where each subtype has been coded in this analysis as 1 and lack of this subtype as 0. For example, CMS1 has been coded as 1 while no CMS1 as 0.

2.2.12 ColoRectal Intrinsic Subtype (CRIS)

Tumour cells are surrounded by other cell types such as immune or stromal cells that have very different transcriptomic patterns. Accordingly, broad transcriptomic classifiers such as CMS are strongly biased depending on cell composition within the tumours. However, a tumour-specific classifier, independent of the cell type content, was recently built by using human xenographs on mouse models. consisting of five CRC intrinsic subtypes (CRIS). CRIS-A is mucinous, glycolytic and enriched for MSI or *KRAS* mutations; CRIS-B shows activated TGFβ pathway, EMT and poor prognosis; CRIS-C has amplified EGFR signalling and therefore responds to EGFR inhibitors; CRIS-D has activated WNT pathway and amplified expression of insulin like growth factor 2 (IGF2); CRIS-E is enriched with *TP53* mutations and Paneth cells (Isella, 2017). Comparing CMS and CRIS classifications, CRIS better classifies biopsies and resections and provides a more robust molecular classification in primary tumours due to its independence on cell type composition (Alderdice, 2018). Because CRIS looks to be a promising biomarker, it was selected to be tested as a stratifier for oxaliplatin. CRIS is a categorical variable where each subtype has been coded in this analysis as 1 and lack of this subtype as 0.

2.2.13 Chromosomal instability (CIN)

CIN refers to abnormal number of chromosomes due to gain or loss of part or whole chromosome during mitosis. CIN is highly prevalent (65% - 70% of sporadic cases) in colorectal cancer leading to genome instability and intra-tumoural heterogeneity. CIN is a major step in tumour development and progression. It can develop from abnormalities in chromosomal segregation, telomere instability or DNA damage response. The multistep progression from adenomas to carcinomas in colorectal cancer features a crucial step of CIN coupled with mutations in tumour suppressor genes and

oncogenes and different signalling pathways (Pino & Chung, 2010). A meta study that involved 63 studies stratifying survival in colorectal cancer by CIN, reported CIN is associated with less a favourable outcome and worse prognosis when compared with MSI-H tumours. CIN tumours showed poorer OS and PFS in colorectal cancer and thus was an interesting candidate to test its role, if any, with oxaliplatin outcome in particular (Walther, 2008).

2.3 Biomarker frequency and distribution per arm

The 13 biomarkers type and frequency are summarised in table 2.2. DDRD negative is higher than DDRD positive samples 5FUFA arm (80% and 20% respectively, table 2.2A) and FOLFOX arm (85% and 15% respectively, table 2.2B). There is a dominance of left side tumours in both datasets (63% in 5FUFA and 52% in FOLFOX). *APC* mutation is the most mutated gene in both datasets, followed by *TP53*, *KRAS* and *BRAF,* as reported in the literature. The latter is more mutated in the FOLFOX arm (18%) compared to the 5FUFA (8%). Hypoxia signature is a continuous variable showing similar mean and standard deviation between both arms. Similarly, for TGF-B fibroblasts, CD8 and CIMP. CMS and CRIS are both categorical variables. CMS2 is the most prevalent in both arms with 35% in 5FUFA and FOLFOX, followed by CMS4, CMS1, and CMS3. Quarter of the data is unclassified with any CMS and 2% is missing in both arms. CRIS-C is the most prevalent CRIS type followed by CRIS-B in both arms. Finally, CIN positive was very high in both datasets (80% in 5FUFA and 92% in FOLFOX).

Variable - 5FUFA arm	Type	Frequency	
DDRD	Binary	DDRD - : 80%	DDRD + : 20%
Sidedness	Binary	Left side: 63%	Right side: 37%
APC	Binary	WT : 12%	Mut : 88%
TP53	Binary	WT : 29%	Mut : 71%
KRAS	Binart	WT : 47	Mut: 53
BRAF-V600E	Binary	WT : 92%	Mut: 8%
Hypoxia Sig	Continuous	M=0.41, SD=0.15	
TGF-B	Continuous	M=0.44, SD=0.18	
CD8	Continuous	M=0.06, SD=0.09	
CIMP	Continuous	M=0.41, SD=0.35	
CMS	Categorical	CMS1 10%, CMS2 35%, CMS3 9%, CMS4 19%, Unclassified 25%, Missing 2%	
CRIS	Categorical	CRIS-A 18%, CRIS-B 17%, CRIS-C 23%, CRIS-D 11%, CRIS-E 14%, Unclassified 15%, Missing 2%	
CIN	Binary	CIN- 20%	CIN+ 80%

Variable - FOLFOX arm	Type	Frequency	
DDRD	Binary	DDRD - : 85%	DDRD + : 15%
Sidedness	Binary	Left side: 52%	Right side: 48%
APC	Binary	WT : 21%	Mut : 78%
TP53	Binary	WT : 28%	Mut : 72%
KRAS	Binary	WT: 52%	Mut: 48%
BRAF-V600E	Binary	WT : 83%	Mut: 17%
Hypoxia Sig	Continuous	M=0.39, SD=0.13	
TGF-B	Continuous	M=0.43, SD=0.17	
CD8	Continuous	M=0.06, SD=0.11	
CIMP	Continuous	M=0.42, SD=0.37	
CMS	Categorical	CMS1 11%, CMS2 35%, CMS3 6%, CMS4 24%, Unclassified 23%, Missing 1%	
CRIS	Categorical	CRIS-A 13%, CRIS-B 23%, CRIS-C 24%, CRIS-D 8%, CRIS-E 17%, Unclassified 14%, Missing 1%	
CIN	Binary	CIN- 8%	CIN+ 92%

Table 2.2 Biomarker frequency and distribution per arm. All the biomarkers tested show a similar frequency in both arms. (A) shows the variable type and frequency in the 5FUFA arm while (B) in the FOLFOX arm. DDRD negative is most prevalent in both arms. Left side tumours are dominant in both datasets. The most mutated gene is APC, followed by TP53, KRAS and BRAF-V600E. Hypoxia, TGF-β, CD8, and CIMP are all continuous variables with similar distribution in both arms. CMS2 is the most prevalent in the CMSs and CRIS-C in 5FUFA and FOLFOX datasets. CIN shows high positivity in both treatment arms. WT=wild type, Mut=mutated, M=mean, SD=standard deviation.

2.4 Missing data

Missing data in the FOCUS trial MSS (n=340) is reported in Figure 2.3. The Figure shows response and sidedness as clinical variables, RNA expression that is used for DDRD signature, hypoxia signature, TGF-β, CMS, and CRIS, mutation data for *APC*, *TP53*, KRAS and *BRAF* mutation status, IHC for CD8 status, methylation for CIMP call and finally CIN. In the FOCUS-MSS, there is 0.58% missing data in response, 9.94% in sidedness, 1.75% in RNA expression, 4.09% in IHC, 9.65% in methylation, and none in mutation and CIN (Figure 2.3A). Looking at the control arm of FOCUS (n=270), there is 0.37% missing data in response, 11.07% in sidedness, 1.85% in RNA expression, 2.95% in IHC, 0.23% in methylation and none in mutation and CIN (Figure 2.3B). Finally, in the FOLFOX arm (n=70), there is 1.41% missing data in response, 5.63% in sidedness, 1.41% in RNA expression, 8.45% in IHC, 11.27%
in methylation and no missing data in mutation and CIN (Figure 2.3C).

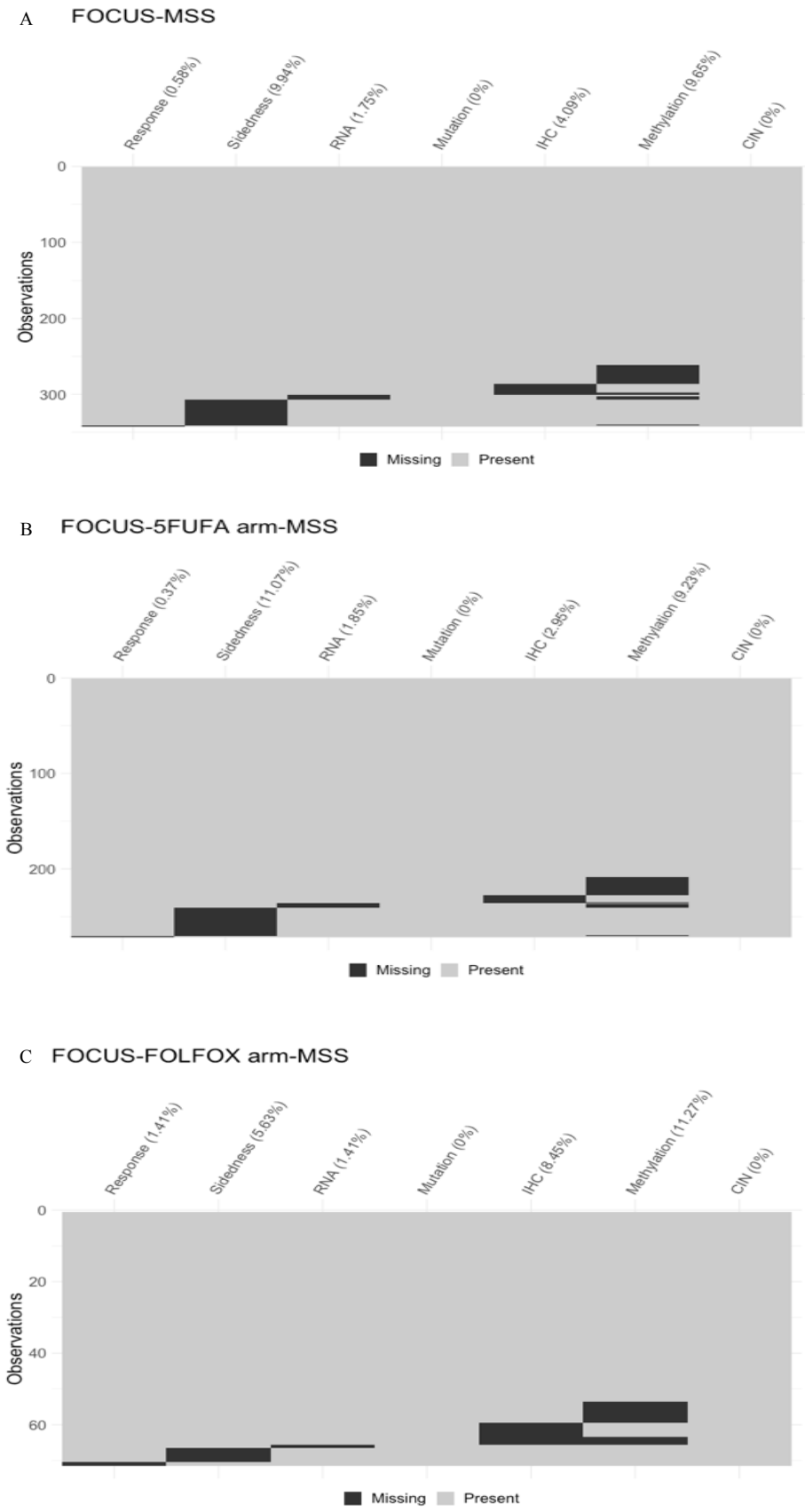


Figure 2.3 Missing data in the FOCUS trial. (A) shows the missing data in response, sidedness, RNA expression, mutation, IHC, methylation and CIN in the whole FOCUS-MSS cases (n=342). (B) shows missing data in the same categories in FOCUS 5FUFA arm (n=271), while (C) shows missing data in FOCUS FOLFOX arm (n=71).

2.5 Outcome of the 13 pre-specified candidates

2.5.1 Linear regression with binarised response as endpoint

The 13 pre-specified biomarkers were profiled as part of the SCORT bioinformatical pipeline and were downloaded from the SCORT data portal. Then, they are tested using logistic regression with binarised response to treatment as the endpoint. This analysis is conducted in MSS patients only. The complete responders (CR) along with the partial responders (PR) have been grouped together as good responders, while the stable disease patients (SD) along with progressive disease patients (PD) have been grouped together as bad responders. First, analysis in the 5FUFA alone control arm (n=270: CR+PR = 91, SD+PD = 179) is performed, showing *TP53* mutation (OR=0.33, 95% CI (0.17-0.62), P=0.0007), CMS2 transcriptomic subtype (OR=0.48, 95% CI (0.28-0.81), P=0.0064) , and CIN positive (OR=0.34, 95% CI (0.16-0.73), P=0.0055) to be associated with good response upon 5FUFA alone treatment while *KRAS* mutation (OR=2.03, 95% CI (1.21-3.38), P=0.0069) and CRIS-A transcriptomic subtype (OR=3.01, 95% CI (1.35-6.75), P=0.0064) shows poor response to 5FUFA (Figure 2.4A). Analysis in the FOLFOX arm (n=70: CR+PR = 47, SD+PD = 23) shows DDRD positive (OR=4.03, 95% CI (1.00-16.17), P=0.0492) to be associated with poor response to oxaliplatin (Figure 2.4B). Following this analysis, prediction to oxaliplatin treatment is assessed by interaction analyses between control arm and treatment arm, which shows DDRD positive patients have a better response when treated with 5FUFA rather than FOLFOX (OR=6.86, 95% CI, (1.49-31.47), P=0.0133) and patients with a KRAS mutation (OR=0.23, 95% CI, (0.07-0.73), P=0.0129) respond better upon FOLFOX treatment (Figure 2.4C).

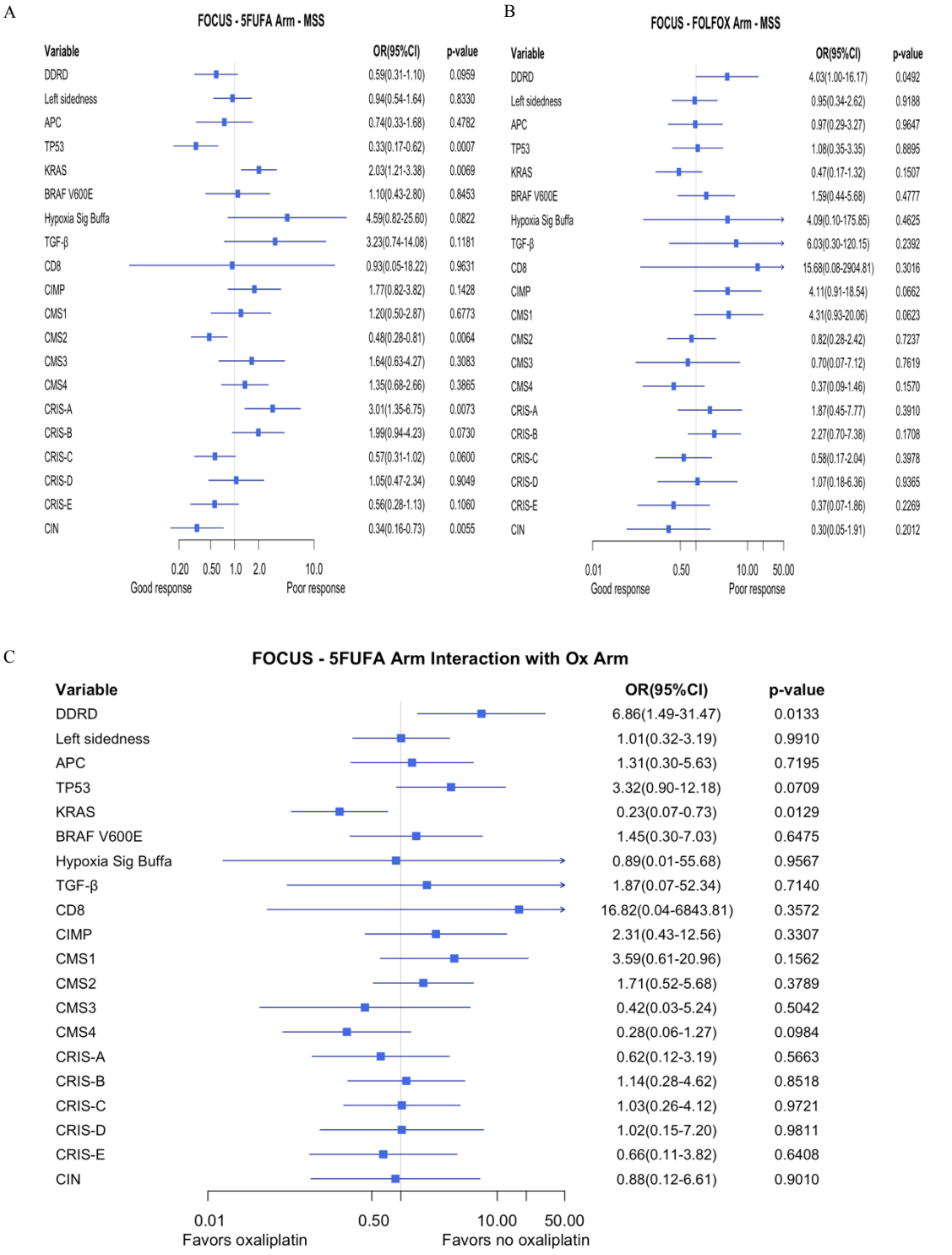


Figure 2.4 CRC biomarkers tested using logistic regression with binarised response to treatment as an endpoint. (A) Analysis in 5FUFA arm shows *TP53* mutation, CMS2 transcriptomic subtype, and CIN positive to be associated with good response upon 5FUFA treatment while *KRAS* mutation and CRIS-A transcriptomic subtype are associated with poor response to 5FUFA. (B) Analysis in FOLFOX arm shows DDRD positive to be associated with poor response to oxaliplatin. (C) Interaction analysis between the two arms shows DDRD positive predictive of oxaliplatin worse response while *KRAS* mutation predicts better outcome to oxaliplatin.

Next, I did multivariate interaction analysis on DDRD and *KRAS* to check if addition of the other significant biomarker to the regression model of the other would affect significance of the results observed in the previous section. This analysis shows that the predictive signal is not affected by adding any of the variables to the other interaction regression models (table 2.3). Independently DDRD positive (OR=6.79, 95%CI (1.47-31.36), P=0.0141) is predictive of worse response to oxaliplatin and *KRAS* mutation is similarly predictive of better response to oxaliplatin (OR=0.19, 95% CI (0.06-0.63), P=0.0063)

Variable	Univariate Analysis	Multivariate Analysis
DDRD	OR=6.86 (1.49-31.47), P=0.0133	OR=6.79 (1.47-31.36), P=0.0141
KRAS	OR=0.23 (0.07-0.73), P=0.0129	OR=0.19 (0.06-0.63), P=0.0063

Table 2.3 Multivariate analysis for significant biomarkers of response. Multivariate interaction regression analysis shows adding DDRD to KRAS regression model does not affect significance of KRAS mutation being a predictive biomarker for oxaliplatin better response. Similarly, adding KRAS to DDRD interaction regression model does not affect the significance of DDRD positive being a predictive marker for oxaliplatin worse response.

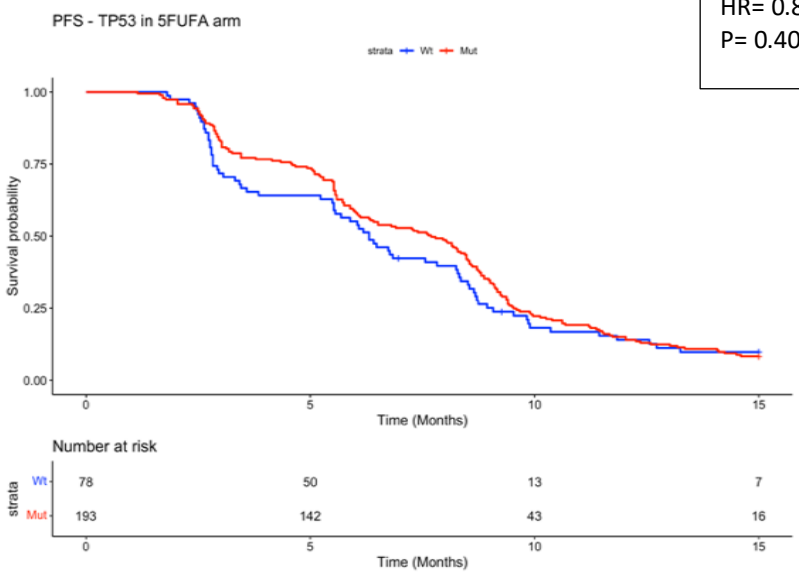
2.5.2 Cox regression with progression free survival as endpoint

All the biomarkers that were significant with response as an endpoint are tested using cox regression with progression free survival (PFS) censored at 15 months as a secondary endpoint. *TP53* mutation, *KRAS* mutation, CMS2, CRIS-A and CIN were all prognostic for response to 5FUFA treatment. However, only *KRAS* mutation is significantly prognostic of poor PFS for 5FUFA treatment (HR= 1.37, 95% CI, 95% CI (1.06 – 1.76) P=0.016) (Figure 2.5B). *TP53* mutation (HR= 0.89 , 95% CI (0.67 – 1.17), P=0.41) (Figure 2.5A), CMS2 (HR= 0.81, 95% CI (0.63 – 1.06), P= 0.13) (Figure 2.5C), CRIS-A (HR= 1.25, 95% CI (0.90 – 1.72), P= 0.18) (Figure 2.5D), and CIN (HR= 0.83,

95% CI (0.60 – 1.14), P= 0.24) (Figure 2.5E) show no significance in prognostic effect for MSS patients with metastatic colon cancer treated with first line 5FUFA alone on PFS. DDRD positive was prognostic for response to oxaliplatin, thus I tested it using cox regression with progression free survival as an endpoint. However, DDRD status (HR= 1.12, 95% CI (0.55 – 2.2), P= 0.78) is not significant in this analysis and is not prognostic of progression free survival upon oxaliplatin treatment (Figure 2.5F).

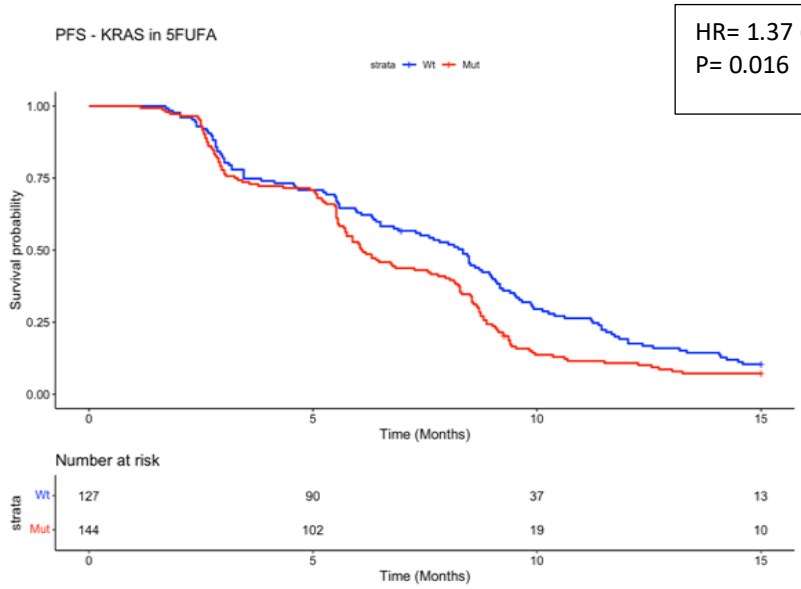
Both predictive biomarkers of response to oxaliplatin are tested to check if they also predict different progression free survival upon oxaliplatin treatment. Cox proportional hazard model of interaction showed no significant difference in progression free survival between 5FUFA and FOLFOX arms for DDRD status in MSS patients with metastatic colon cancer treated with oxaliplatin (HR= 1.0839, 95%CI (0.49 – 2.36), $P_{interaction}$= 0.84) (Figure 2.6A/B). A similar result is observed for KRAS mutation (HR=0.84, 95%CI (0.48 – 1.47), $P_{interaction}$= 0.54) (Figure 2.6C/D).

A



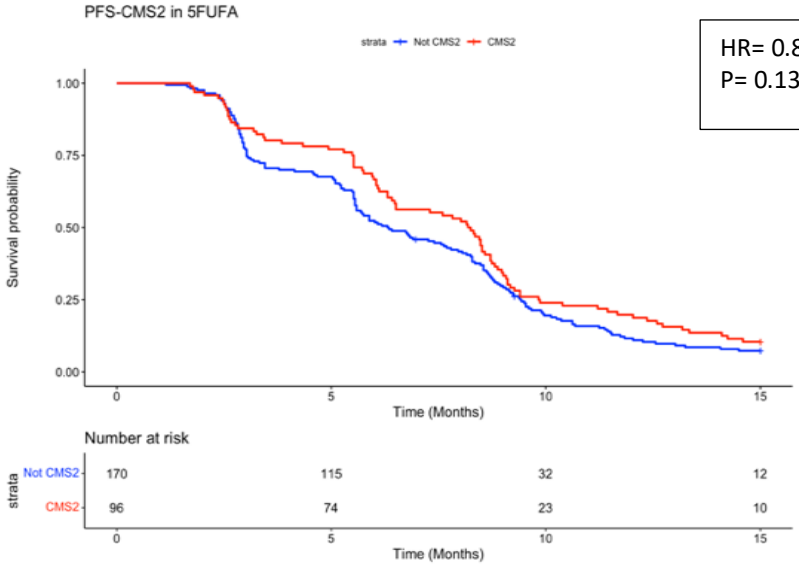
HR= 0.89 (0.67-1.17)
P= 0.405

B



HR= 1.37 (1.06- 1.76)
P= 0.016

C



HR= 0.81 (0.63- 1.06)
P= 0.13

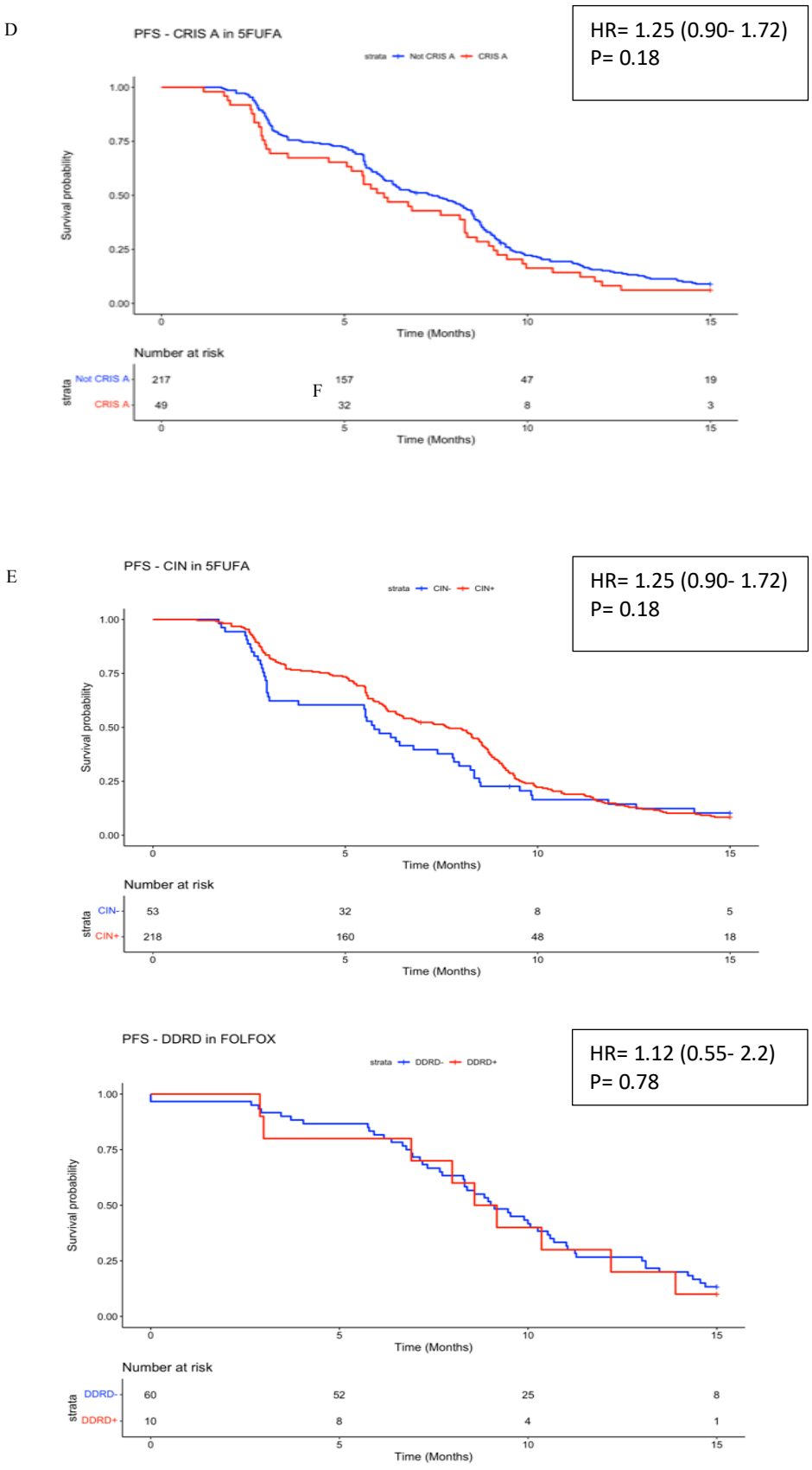
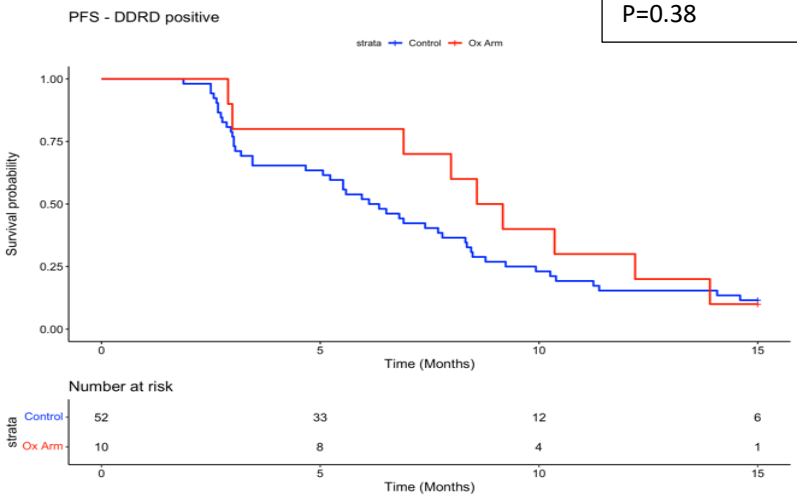


Figure 2.5 Survival analysis for significant prognostic biomarkers of response. (A)*TP53* mutation status, CMS2 (C), CRIS-A (D), and CIN (E) are not significantly associated with PFS for MSS patients with metastatic colon cancer treated with 5FUFA alone on PFS at 15 months. Only KRAS mutation (B) is prognostic of 5FUFA treatment in MSS patients. (F) DDRD status is not significantly associated with PFS to oxaliplatin.

A

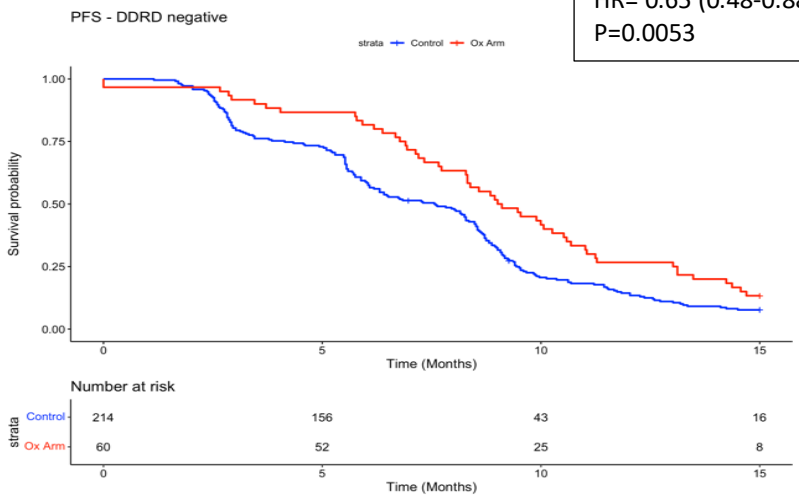
$P_{interaction} = 0.8391$

HR= 0.73 (0.35- 1.48)
P=0.38



B

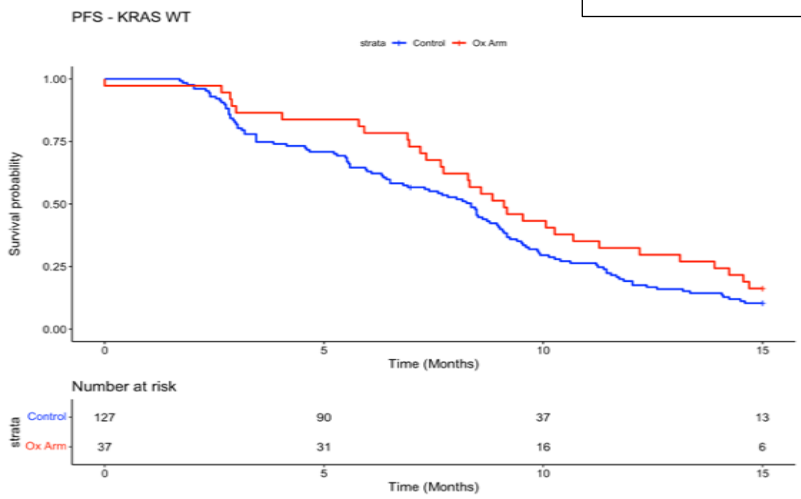
HR= 0.65 (0.48-0.88)
P=0.0053



P *interaction*= 0.5376

C

HR= 0.60 (0.40-0.89)
P=0.0107



D

HR= 0.73 (0.49 -1.09)
P=0.124

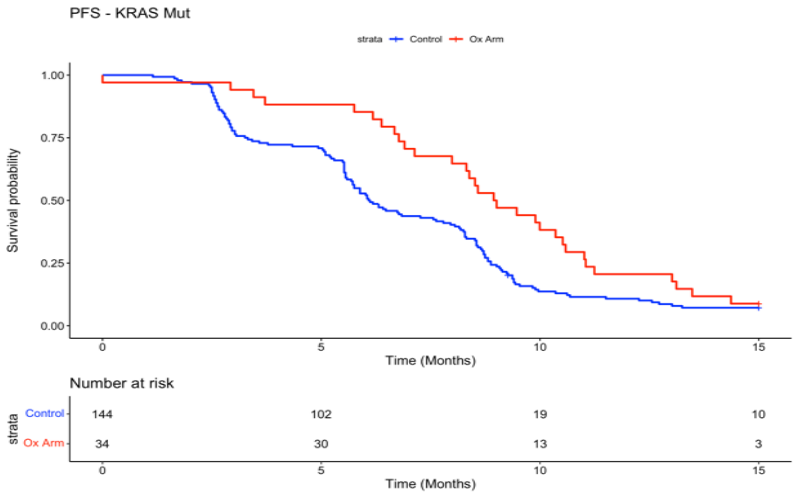


Figure 2.6 Survival analysis for significant predictive biomarkers of response. Neither DDRD status (A/B) nor KRAS mutation status (C/D) is significantly predictive of PFS for oxaliplatin treatment in MSS metastatic patients.

2.5 Summary

In this chapter, I investigated the potential role of known colorectal cancer biomarkers effect on oxaliplatin response. This analysis is conducted in MSS patients only due to the fact that MSI patients benefit from immunotherapy regimens. Response was selected as a primary end point for this investigation because response gives the most optimal view on oxaliplatin outcome as it gives a more direct measure of drug effect compared to survival endpoints where several confounders may result in background noise already in the variable that we want to test. PFS censored at 15 months is assessed as a secondary endpoint. We identify in this chapter that *TP53* mutation, CMS2 transcriptomic subtype, and CIN positive associated with good response for MSS patients with metastatic colon cancer treated with 5FUFA alone treatment while *KRAS* mutation and CRIS-A transcriptomic subtype show poor response to 5FUFA. Analysis in the FOLFOX arm shows DDRD positive to be associated with poor response to oxaliplatin in MSS patients with metastatic colon cancer. Only *KRAS* mutation has a bad prognostic effect on PFS for MSS patients treated with 5FUFA alone. We also identify that high DDRD score is predictive of worse response to oxaliplatin, while *KRAS* mutation is predictive of better response to oxaliplatin. Table 2.4 shows the exact numbers of *KRAS* and DDRD phenotype in every arm of the FOCUS trial.

	5FUFA		FOLFOX	
	Poor response	Good response	Poor response	Good response
KRAS Wt	38	53	25	22
KRAS Mut	106	73	8	15
DDRD negative	22	66	4	43
DDRD positive	148	29	6	16

Table 2.4 *KRAS* and DDRD phenotypes frequencies in each arm of FOCUS

Chapter 3 Identifying New Candidates for Oxaliplatin Resistance from a Systematic Literature Review and Testing their Role in Stratifying Patients to Oxaliplatin

Further to the 13 biomarkers studied in the previous chapter, we speculated that a systematic literature review of recent publications would identify new hypothesis-driven candidate biomarkers for oxaliplatin resistance or molecular mechanisms ascribed to the resistance to the drug. In this chapter, I conducted a systematic literature review to find new candidate biomarkers for oxaliplatin resistance and will test their role in stratifying patients to oxaliplatin. In order to investigate those candidate biomarkers, I will be using the FOCUS and Tsuji trials clinical and expression data for reasons explained further in the chapter.

3.1 Methods

3.1.1 Tsuji

The Tsuji (GSE28702) dataset is a publicly available cohort comprised of 83 patients with unresectable stage IV CRC undergoing FOLFOX treatment. Patients were recruited between 2007 and 2010 in Teikyo University Hospital at Mizonokuchi and Gifu University Hospital. Samples included 56 primaries and 27 metastatic CRCs (23 liver, 1 lung, and 3 peritoneum). All patients were treated with mFOLFOX6, following de Gramont regimen, 85mg m^{-2} oxaliplatin, 200mg m^{-2} LV, and 400mg m^{-2} 5-FU bolus on day 1, and 2400mg m^{-2} 5-FU as a 46-h continuous infusion starting on day 1, which is repeated every 2 weeks. Lesions were assessed after four cycles of therapy by CT scans

and responses was recorded following RECIST 1.1 version (The Response Evaluation Criteria for Solid Tumours) (Eisenhauer, 2009; Tsuji, 2012).

3.1.2 Gene expression profiling

For Tsuji cohort, RNA extraction was conducted with an ISOGENE™ kit (Nippon Gene, Tokyo, Japan), according to the manufacturer's protocol. RNA profiling was performed using 3 μg of total RNA on the HG-[U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array. All data are available at NCBI with GEO Accession GSE28702 (Tsuji, 2012).

I used GEOquery package (https://www.bioconductor.org/packages/release/bioc/html/GEOquery.html) , I got the expression matrix then I log transformed the data. Following that, I did quality control assessment of the expression value distribution (Figure 3.1A, showing Tsuji dataset) and the mean variance trend (Figure 3.1B, showing Tsuji dataset) that both showed good quality. Next, I loaded annotation file and linked probe sets to genes. Probe set IDs associated with more than a single gene were removed. Next, I joined gene names with expression and got mean per gene.

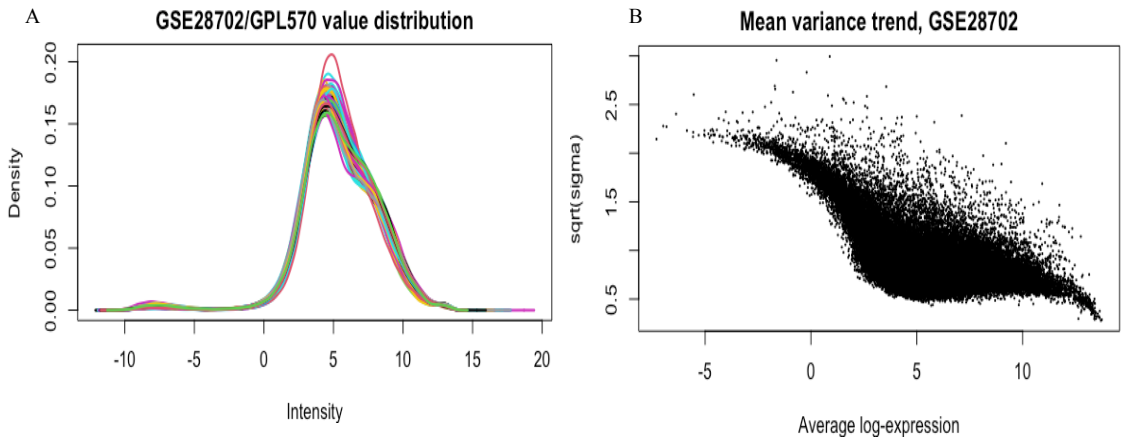


Figure 3.1 Quality control graphs of Tsuji dataset. Quality control assessment of the expression value distribution (A), and the mean variance trend (B), confirming good quality.

3.1.3 Combatted FOCUS Tsuji

Focus and Tsuji datasets were combined for analysis that will be explained next in this chapter. Batch effects adjusted with the combat method (https://rdrr.io/bioc/sva/man/ComBat.html) for further analysis of identified biomarkers of interest. Quality control Figures using Principal Components Analysis (PCA) (https://cran.r-project.org/web/packages/ggfortify/vignettes/plot_pca.html) are used to visualise good performance of the batch effect correction. The first four PCAs are plotted individually for comparison between uncorrected combat and corrected combat (Figure 3.2A, 3.2B and 3.2C). The three candidate biomarkers, which are identified from the analysis of the FOCUS trial and will be presented in the result section of this chapter, are checked between single cohort vs combatted in both cohorts separately and the correlation is one (Figure 3.2D).

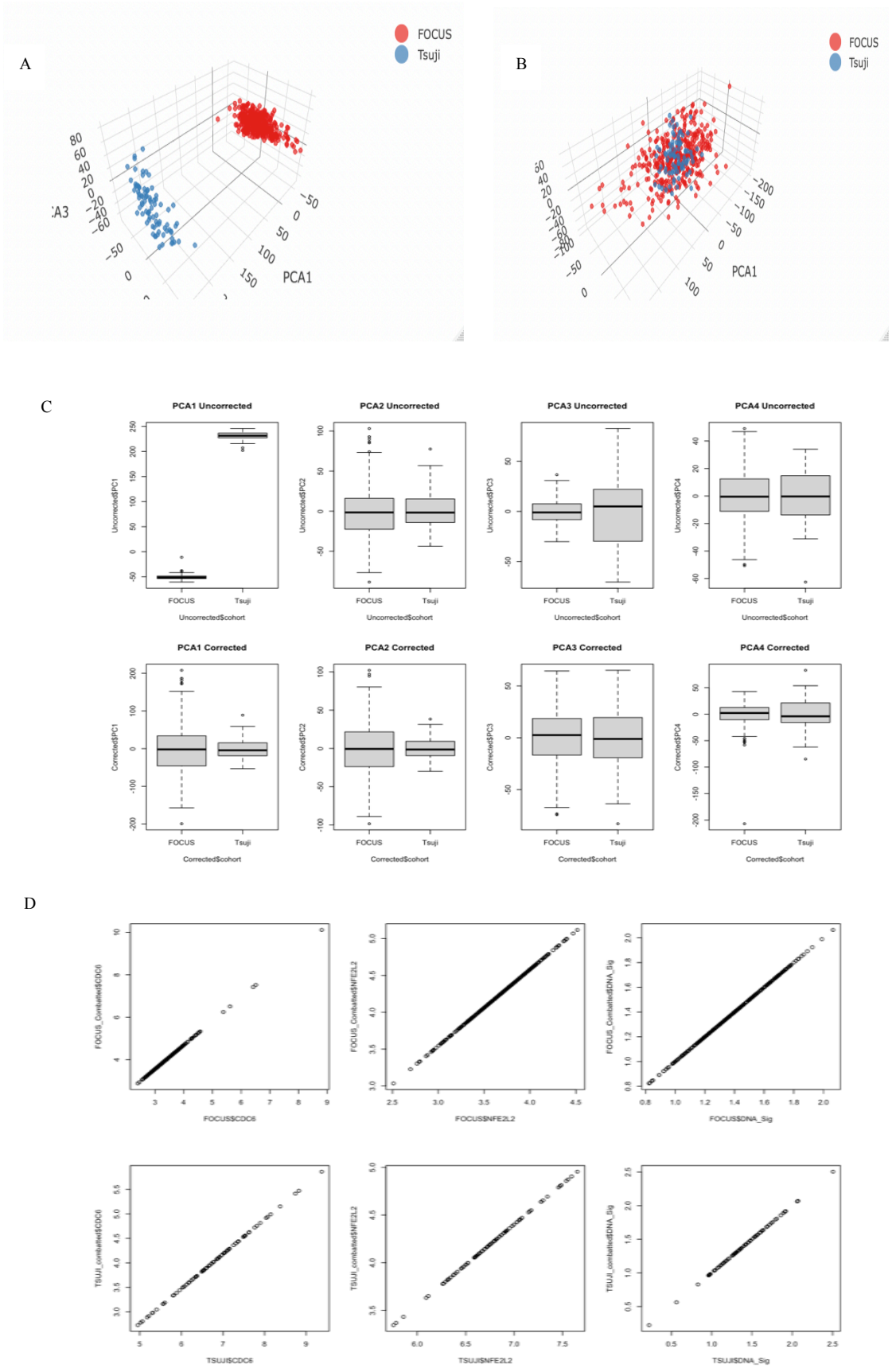


Figure 3.2 Principle component analysis of combined FOCUS and Tsuji datasets. Quality control graphs show the batch effect correction, (A) before correction and (B) after correction. (C) Boxplots by cohort in the first PCAs before and after batch correction. D) Correlation of main genes before and after batch correction in FOCUS (top) and Tsuji (bottom).

3.1.4 Regression model and ROC curves

For analysis in this chapter, I used R version 3.5.1 for bioinformatic analyses, available for download at https://www.r-project.org/. I conducted logistic regression analysis, similar to what I did in chapter 2, with binarised response to treatment (5FUFA arm versus FOLFOX treatment arm) as an endpoint using the function glm (https://cran.r-project.org/web/packages/cranly/vignettes/glms.html) from the stats R library. I scaled all tested variables from 0 to 2 to make them statistically comparable. I conducted cox regression in survival analysis (https://cran.r-project.org/web/packages/survival/survival.pdf) (https://cran.r-project.org/web/packages/survminer/index.html) with progression free survival as an endpoint and calculated statistical significance using cox proportional hazards progression (coxph). I also conducted multivariate analysis on significant candidates that came up from univariate interaction regression analysis to check if adding any of the significant variables to the model of the other one, would affect the significance of any of the significant variables.

3.1.5 DNA repair signature score

One of the selected candidates in the literature review is a signature score (Combes, 2019). It is composed of 12 genes. To generate such score, the gene expression datasets from the different cohorts are used. The bad prognostic genes (*HMGN1, RPA2, GTF2H2, SMUG1, ERCC1*) are assigned a positive direction and the good prognosis genes (*SRCAP, XRCC2, DMC1, GTF2H3, BARD1, MSH3, FAAP24*) are assigned an opposite negative direction (Combes, 2019). The mean of expression of the positive genes is generated, along with the mean of expression of the negative genes. Each sample DNA repair signature score is then generated by subtracting the two means.

3.2 Profiling, quality control and final list of candidates

On June 29th 2020, I performed a PubMed search (https://pubmed.ncbi.nlm.nih.gov) of the following keywords ((oxaliplatin resistance) OR (oxaliplatin resistant)) OR (oxaliplatin biomarker)) and identified 2521 papers. The first filter that has been applied to the search is to select papers from the last 5 years. This decreased the number of publications to 1245. Journals with impact factor five and below is eliminated from the results to end up with seventy-nine papers to be assessed. This final list of manuscripts has been assessed and examined thoroughly for variables reported for oxaliplatin resistance. This yielded sixty-six biomarkers. The latter are grouped into biological pathways and processes and given scores from zero to three based on biology, type of cancer, and feasibility of testing. For example, all biomarkers identified not in colorectal cancer are given a score of 0 and not included in further analysis. Our exclusion criteria are biomarkers with scores zero and one, those based on genes not profiled on XCEL array (including microRNA), and finally biomarkers not expressed in eighty percent of samples in the FOCUS trial. This triage has resulted in 32 biomarkers for further evaluation.

Nine of the identified biomarkers are under DNA damage repair pathways (DNA repair signature score, HSPH1, XIAP, NFYβ, E2F1, CHEK1, AKT1, GSK3β, ERCC1), five under microenvironment/EMT (MALAT1, E2H2, BMP4, STAT3, CXCL8), three within metabolism pathways (LPCAT2, G6PD, ID1), nine under signalling pathways (CRY2, FBXW7, MUC5AC, RAC1, SMAD4, PIM1, IDH1, CDC6, NFE2L2), three in drug transport (ABCC2, NR1I2, ABCC3), and some are not linked to any specific pathway (CDK5, EIF4A2, FOXC1) (Figure 3.3). A table showing those biomarkers and their characteristics is shown in table 3.1.

From the literature review, candidates excluded from further analysis are detailed in table 3.2.

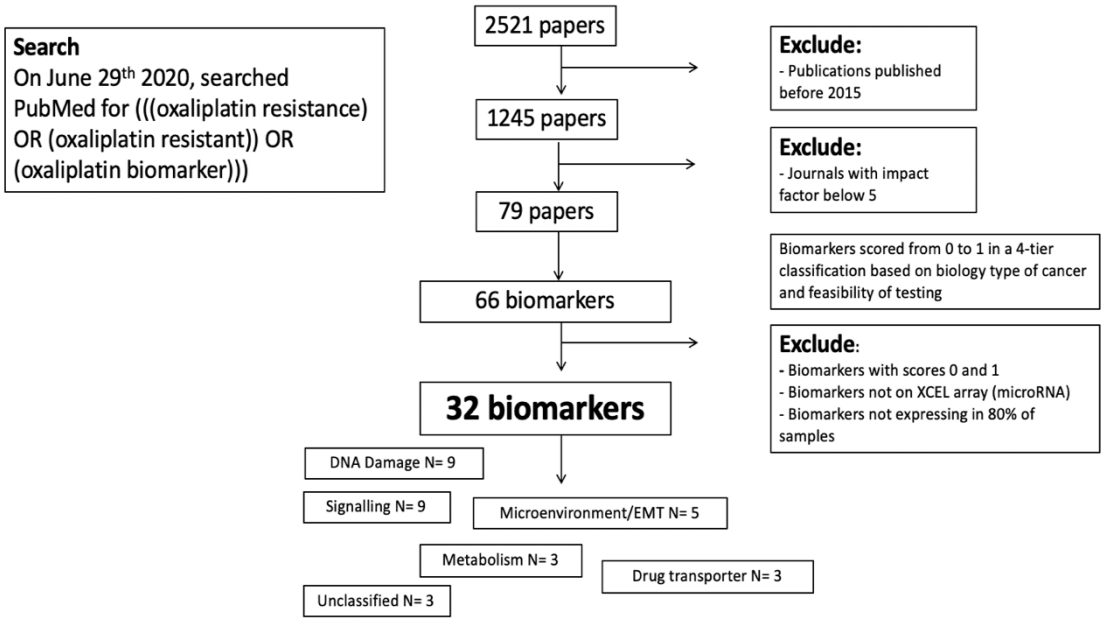


Figure 3.3 Diagram summarizing literature review steps and outcome. Pubmed search of certain keywords for oxaliplatin resistance gave 2521 papers. The first filter that has been applied is to select papers from the last 5 years only. This decreased the number of publications to 1245. Journals with impact factor five and below is eliminated from the results yielding 79 papers of which 66 biomarkers for oxaliplatin resistance were identified. Those are grouped into biological pathways and processes and given scores from zero to three based on biology, type of cancer, and feasibility of testing. Biomarkers with score 0 and 1, not expressed on XCEL array and not expressed in 80% of the FOCUS samples were all excluded to end up with 32 biomarkers qualified for further testing.

Candidate	Original Reference	Pathway	Score
DNA Repair Signature	Combes E, Cancer Res, 2019	DNA Damage	3
HSPH1	Causee S, Oncogene 2019	DNA Damage	3
	Dorard C, Nat Med 2011		
XIAP	Hua Y, Cell Physio &Biochem 2019	DNA Damage	3
NFYB/E2F1/CHEK1	Fang Z, Cancer Letters, 2018	DNA Damage	3
AKT/GSK3β/ERCC1	Wei W, Onc Res, 2020	DNA Damage	2
	Li P, Br J Cancer 2013		
	Uchida K, Cancer 2008		
MALAT1/EZH2	Li P, Mol Cancer and Ther, 2017	Microenvironment/ EMT	3
BMP4	Ma J, Cancer Letters, 2017	Microenvironment/ EMT	3
STAT3	Yin Y, Clin Cancer Res, 2017	Microenvironment/ EMT	3
CXCL8	Dabkeviciene D, Medical Oncology 2015	Microenvironment/ EMT	3
LPCAT2	Cotte A, Nature Comm 2018	Metabolism	3
G6PD	Ju H, Oncogene 2017	Metabolism	3
ID1	Yin X, Exp Clin Cancer, 2017	Metabolism	2
ABCC2	Biswas R, Cancers(Basel), 2019	Drug Supporter	3
NR1I2/ ABCC3	Dong Y, Molecular Cancer 22017	Drug Supporter	3
CRY2/FBXW7	Fang L, Mol Cancer Ther 2015	Signaling	2
MUC5AC	Pothuraju, Mol Cancer 2020	Signaling	2
RAC1	Goka E, Mol Cancer Therapeutics, 2019	Signaling	2
SMAD4	Jiang T, Exp Clin Cancer Res, 2017	Signaling	3
PIM1/IDH1/CDC6	Cai J,Mol Cancer Therapeutics, 2019	Signaling	3
	Blanco F, Oncogene, 2016		
NFE2L2	Cheng Y, Oncotarget 2018	Signaling	3
CDK5	de Porras, Cancers, 2019	Not classified	3
EIF4A2	Chen ZH, Exper and clin cancer research, 2019	Not classified	3
FOXC1	Hsu H, Cancer(Basel), 2019	Not classified	3

Table 3.1 Final list of candidate biomarkers from literature review. The table shows the final list of biomarkers from the literature review that are selected for further testing in this chapter, along with the original reference where the biomarker was selected from, the pathway it belongs to and the score that was given based on biology, type of cancer and feasibility of testing.

Candidate	Oroginal Reference	Reason for Exclusion
ATR	Combes E, Cancer Research, 2019	Not expressed in 80% of samples
ABCB1	Wang H, Cancer Science, 2020	Not expressed in 80% of samples
miR-19b-3p	Jiang R, Experimental & Clinical Cancer Res,2020	Not profiled in XCEL array
IL-6 (IL6)	Xu K, Molecular Therapy, 2020	Not expressed in 80% of samples
CXCR2	Dabkeviciene D,Medical Oncology, 2015	Not expressed in 80% of samples
miR-625-3p	Dabkeviciene D,Medical Oncology, 2016	Not profiled in XCEL array
PIM1	Xia J, Mol Cancer Therpeutics,2019	Not expressed in 80% of samples
LATS2	Hsu H, Cancers, 2019	Not expressed in 80% of samples
Slug (SNAI2)	Wei W , Oncology Research, 2020	Not expressed in 80% of samples
OCT2 (SLC22A2)	Chen L, Theranostics, 2020	Not expressed in 80% of samples
PTGER2	Iwamoto K,British J of Cancer,2020	Not expressed in 80% of samples
hnRNP L	Hu W,Cell Disease and Death ,2019	Score 1
FOXC2	Cheng Y, Onco Target, 2020	Score 0
ZEB2	Francescangeli F,Experimental Cancer Res, 2020	Score 0 Possible contamination by stromal content
TLR7/8	Liu Z, Cancer Letters, 2020	Score 0
CCAL	Deng X, International Journal of Cancer, 2019	Score 0
IL1β/TGFβ1	Diaz-Maroto N, Clinical Cancer Research, 2019	Score 0
CYTOR	Yue B, Molecular Therapy, 2018	Score 0
TERT	Galaine J. International J of Cancer, 2019	Score 0 Difficulty in Testing
COA-1	Galaine J. International J of Cancer, 2019	Score 0 Difficulty in Testing
circCCDC66	Lin Y, Cancers (Basel), 2020	Score 0 Difficulty in detection on XCEL array
PSMD4	Cheng Y. M, Oncotarget 2018	Score 0
PKM2	Wang X, Molecular Oncology, 2020	Score 0
HIF 1a	Wei T, Oncogene, 2019 ; Xu k, Mol Ther,2019	Covered in previous studies in the group
PRX2	Cerda M, Cancers Letters, 2017	Score 0
Nuclear IGF-1R	Servant J, British Journal Of Cancer, 2017	Score 0 Difficulty in detection on XCEL array
NOS3	Jeong S, Cancers (Basel), 2019	Score 0
CCN2-MAPK-Id-1	Liao X , Hepatlogy International, 2019	Not in CRC
HSF1	Shen JH, American J of Cancer Res, 2019	Not in CRC
JNK1-JUN	Lipner M, JCI, 2020	Not in CRC
IL-8	Yamada S, J Gastrointest Onco 2019	Not in CRC
PML	Swayden M, FASEB, 2019	Not in CRC
GKN2	Zhang Z, J of Exp & Clinical Cancer Research, 2019	Not in CRC
PRKRA	Hisamatsu T, Molecular Cancer Therapeutics, 2019	Not in CRC
EVA1A	Ren W, Cell Death and Disease, 2019	Not in CRC
CLDN18-ARHGAP26/6	Shu Y, Cell Communication, 2018	Not in CRC
PAK	Jiang Y, EBioMedicine, 2017	Not in CRC
lncRNA HULC	Xiong H, Oncogene, 2017	Not in CRC
ATP7A	Hua-Li Z, Journal of Cancer, 2016	Not in CRC
Hsp72	Takahashi K, Cancer Letters, 2016	Not in CRC
CHKα	Mazarico J, Molecular Cancer Therapeutics, 2016	Not in CRC

Table 3.2 List of excluded biomarkers from literature review. The table shows the list of excluded biomarkers from the literature review and are not carried on for further testing in this chapter, along with the original reference where the biomarker was selected from, and the reason for exclusion.

3.2.1 Biomarkers in the DNA repair pathway

DNA repair signature score: In the DNA damage repair pathway, it has been shown that oxaliplatin sensitivity was restored in oxaliplatin resistant cell lines upon silencing ataxia-telangiectasia mutated and RAD3-related (ATR) which is a serine/threonine kinase important in response to DNA stress. Upon administering ATR inhibitor (VE-822) and oxaliplatin to 6 different CRC cell lines, a synergistic effect was seen in single strand and double strand break formation, growth arrest and apoptosis. Similar results were seen in CRC mouse model. Finally, a DNA repair signature was able to differentiate patients with oxaliplatin sensitivity versus patients with oxaliplatin resistance (Combes, 2019).

HSPH1: A multicentre clinical study showed that 15% of patients showed a loss of function mutation in *HSP110* (*HSPH1*) and patients expressing low levels of the wild type HSP110 showed better outcome to oxaliplatin treatment. Experiments done in vitro, in vivo and from biopsies of patients showed that *HSP110* co-localizes with γ-H_2AX (DNA damage) and translocate to the nucleus upon oxaliplatin treatment and promotes DNA repair (Causse, 2019).

***XIAP*:** oxaliplatin resistant CRC cell lines has shown higher levels of X-linked inhibitor of apoptosis protein (*XIAP*) which is known to be an important signalling molecule for the regulation of DNA repair related proteins like γ-H_2AX, and lower levels of miR-122 compared to the parental cell line (Hua, 2018).

***NFYB/E2F1/CHEK1*:** oxaliplatin resistant cell lines have shown higher levels of *NFYB* and *E2F1* with *CHK1* being a target for *E2F1*. Knocking down *CHK1* sensitised the cells to OX. Clinically patients treated with oxaliplatin had high levels of *NFYB, E2F1 and CHK1* and resulted in poor survival (Fang, 2018).

***AKT/GSK3β /ERCC1*:** Excision Repair Cross-Complementing group-1 (*ERCC1*) upregulation was linked to oxaliplatin resistance in HCT116 cells leading to epithelial to mesenchymal (EMT) induction, activation of *AKT/GSK3β*, and overexpression of slug. This phenotype was reversed by slug silencing consequently mediating oxaliplatin outcome (Wei, 2020). *ERCC1* was induced in patients undergoing oxaliplatin treatment and resulted with poor outcome (Rao, 2019).

3.2.2 Biomarkers in the microenvironment/EMT

***MALAT1/EZH2*:** *MALAT1* is a long noncoding RNA (lnRNA) and its 3' end region is associated with *EZH2*. *MALAT1* has been found to be highly expressed in patients responding poorly to oxaliplatin based chemotherapy and also high, along with EMT phenotype, in oxaliplatin resistant CRC cell lines. When this lncRNA is knocked down in cells, oxaliplatin induced EMT and chemo-resistance are inhibited (Li, 2017).

***BMP4*:** Extensive research has been performed on *BMP4* to detect its role on EMT and oxaliplatin sensitivity in human hepatocellular carcinoma specimen, CRC cell lines, and subcutaneous tumour model receiving oxaliplatin treatment. *BMP4* was correlated with oxaliplatin resistance and EMT activation. Knockdown of *BMP4* rescued phenotype and oxaliplatin sensitivity. Mechanistically, *MEK/ERK/ELK1*mediates *BMP4* to induce *EMT* and cause oxaliplatin resistance. Blocking this axis or BMP4 restores sensitivity to oxaliplatin(Ma, 2017).

STAT3: HIF-1α induces hypoxia induced oxaliplatin resistance. At the same time, chemo-resistance relating to the micro-environment is seen through macrophage derived IL-6 which activates IL6R/STAT3 pathway and HIF1α, in this context, inhibits tumour suppressor miR-338-5p. Thus there is HIF-1α/miR-338-5p/IL-6 feedback axis contributing to the resistance seen (Xu, 2019).

CXCL8: Immunologically, CXXC cytokine family receptor 2 (CXCR2) and interleukin 8 (CXCL8) upregulation in CRC cell line HCT116 caused chemo-resistance (Dabkeviciene, 2015).

3.2.3 Biomarkers in metabolism

***LPCAT2*:** Lipid droplet content in CRC cell lines is linked to the expression of LPCAT2, an enzyme involved in phosphatidylcholine synthesis. High lipid content has also been shown to cause resistance to 5-FU and oxaliplatin (Cotte, 2018).

***G6PD*/ID1:** Similarly, ID1 high expression was detected in oxaliplatin resistant cell lines. Knocking it down decreased the expression of G6PD, an enzyme catalysing pentose phosphate pathway a major glucose catabolism pathway. This resulted in high reactive oxygen species content in the cells driving apoptosis (Ju, 2017; X. Yin, 2017).

3.2.4 Biomarkers in drug transport

***ABCC2*: oxaliplatin** resistance has also been linked to drug transporters function and status. Increased expression of ABCC2 gene that encodes MRP2 protein has been linked to decreased oxaliplatin accumulation and toxicity in colorectal and pancreatic cancer cell lines. Silencing MRP2 restored oxaliplatin levels in the cells and consequently drove cytotoxicity (Biswas, 2019).

***NR1I2/ABCC3*:** Another regulator of chemotherapy response is PXR gene (NR1I2) which regulates downstream genes involved in drug transport. PXR overexpression increase transport of oxaliplatin while reducing its content by binding to MRP3 (ABCC3) while knocking it down restores oxaliplatin cytotoxicity (Dong, 2017).

3.2.5 Biomarkers in signalling pathways

CRY2/FBXW7: Other pathways that seem to affect oxaliplatin response are different signalling pathways. Circadian clock protein called cryptochrome 2 (CRY2) was shown to be overexpressed in patients treated with oxaliplatin and having a poor outcome, its knockdown in cells have restored oxaliplatin sensitivity. A ubiquitin ligase FBXW7 was identified to target CRY2 for degradation. High expression of FBXW7 downregulates CRY2 and increases sensitivity to chemotherapy in CRC (Fan, 2015).

***MUC5AC*:** *MUC5AC* is another gene when over-expressed in patients and cell lines results in higher malignancy and resistance to 5-FU and OX. *MUC5AC* interacts with CD44 activating Src signaling resulting in oxaliplatin resistance (Pothuraju, 2020).

RAC1: Another signaling mechanism that causes decrease in apoptosis and thus oxaliplatin resistance is NF-κB. Rac1b, an isoform of Rac GTPase, is upregulated and consequently causes NF-κB signaling increasing proliferation and inhibiting apoptosis (Goka, 2019).

SMAD4: Several microRNAs are linked to malignant biology in CRC, one of the most prominent is miR-19b-3p where its high expression has been associated with worse phenotype from higher N stage to higher liver metastasis. miR-19b-3p has also been shown as a prognostic marker associated with overall and disease-free survival. In vitro, it has been linked to chemo-resistance particularly oxaliplatin through targeting SMAD4 (Jiang, 2017).

PIM1/IDH1/CDC6 : Human antigen R (HuR) regulates the DNA replication gene CDC6 by binding to its 3' region and both are overexpressed in oxaliplatin resistant in human CRC tissue (Cai, 2019). Isocitrate dehydrogenase 1 (IDH1) has been identified as an antioxidant enzyme under HuR regulation. High expression of IDH1 promoted chemo-resistant in nutrient deficient PDAC cells, showing a role of IDH1 and HuR in causing chemo-resistance (Zarei, 2017). Mechanistic studies showed that HuR promotes

hypoxia induced chemo-resistance through post transcriptional regulation of proto-oncogene PIM1 (Blanco, 2016).

***NFE2L2*:** Similarly, NRF2 has been linked to cause chemo-resistance in CRC and recently NRF2 location in the cytoplasm versus the nucleus has been specifically studied and linked to this resistance. Cytoplasmic NRF2 upregulated PSMD4 expression that consequently had an effect on NF-κB/AKT/β-catenin/ZEB1axis causing resistance to 5-FU and oxaliplatin(Cheng, 2018) .

3.2.6 Biomarkers in unclassified pathways

CDK5: CDK5 is a kinase also highly expressed in CRC cell lines and tissues. In patients high expression correlated with poor prognosis, while in the metastatic setting, high expression was only seen in those receiving oxaliplatin(de Porras, 2019).

EIF4A2: EIF4A2 is a protein which showed higher expression in patient samples tumour site upon immunohistochemistry staining. Its knockdown increased oxaliplatin sensitivity in vitro and in vivo (Chen, 2019).

FOXC1: Transcription factor FOXC1 increases the transcription of microRNA miR-31-5p by binding its promoter region which in it turn targets LATS2 causing resistance to oxaliplatin (Hsu, 2019)

3.3 Outcome analysis of literature review candidate biomarkers

32 biomarkers, which are all continuous variables, identified from the literature review are tested using logistic regression with binarised response to treatment as the endpoint. The complete responders (CR) along with the partial responders (PR) are

grouped together as good responders, while the stable disease patients (SD) along with progressive disease patients (PD) are grouped together as bad responders. First, I performed an analysis in the 5FUFA control arm (n=265: CR+PR n=88, SD+PD n=177), selected for MSS patients. This analysis shows *BMP4* (OR=2.92, 95%CI (1.33- 6.39), P= 0.0074), *CXCL8* (OR=2.09, 95%CI (1.05- 4.17), P= 0.00364) , and *AKT1* (OR=2.44, 95%CI (1.02- 5.83), P=0.0453) high expression is associated with poor response to 5FUFA arm in colon cancer metastatic patients (Figure 3.4A). The same analysis is conducted in the FOLFOX arm (n=70: CR+PR =47, SD+PD n=22) and has yielded *NR1I2* (OR=0.21, 95%CI (0.05- 0.89), P=0.0339), and DNA repair signature score (OR=5.52, 95%CI (1.02- 29.78, P= 0.0468), as statistically significant markers for oxaliplatin response (Figure 3.4B). Stronger expression of NR1I2 is associated with good response, while high score of the DNA repair signature is associated with poor response to oxaliplatin. Finally, interaction analysis between the two arms (5FUFA versus FOLFOX) showed *CDC6* (OR=25.46, 95%CI (1.03- 631.41), P=0.0482), *NFE2L2* (OR=0.10, 95%CI (0.02- 0.61), P= 0.0120) , and DNA repair signature score (OR=6.60, 95%CI (1.07- 40.82), P= 0.0423) have statistically different levels of response (Figure 3.4C). High levels of CDC6 expression and high DNA repair signature score have resulted in worse response when treated with FOLFOX than 5FUFA, while high expression of NFE2L2 has resulted in better response when treated with FOLFOX than 5FUFA.

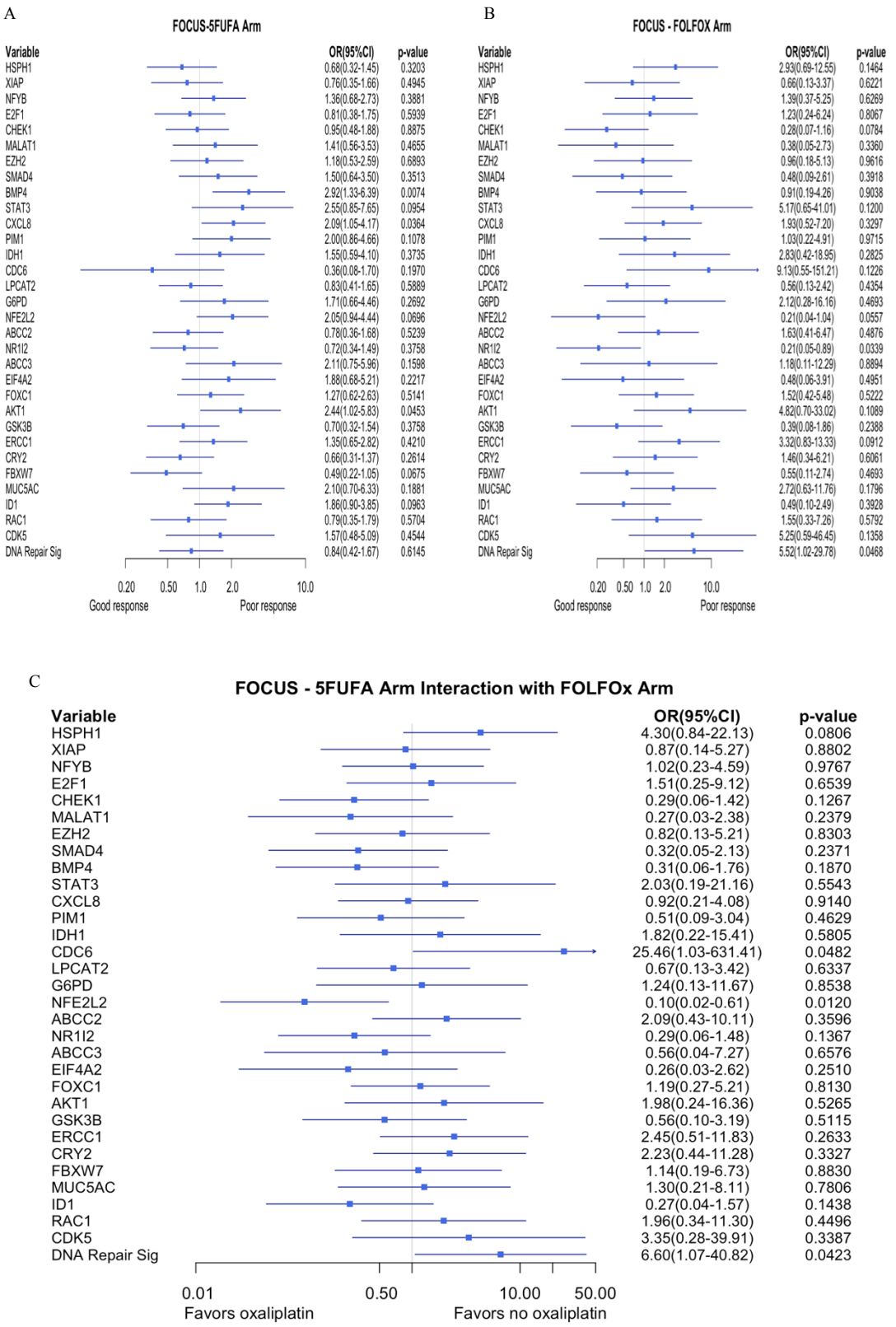


Figure 3.4 Literature review biomarkers tested using logistic regression with binarised response to treatment as an endpoint. (A) Analysis in 5FUFA arm shows BMP4, CXCL8, and AKT1 associated with poor response to 5FUFA in MSS metastatic colon cancer patients. (B) Similar analysis in FOLFOX arm shows NR1I2 high expression associated with good response to oxaliplatin and DNA damage repair signature high score associated with bad response to oxaliplatin. (C) Interaction analysis between the two arms shows CDC6 high expression and DNA damage repair signature high score predictive of oxaliplatin worse response, while NFE2L2 high expression is predictive of oxaliplatin better response.

Next, I did multivariate interaction analysis on *CDC6*, *NFE2L2*, and DNA repair signature to check if adding any of the significant biomarker to the regression model of the other would affect significance of the results observed in the previous section. This analysis shows that the predictive signal is not affected by adding any of the variables to the other interaction regression models (table 3.3). Independently CDC6 high expression positive (OR=26.45, 95%CI (1.05- 667.09), P=0.047) and DNA repair signature high score (OR=6.63, 95%CI (1.05- 41.72), P=0.043) is predictive of worse response to oxaliplatin and *NFE2L2* high expression is similarly predictive of better response to oxaliplatin (OR=0.09, 95%CI (0.02- 0.57), P=0.010)

Variable	Univariate Analysis	Multivariate Analysis
CDC6	OR=25.46(1.03-631.41), P=0.048	OR=26.45(1.05-667.09), P=0.047
NFE2L2	OR=0.10 (0.02-0.61), P=0.012	OR=0.09(0.02-0.57), P=0.010
DNA Repair Sig	OR=6.60 (1.07-40.82), P=0.042	OR=6.63 (1.05-41.72), P=0.043

Table 3.3 Multivariate analysis for significant biomarkers of literature review. Multivariate interaction regression analysis shows adding any of the significant biomarkers to regression model of the others does not affect significance of any of the biomarkers. The 3 biomarkers are significant in both univariate and multivariate analyses.

3.4 Outcome analysis following binarisation of the three identified biomarkers

3.4.1 Linear regression with binarised response as endpoint

For a biomarker to be clinically relevant, it is best to have a clear threshold above which is considered high and below which is considered low. Thus, I wanted to binarise the 3 significant biomarkers for response to oxaliplatin. To do so, we selected a publicly available dataset called Tsuji that have similar clinical settings to FOCUS. Both are

metastatic disease, Tsuji has oxaliplatin based chemotherapy as a treatment of choice, and both follow the RECIST criteria for recording responses. We did not want to derive a cut point from FOCUS being it the same cohort used in the testing and the derivation of those three significant biomarkers. We wanted to have an independent cohort that we can use to derive a threshold around which we binarise in FOCUS. Thus, the first thing that I did is combatted Tsuji and FOCUS (described in methods section 3.1.4) for both datasets to be comparable transcriptionally. All quality control testing that I performed are presented in section 3.1.4.

Further quality control testing was conducted where regression analysis is run for the three candidate biomarkers in the original datasets of both FOCUS and Tsuji compared to the regression analysis in the combatted dataset. In the FOCUS dataset, *CDC6* and *NFE2L2* show the exact same results with both significant p-values, while the DNA repair signature score shows the same trend but lost significance in the combatted dataset (Figure 3.5A/B). In the Tsuji dataset, *CDC6* and *NFE2L2* show the same trend in the original dataset and the combatted but with no significance in both, while the DNA repair score have the same trend and significant in both original Tsuji dataset and in the combatted one (Figure 3.5C/D). Those quality control analyses gave confidence in the combatted dataset and deemed it suitable for further analysis.

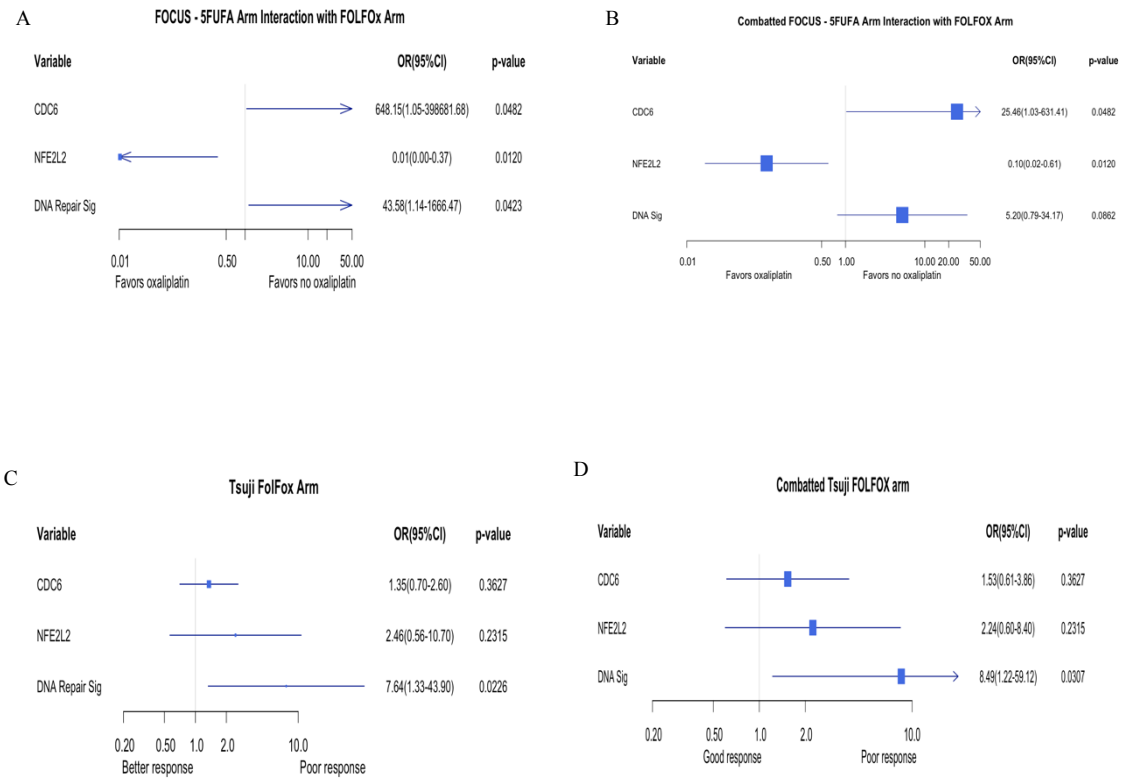


Figure 3.5 Regression analysis of the three significant biomarkers in the original datasets versus the combatted dataset. (A) In the FOCUS dataset, *CDC6* and *NFE2L2* show the exact same results with significant p-values, while the DNA repair signature score shows the same trend but lost significance in the combatted dataset (B). In the Tsuji dataset, *CDC6* and *NFE2L2* show the same trend in the original dataset (C) and the combatted (D) but with no significance in both, while the DNA repair score have the same trend and significant in both original Tsuji dataset and in the combatted one.

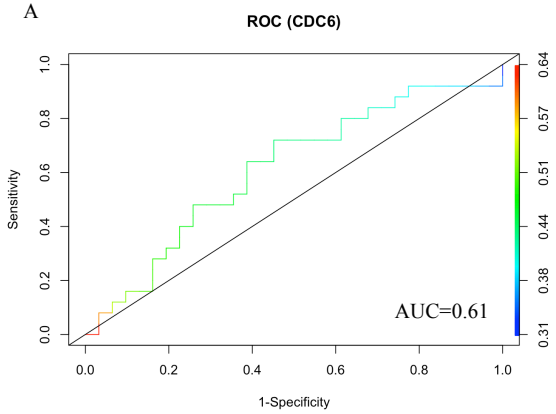
Now we have a homogenised transcriptomic dataset where I have run the ROC (receiver operating characteristic) regression curve for each candidate on the Tsuji subset of the combined combatted dataset. The optimal cut point for each biomarker has been calculated. For *CDC6*, the cut-off point is 0.43, the area under the curve (AUC) is 0.61 (sensitivity= 54.8% and specificity= 72%) (Figure 3.6A). For *NFE2L2*, the cut-off point was 0.45, the area under the curve (AUC) is 0.58 (sensitivity= 64.5%and specificity= 60%) (Figure 3.6B). For DNA repair signature score, the cut-off point is 0.63, the area under the curve (AUC) is 0.65 (sensitivity= 100% and specificity= 28%) (Figure 3.6C). Then I have applied the regression model in the FOCUS dataset and used the same cut-off identified in the Tsuji dataset in the FOCUS cases. Upon binarising the whole

combatted dataset, I then ran logistic regression analysis with binarised response as an endpoint in the FOCUS as an independent cohort. The number of patients in each category of response and each biomarker phenotype is shown in table 3.4. This analysis of the binarised variables shows non-significant results in the 5FUFA arm of FOCUS (Figure 3.7A), significant results of the three biomarkers (CDC6, NFE2L2, and DNA repair signature score) in the FOLFOX arm of FOCUS (Figure 3.7B), and significant interaction analysis between the 5FUFA arm and the FOLFOX arm of FOCUS. This result shows that those three biomarkers are predictive of response to oxaliplatin (Figure 3.7C). High *CDC6* expression (OR=4.35 (1.32- 14.31), P=0.0155) and high DNA repair signature score favours no oxaliplatin (OR=3.27 (1- 10.67), P=0.0496) favours no oxaliplatin, while high *NFE2L2* expression favours oxaliplatin (OR=0.16 (0.05- 0.58), P=0.0048).

	5FUFA		FOLFOX	
	Poor response	Good response	Poor response	Good response
High CDC6	76	101	11	11
Low CDC6	44	44	11	36
High NFE2L2	93	84	5	17
Low NFE2l2	38	50	26	21
High DNA sig	64	113	13	9
Low DNA sig	31	57	14	33

Table 3.4 Binarised *CDC6, NFE2L2,* and DNA signature phenotypes frequencies in each arm of FOCUS

A

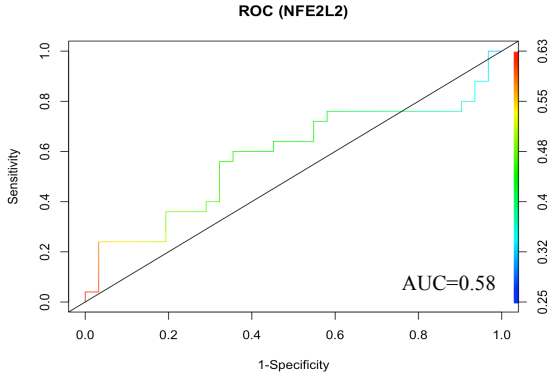


		Predicted Outcome		
		Good Response	Poor Response	Total
Observed Outcome	Good response	17 (TP)	14 (FN)	31
	Poor response	7 (FP)	18 (TN)	25

Sensitivity = 54.8%
Specificity = 72%

Cut off: 0.43

B

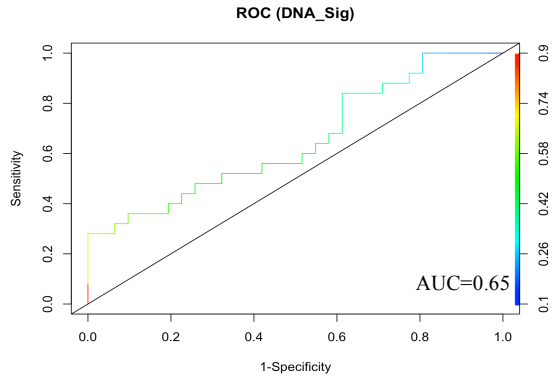


		Predicted Outcome		
		Good Response	Poor Response	Total
Observed Outcome	Good response	20 (TP)	11 (FN)	31
	Poor response	10 (FP)	15 (TN)	25

Sensitivity = 64.5%
Specificity = 60%

Cut off: 0.45

C



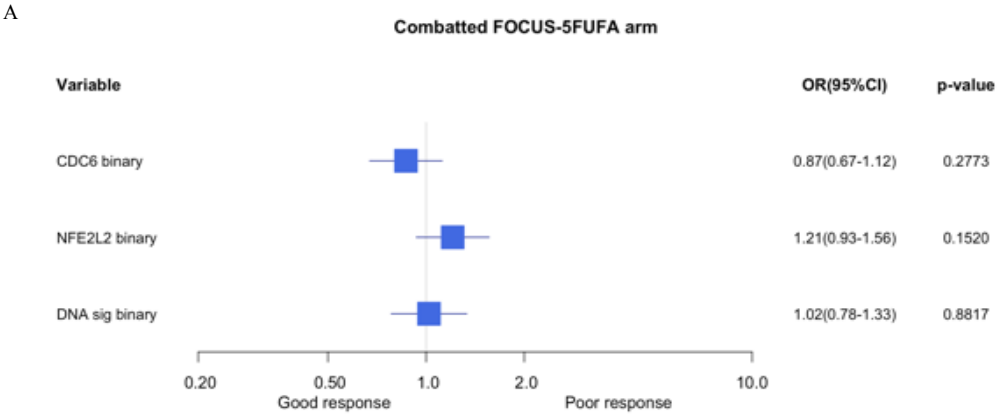
		Predicted Outcome		
		Good Response	Poor Response	Total
Observed Outcome	Good response	31 (TP)	0 (FN)	31
	Poor response	18 (FP)	7 (TN)	25

Sensitivity = 100%
Specificity = 28%

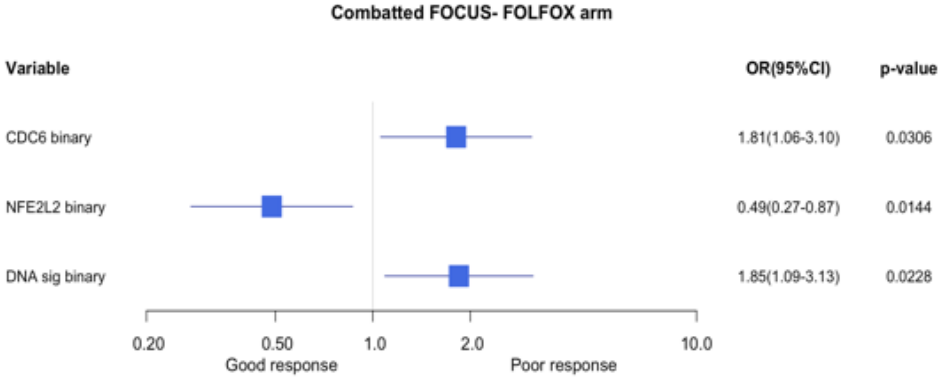
Cut off: 0.63

Figure 3.6 ROC regression curves for the 3 biomarkers. ROC regression curve for each candidate on the Tsuji subset of the combined combatted dataset. (A) For CDC6, the cut-off point is 0.43, the area under the curve (AUC) is 0.61 (sensitivity= 54.8% and specificity= 72%). (B) For NFE2L2, the cut-off point is 0.45, the area under the curve (AUC) is 0.58 (sensitivity= 64.5% and specificity= 60%). (C) For DNA repair signature score, the cut-off point is 0.63, the area under the curve (AUC) is 0.65 (sensitivity= 100% and specificity= 28%).

A



B



C

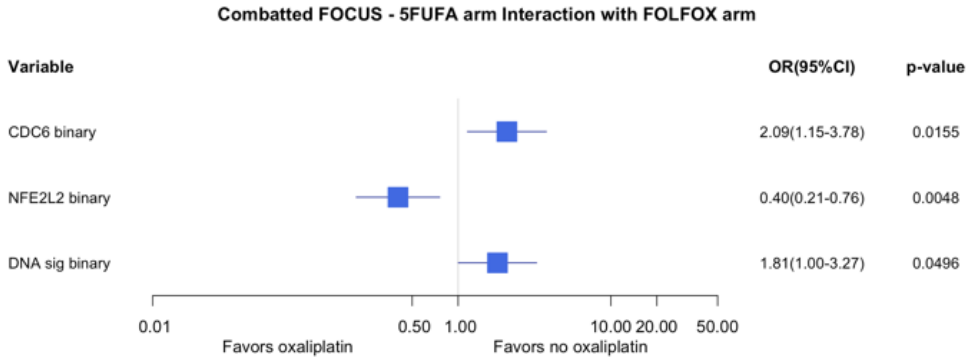


Figure 3.7 Logistic regression of binarised biomarkers in the FOCUS dataset. Logistic regression analysis with binarised response as an endpoint in the FOCUS shows flat results in the 5FUFA arm (A), significant results of the three biomarkers (CDC6, NFE2L2, and DNA repair signature score) in the FOLFOX arm (B), and significant interaction analysis between the 5FUFA arm and the FOLFOX arm of FOCUS where the three binarised biomarkers are predictive of response to oxaliplatin (C). High *CDC6* expression (OR=4.35 (1.32- 14.31), P=0.0155) and high DNA repair signature score favours no oxaliplatin (OR=3.27 (1- 10.67), P=0.0496) favours no oxaliplatin, while high *NFE2L2* expression favours oxaliplatin (OR=0.16 (0.05- 0.58), P=0.0048).

3.4.2 Cox regression with progression free survival as endpoint

Given the association with response as the endpoint, I further tested these 3 binarised biomarkers for progression free survival for oxaliplatin. Cox proportional hazard model of interaction between the two arms of FOCUS 5FUFA versus oxaliplatin showed no significance for CDC6 binarised score to predict progression free survival for MSS patients with metastatic colon cancer treated with oxaliplatin (HR= 1.26, 95%CI (0.93 – 1.71), $P_{interaction}$= 0.129) (Figure 3.8A/B). A similar result is observed for NFE2L2 binarised score (HR=1.068, 95%CI (0.81 – 1.42), $P_{interaction}$= 0.648) (Figure 3.8C/D). And finally, DNA repair signature score did not predict significantly PFS in MSS patients with metastatic colon cancer treated with oxaliplatin (HR= 0.99, 95%CI (0.74 – 1.33), $P_{interaction}$= 0.943) (Figure 3.8E/F)

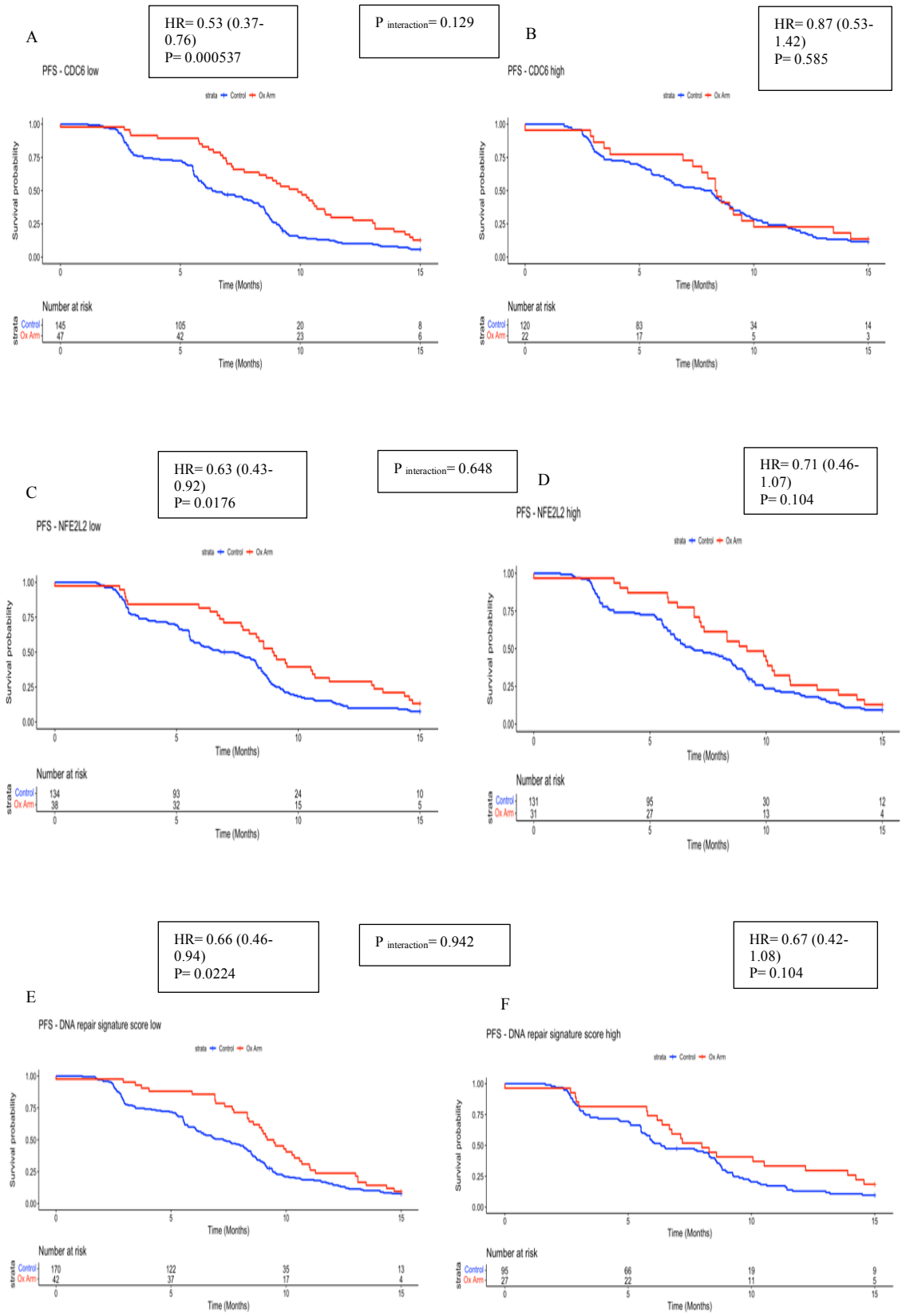


Figure 3.8 Survival analysis of the three binarised biomarkers. Cox proportional hazard model of interaction showed no significance for CDC6 binarised score to predict progression free survival for oxaliplatin ($P_{interaction}$= 0.129) (A/B). A similar result is observed for NFE2L2 binarised score ($P_{interaction}$= 0.648) (C/D). And finally, DNA repair signature score did not predict significantly PFS in MSS patients with metastatic colon cancer treated with oxaliplatin ($P_{interaction}$= 0.943) (E/F).

3.5 Summary

A systematic literature review has been conducted in search for biomarkers ascribed to oxaliplatin resistance or molecular mechanisms of resistance to the drug. The literature review identified 32 biomarkers that have been tested in the FOCUS trial to assess their role in oxaliplatin response. Results have shown that *BMP4*, *CXCL8*, and *AKT1* high expression is associated with poor response to 5FUFA arm in colon cancer metastatic patients while *NR1I2* high expression is associated with good response to oxaliplatin and DNA repair signature score is statistically significantly associated with bad response to oxaliplatin. Interaction analysis between the two arms of the FOCUS dataset shows *CDC6* high expression and high DNA repair signature score predict worse response to oxaliplatin while high expression of *NFE2L2* predict better response when treated with oxaliplatin than 5FUFA. Using the Tsuji dataset to binarise the three predictive biomarkers for them to be more clinically relevant, the significance to predict response to oxaliplatin was retained with the binarised biomarkers. Analysis of these 3 binarised biomarkers to predict PFS to oxaliplatin showed no significance.

Chapter 4 Biomarkers for Oxaliplatin Stratification: Hypothesis Free Analyses

After exploring important biology colorectal cancer biomarkers and biomarkers recently reported in the literature and their role in oxaliplatin resistance, I have taken a different approach in this chapter. In an attempt to find a molecular biomarker for oxaliplatin, I have extracted RNA from formalin fixed paraffin embedded (FFPE) blocks of colorectal cancer patients from a relevant cohort, that will be explained in details in the next sections, profiled and interrogated the transcriptome and performed differential gene expression analysis to develop a new RNA signature for further validation.

4.1 Methods

4.1.1 COIN

COIN is a randomised controlled clinical trial that enrolled patients with confirmed adenocarcinoma for the colon or rectum, had inoperable metastatic disease and were fit for but had not received previous metastatic chemotherapy(Adams, 2011; Msc, 2011). The trial constituted of three arms: arm A (n=815) is the control arm of the trial and had patients on continuous oxaliplatin-based chemotherapy (oxaliplatin plus capecitabine or oxaliplatin plus fluorouracil and folinic acid), arm B (n=815) is the first interventional arm of the trial where patients received continuous chemotherapy plus cetuximab, and arm C (n=815) is the second interventional arm of the trial where patients received intermittent oxaliplatin based chemotherapy for 12 weeks then treatment is stopped, monitoring is conducted, if progression is noted assessed by RECIST then treatment is restarted for another 12 weeks. Oxaliplatin plus capecitabine was given as a 3-weekly regimen of intravenous oxaliplatin 130 mg/m^2 over 2 h followed by oral

capecitabine 1000 mg/m^2 twice a day for 2 weeks. Oxaliplatin plus fluorouracil and folinic acid was given as a 2-weekly regimen of intravenous L-folinic acid 175 mg or D,L-folinic acid 350 mg over 2 h given concurrently with oxaliplatin 85 mg/m^2 over 2 h, followed by intravenous bolus fluorouracil 400 mg/m2, and finally fluorouracil 2400 mg/m2 infusion over 46 h via an ambulatory pump (Adams, 2011).

4.1.2 Response Criteria in COIN

The response criteria used for CT scans followed RECIST 1.1 version (The Response Evaluation Criteria for Solid Tumours), similar to FOCUS. Radiologists call complete response (CR) when the target lesion completely disappears with any lymph node reduced to <10mm. Partial response (PR) is when the tumour has decreased by at least 30% in diameter in reference to the baseline tumour/lesion size. A response is called as progressive disease (PD) if the tumour has increased by 20% (with an absolute value of at least 5mm) or if there is the appearance of one or more new lesions. A response to treatment was called as stable disease (SD) if there is not sufficient reduction to call it PR neither enough increase to call it PD (Eisenhauer, 2009). For our purpose of finding a biomarker for oxaliplatin, samples have been selected from arms A and C. This is because arm B would add a strong confounder with the addition of the EGFR inhibitor cetuximab. Patients who progressed within the first 12 weeks have been labeled as primary progressors (n=89) while patients who progressed after 18 weeks have been labeled as non progressors (n=576) and their response based on RECIST categorized as stable disease patients, partial responders or complete responders. Those who progressed between 13 and 17 weeks (n= 965) have been excluded from the selection process. Only primary resections were kept for analysis. From the 89 primary progressors, 20 had no available tumour block and 9 have metastasized blocks thus removed for a final number

of 60 primary progressors. As for the non progressors, 150 out of 576 samples were picked randomly for sequencing (Figure 4.1).

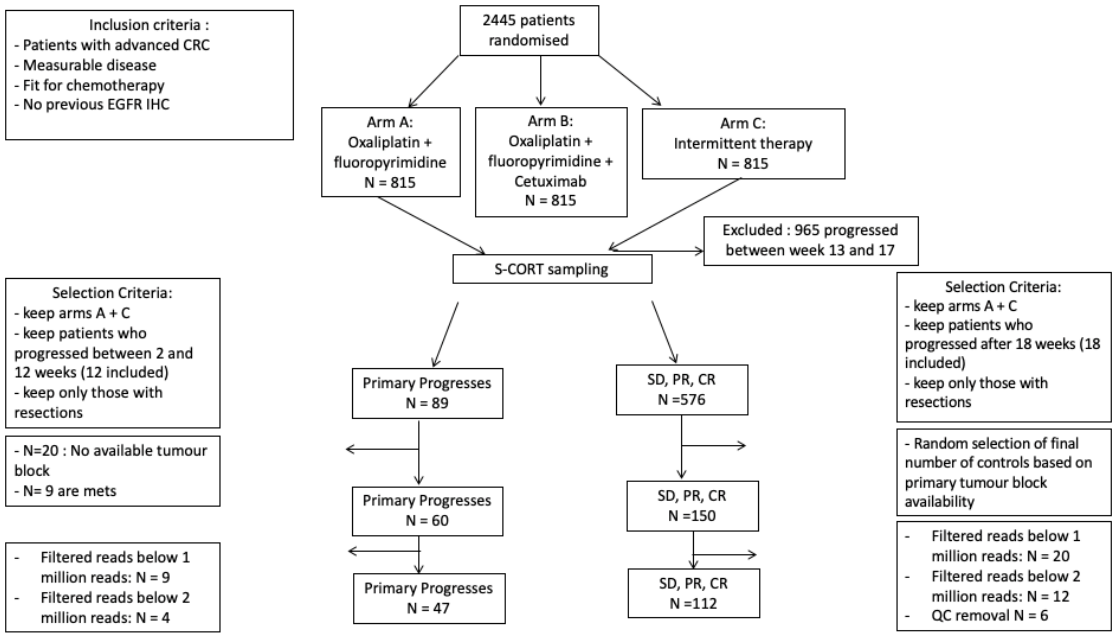


Figure 4.1 COIN consort diagram. COIN is a randomised controlled clinical trial that enrolled patients with confirmed adenocarcinoma for the colon or rectum and is comprised of three arms. Arm A (n=815) had patients on continuous oxaliplatin-based chemotherapy, arm B (n=815) had patients on continuous chemotherapy plus cetuximab, and arm C (n=815) had patients who received intermittent oxaliplatin based chemotherapy. For this chapter analysis, samples have been selected from arms A and C. Patients who progressed within the first 12 weeks have been labeled as primary progressors (n=89) while patients who progressed after 18 weeks have been labeled as non progressors (n=576). Those who progressed between 13 and 17 weeks (n= 965) have been excluded from the selection process. Only primary resections were kept for analysis. From the 89 primary progressors, 20 had no available tumour block and 9 have been metastasized blocks thus removed for a final number of 60 primary progressors and 150 non progressors. I have extracted RNA from 210 FFPE blocks and sent them for 3'sequencing. I then performed quality control assessment on the counts data received. Reads below 2 million have been removed with an addition of 6 samples with poor quality to have a final number of N=47 primary progressors and N=112 of non progressors available for downstream analysis.

4.1.3 RNA extraction

For the COIN trial, FFPE blocks containing CRC were identified, sectioned and processed by personnel at Wales Cancer Bank. Presence of tumour in the sample was confirmed by a pathologist. five slides of five um thickness were cut. One slide has been processed for H&E to mark the tumour only on the slide and guide RNA extraction. The H&E slides are assessed and marked by drawing a circle around tumour cells to highlight a dense area of tumour cells on the slide to increase yield when macro dissecting prior to extraction. There still may be inflammatory cells, stroma, muscle, blood vessels, lymphatics, neural and normal glandular cells within that, but the highlighted area will contain the greatest density of tumour cells in the tissue section (Dr. Alison Perry-Jones).

I extracted RNA from 3 slides per patient, only from the marked region on the slides that contains the tumour, using Roche High Pure FFPE RNA Isolation kit (Lot number 6650775001). The first step is deparaffinization of the tissue of 5µm thickness. Dip the slides in xylene for two minutes, tap excel liquid then place the slides in 100% ethanol for one minute, then 96% ethanol for another minute, followed by 70% ethanol one minute then water for the final one minute. Following this step, the tissue is soft and ready to be scraped off the slides. 50 µL of RNA tissue lysis buffer was used to scrape the tissue from the slide, then another 50 µL of the same lysis buffer with 16 µL of 10% SDS and 40 µL of proteinase K are added to the 1.5mL tube. Vortex, spin down briefly and incubate for 30 minutes at +85^{o}C while shaking on 600 rpm. The tube is set aside to cool down for few minutes before the addition of 80 µL of proteinase K. The tube is vortexed, spun down briefly and incubated for another 30 minutes at +55^{o}C while shaking at 600 rpm. The lysate will look clear at this point. Then 325 µL of RNA binding buffer and 325 µL of absolute ethanol are added to the tube. After vortexing, the lysate is pipetted into the upper reservoir of High Pure filter tube assembly. The tube is centrifuged at 6000

g for 30 seconds, the collection tube is discarded, and the filtered tube is placed in a new collection tube to centrifuge again for two minutes at 16,000 g to dry the filter. 100 µL DNAase working solution is added onto the filter, and the tube is incubated for 15 minutes at +15 to +25ºC. 500 µL Wash Buffer I is added to the tube that is centrifuged for 20 seconds on 6000 g, the flow through is discarded. Next, 500 µL of Wash Buffer II is added to the tube that is centrifuged for 20 seconds on 6000 g, the flow through is discarded. This step is repeated twice. Finally, the tube is centrifuged to dry the filter for two minutes at 16,000 g. The filter tube is placed in a fresh clean 1.5 mL tube, with 30 µL of elution buffer added to it and incubated for five minutes at +15 to +25ºC. The tube is centrifuged at 6,000g for one minute to collect pure total RNA. I quantified the RNA by Nanodrop where also A260/A280 and A260/230 ratios are measured and impurities may be detected. Then RNA is quality controlled and checked by The Qubit® RNA HS Assay using the kit (lot number: Q32852, Q32855). Standards should be prepared to calibrate the Qubit® Fluorometer. The assay requires two standards. The Qubit working solution is prepared by diluting the Qubit RNA HS Reagent 1:200 in Qubit RNA HS Buffer in a clean plastic tube. The final volume in each tube must be 200 µL. Each standard tube requires 190 µL of Qubit working solution, and each sample tube requires 199 µL. Every run, sufficient Qubit working solution is prepared to accommodate all standards and samples. 10 µL of each standard is added to the working solution and 1 µL of each sample is added to the working solution for a total of 200 µL in each tube. On the home screen of the fluorometer, RNA is selected from the nucleic acid types, then RNA high sensitivity as the assay type. Following this step, the standards are read for machine calibration. After this step, the samples are placed in the samples chamber, the lid is closed and the samples can now be run and read, one at each time. The fluorometer gives the values in ng/mL where the reading is multiplied by 200 and divided by the number of

microliters of sample added to the assay tube. It automatically calculates the concentration in the original sample and displays it. I re-extracted and re-quantified samples with a concentration lower than 20 ng/μL on the Qubit. The majority of the samples were normalized to a concentration of 100 ng/μL which corresponded to 1.2-1.5 μg of total RNA except for very few cases where the concentration was less but not lower than 20 ng/μL.

4.1.4 3'mRNA Sequencing

I sent RNA samples to the Oxford Genomics Centre (OGC) for sequencing. Pre-sequencing, samples were further put for quality control using Agilent High Sensitivity RNA Screen Tape System (catalogue number: 5067-5579). First, the 2200 Tape Station Controller Software is launched, then High Sensitivity RNA Screen Tape device and loading tips are loaded into the instrument. RNA protocol is selected and the reagents are set to equilibrate at room temperature for 30 minutes (sample buffer cat number 5067-5580 and ladder cat number 5067-5581). RNA samples are thawed on ice after which 2 μL of sample is mixed with 1 μL of buffer. Same quantity is used to prepare the ladder tube. The tubes are vortexed at 2000rpm for one minute, and then samples are denatured at 72^{o}C for three minutes then placed on ice for two minutes. Finally, samples are loaded into the instrument for total RNA reading. Libraries are normalized using 10 mM Tris-HCl, pH 8.5 with standard library quantification and quality control procedures. Libraries are denatured for sequencing using 0.2N NaOH and incubated for eight minutes at room temperature. Improper denaturation can decrease yield. 400 mM Tris-HCl, pH 8.0 is added to neutralize the solution. The libraries are vortexed briefly, then centrifuged down at 280 g for one minute. The full volume of denatured library is transferred to the library tube provided with the NovaSeq 6000 Reagent Kit. Immediately the library tube is loaded

into the cluster cartridge and the run is set up for 100 cycle lane on the NovaSeq 6000 sequencer.

4.1.5 Quality control of RNA sequencing

Raw data were received from the OGC and downloaded to the S:CORT server. First step is to use MultiQC to produce and aggregate all the outputs from different tools to assess quality control metric and generate one html file detailing the quality of all the libraries. Second step is assigning a threshold value to low quality base calls to perform quality trimming. One or both ends of a low-quality sequencing read is trimmed to sustain a region with high quality bases. RNASeq reads are mRNA reads that map to exons and thus the aligner should be a splice aware mapper to splice the introns. Third step is to supply reference genome sequences in the form of FASTA files and annotations in the form of GTF files from which STAR was used to generate genome indexes. Those indexes are used to map RNAseq reads to the genome. The STAR (Dobin, 2013) generates output alignment files in the form of sam which are files in a format that contains the alignment information of various sequences that are mapped against the reference sequences or bam which are the binary version of sam files, mapping summary statistics, unmapped reads and several other outputs.

I have accessed the multi-QC files to quality control the read counts. I have removed filtered reads below 1 million reads (n=29) for poor quality. To make sure of the quality of the remaining data, I have conducted some data exploration to prepare the data for differential gene expression. Taking the raw counts (57,773 entries linked to Ensembl IDs), all genes where the total number of reads across all samples is greater than 5 have been kept and everything below 5 have been filtered out (29,523 remaining entries after filtering linked to Ensembl IDs). First, I plotted raw counts to assess distribution (Figure

4.2 A-B). Because raw counts are large and because the variance increases with mean gene expression, data transformation is necessary to study differential gene expression. I plotted log transformation to check the distribution of the read counts (Figure 4.2 C). Another kind of transformation called variance stabilizing transformation (VST) was also assessed. VST is a transformation that calculates the fitted dispersion mean relation and this is specifically important for low mean samples where it generates a matrix of values for which variance is constant across all mean values. I plotted VST to assess the distribution and data quality of the read counts (Figure 4.2D). From those Figures, assessment was made for further removal of samples below 2 million reads (n=165 remaining samples after filtering: 47 progressive disease patients (PD) and 118 non-PD). All data exploration and quality check are reported in Figure 4.3 A-D. From the log transformation and the vst distribution plots, 6 samples showed very low expression patterns more likely due to technical issues than true biological patterns and were excluded for further analysis to have the final count of COIN samples at 159 (47 PD and 112 non-PD). Data visualization after all QC steps are reported in Figure 4.3 E-F.

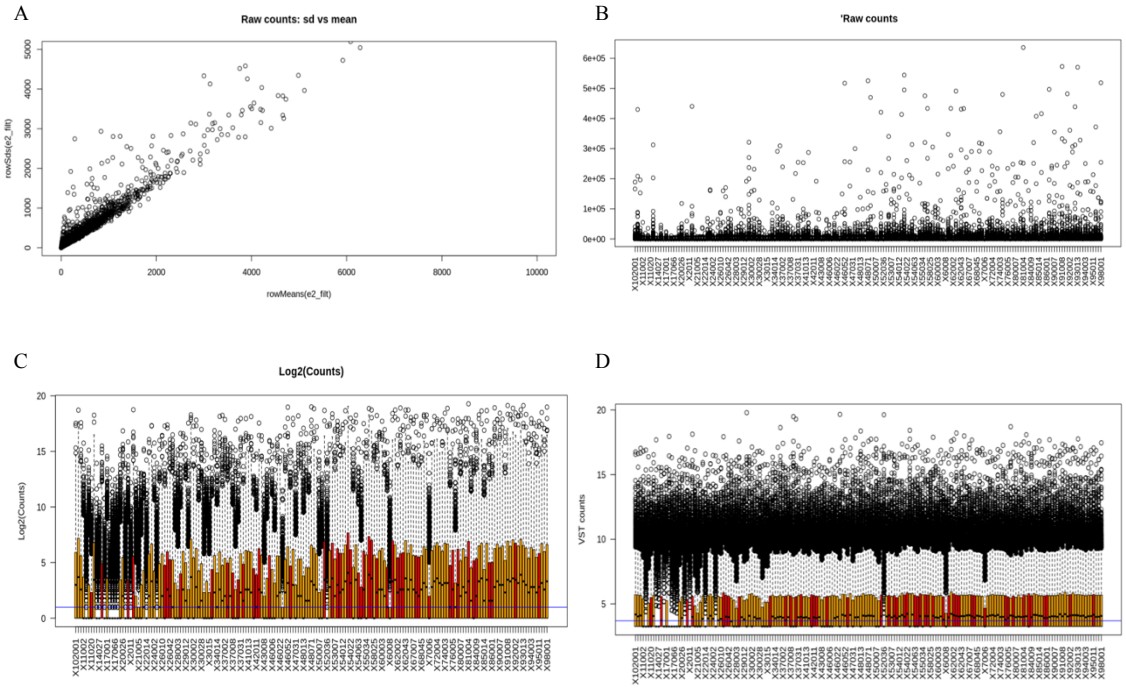


Figure 4.2 RNA counts exploration and assessment. Taking the raw counts (57,773 entries linked to Ensembl IDs), all genes where the total number of reads across all samples is greater than 5 have been kept and everything below 5 have been filtered out (29,523 remaining entries after filtering linked to Ensembl IDs). (A/B) Raw counts have been plotted to assess distribution after removing low quality reads below 1 million reads. (C) Log transformation has been plotted to check the distribution of the read counts. (D) VST is a transformation that calculates the fitted dispersion mean relation and this is specifically important for low mean samples where it generates a matrix of values for which variance is constant across all mean values. VST has been plotted to assess the distribution and data quality of the read counts.

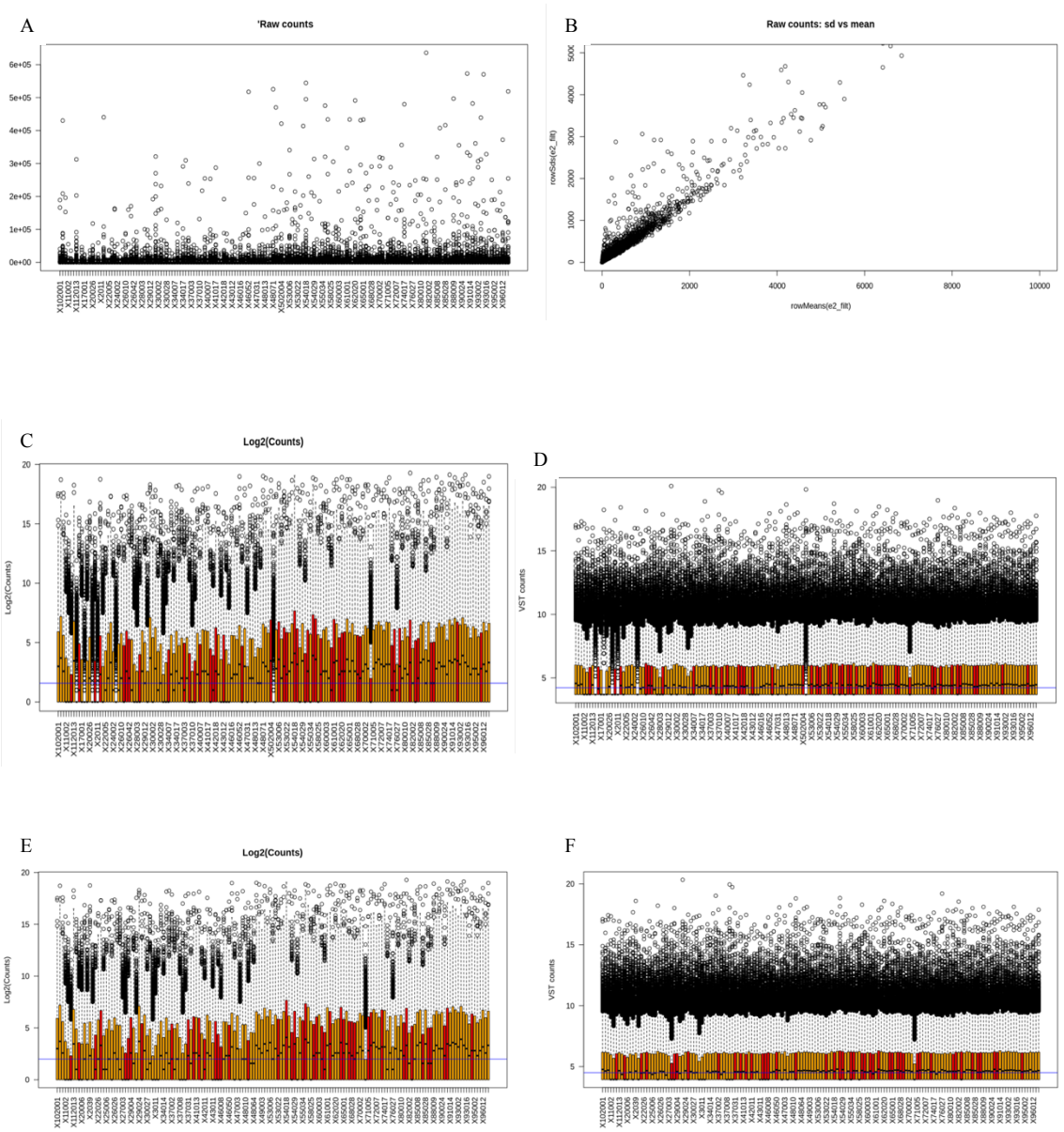


Figure 4.3 RNA counts exploration and assessment-part 2. (A/B) Raw counts have been plotted to assess distribution after removing low quality reads below 2 million reads. (C) Log transformation has been plotted to check the distribution of the read counts. Further 6 samples show poor quality. (D) VST has been plotted to assess the distribution and data quality of the read counts confirming the 6 poor quality samples detected in log2 graph. (E) Final log transformed graph after removal of all poor-quality samples. (F) Final VST transformed graph after removal of all poor-quality samples.

4.1.6 Differential gene expression analysis (DESeq2)

Differential gene expression analysis on the COIN trial was conducted using the R package DESeq2. RNASeq provides quantitative records in the form of read counts data. Each read reflects a target transcript mapped to a target gene. The higher the read

count, the more abundant this transcript is. The interest of studying RNASeq lies in comparing the read counts to different biological processes. DESeq stands for Description Estimate variance-mean in count data and is a package that analyses this count data, inputted as a matrix with non-negative integers: one row for each observational unit and one column for each sample. The package processes the normalisation, visualisation and differential expression analysis of count data. To calculate log fold change and dispersion, this package uses empirical Bayes techniques. The differential expression function (DESeq) performs the analysis through 3 steps: it estimates size factors, it estimates dispersion and finally performs differential analysis using a gerenalised linear model with negative binomial (NB) distribution.

Kij ~NB(μij,αi) where Kij is the counts for gene i and sample j, μij is the fitted mean, and αi is the gene-specific dispersion parameter. After running this function, it returns a DESeqDataSet object where the log2 fold changes and p-value results can be extracted (Love, 2014). The package is available via Bioconductor and can be conveniently installed with biocLite.

4.1.7 Differential expression from RNA arrays (Limma)

Almac and Queens provided RNA profiling data as raw CEL files, for both FOCUS and FOxTROT, which were uploaded to Oxford using the data upload portal on the 17th February 2016. For differential gene expression analysis on the FOCUS trial, I have used the limma package from R (https://github.com/cran/limma). Limma stands for linear models for microarray where linear models are used to assess differential expression analyses. One advantage of limma is the ability to report stable result despite having experiments with small size of arrays. I have used the Bioconductor package affy that allows reading (readaffy) and normalisation (rma) of Affymetrix CEL file data to

form an expression set or matrix object with log transformed expression values. I have filtered out low expression probes by assigning a cut off (80% of samples are above the median of median of the expression) where a probe is removed if not expressed in 80% of the samples of the FOCUS dataset. Then I generated a design matrix with one column having the first condition tested (in this analysis PD) and a second column is the contrast matrix with the other condition being studied (Non-PD). A linear model is then fitted based on the design matrix, limma smoothens standard errors of the log-fold changes by using an empirical Bayes method, and finally differential analysis reveals the top genes between the first and second groups compared. By having this flow of analysis, limma has the power to give stable results irrespective of array size (Ritchie, 2015).

4.1.8 Differential DNA methylation (ChAMP)

Cytosine methylation (5-mC) and hydroxymethylation (5-hmC) have major consequences on gene expression and chromatin spatial remodelling. 5-mC forms by the addition of -CH_3 to N5 of cytosine, if it occurs in or around promoter areas, it may lead to gene silencing and transcription repression. CpG islands are regions in the genome which are > 500 base pairs (bp) and are rich in GC content (55%) and CpG sites. They are highly abundant in promoter regions (40% of promoters contain CpG islands). Next to CpG islands, there are CpG shores which are ~2KB from islands and usually contain the highest tissue specific differentially methylated regions (DMR). Next to CpG shores are the CpG shelves which are ~4KB from islands(Guibert & Weber, 2013).

For differential DNA methylation analysis, I used the ChAMP package in R. ChAMP stands for Chip Analysis Methylation Pipeline and allows the pre-processing, normalisation and analysis of Illumina Methylation beadarray data (450K or EPIC). Idat files (which are files that store beadarray data) have been received along with the sample

sheet which is a file in .csv format that shows the location of the samples within each beadchip. I have loaded those files and read them into ChAMP. The loading function is a duel function that performs importing and filtering. ChAMP will filter out first probes that have a detection p-value above 0.01, thus ChAMP would consider those probes as failing probes. Second, probes with <3 beads in at least 5% of samples per probe are removed. Third, it will, by default, filter out all non-CpG probes contained in the dataset. Fourth, it will filter all SNP-related probes (Zhou, 2017). Fifth, all multi-hit probes will be removed and finally, all probes located in chromosome X and Y will also be removed. Then there is a further QC step to assess if the data is suitable for consequent analysis, with a special function (QC.GUI) that generates QC plots for visualisation and data quality checking. Following this, there is a normalisation step. This step is very crucial because probes on beadarrays come in two distinctive designs (type-I and type-II) and they hybridize differently and thus have different distributions due to technicality and not biological significance, thus it is so crucial to correct for such a technical difference. Specifically, type-II probes have a shorter range compared to type-I and this would create a technical bias which can be corrected for using the normalisation function in ChAMP. Post this step, QC.GUI is run again to visualise the data quality after normalisation. This is followed by the singular value decomposition method (SVD) for methylation data which is a powerful tool for assessing the number and nature of the significant components of variation in a dataset with some of these variations correlating to batch effects. After all those pre-processing steps, analysis is started by running a function (champ.DMP) from the package to extract the differentially methylated probes (DMPs). This function works on numerical variables where it runs linear regression on CpG regions to try and find covariates CpGs, and it can also work on categorical variables where it runs contrast comparison between the different groups of the variable. DMP.GUI

function is run afterwards to assess and visualise the results. A panel will be displayed having a table with all the DMPs. There is another section of the panel with P-value and logFC as cutoffs where the latters can be specified and a job submitted for specific analysis. Another option on the panel is a tab with gene name where a specific gene name can be inputted and visualised or finally the examination can be through inputting a CpG ID. A second tab displays a heatmap of the parameters set from the first tab. The third tab shows three different bar plots, one that shows where the DMPs fall in terms of the CpGs whether on islands, shelves, shores or opensea, another bar plot shows the position of the DMPs whether in untranslated regions (3' or 5'), first exon, transcription start sites or body. Finally, a third bar plot combining the information from first and second bar plots for complete visualisations and assessment. The fourth tab on this panel displays the gene selected in first tab where there is information about it and how the DMPs in that particular gene look. Finally, the last fifth tab shows an enrichment plot with the top 70 genes with hyper and hypo differentially methylated CpGs within that gene. A very similar analysis is conducted on DMRs with similar functions and results assessment (The Chip Analysis Methylation Pipeline, n.d.; Zhou, 2017).

4.1.9 Analysis of unselected mutations and copy number alterations (regression)

The targeted NGS panel contains probes spanning SNPs evenly distributed along the human genome (average of 1 SNP per 3 Mb) and also 66 chromosomal regions recurrently gained or lost in CRC. This design allows generation of copy number estimations from targeted NGS at low resolution, mindful it may not be comparable to techniques interrogating the whole genome. The quality of the analysis may be further enhanced by using CNVkit, a bioinformatical tool specifically designed to be used on

targeted NGS data by analysis of both targeted and off-target reads (Talevich, 2016). Accordingly, CNVkit was used adjusted by tumour purity that had been estimated using the RNA signature Estimate (Yoshihara, 2013). Copy number segments with estimations ≥3 were classified as gain, 2 as neutral and ≤1 as loss. In order to get CNV calls at the chromosome arm level, the length of all the segments for each of these 3 copy number categories was added. The one with highest total length in each chromosome arm was called. To generate Chromosomal Instability (CIN) calls, it was measured by calculating the sum of the lengths of calls for either loss or gain divided by the whole length. CIN calls were then performed with a threshold of >0.2 as previously defined (Burrell, 2013).

I analysed unselected mutations and copy number variation using regression analysis. Similar to what was done in the previous chapters, I conducted logistic regression analysis with binarised response to treatment (5FUFA arm versus FOLFOX treatment arm) as an endpoint using the function glm (https://cran.r-project.org/web/packages/cranly/vignettes/glms.html) from the stats R library. I scaled all tested variables from 0 to 1 to make them statistically comparable. Copy number (CN) is a categorical variable comprised of three categories: neutral coded as zero, gain coded as 1 and loss coded as -1 in the analysis. This analysis is exploratory rather than candidate analysis originating from a hypothesis, thus for those reasons false discovery rate (FDR) which is a statistical test to correct for false positive rate when performing hypothesis testing. This method of adjustment is particularly important in health and medical studies (Glickman, 2014).

4.2 Transcriptomic signatures

4.2.1 Signature 1: generation in COIN

First, DESeq2 analysis has been conducted using counts Ensembl IDs in COIN MSS cases but yielded only two genes (*MUC5AC* and *ANGPTL4*) with an insignificant correlation with outcome when this signature was generated. So next, I have conducted DESeq2 analysis using counts Ensembl IDs in 159 COIN trial samples comparing 47 PD samples to 112 Non-PD samples. This yielded a list of 29 differentially expressed genes (DEGs) considering the p-values < 0.05 as attained by the Wald test and corrected for multiple testing using the Benjamini and Hochberg method (Table 4.1). Pathway analysis, using over representation analysis with Kegg pathways was conducted to check if those genes belong to a common pathway, but it was not informative. Furthermore, Gene Set Enrichment Analysis (GSEA) was performed by sorting the logFC of the genes and using gseKEGG() function for KEGG pathways as GSEA gene sets but again the analysis did not yield any informative outcome.

A signature, termed here signature one, was generated combining all DEGs with log2 fold change above zero as positive, and all DEGs with log2 fold change below zero as negative. Scores for this signature were generated by subtracting the mean of expression in the positive DEGs from the mean of expression of the negative DEGs. As a confirmation test, this signature was tested on outcome (12-week response) in COIN and yielded a significant result (OR=3.66, 95% (2.23 -5.99), P=0.00) with higher score associated with poor response to oxaliplatin (Table 4.2).

X	baseMean	log2FoldChange	lfcSE	stat	pvalue	padj	geneSymbols
ENSG00000107954	18.54658952	1.766382782	0.356138975	4.959813184	7.06E-07	0.005346068	NEURL1
ENSG00000185479	30.70190944	-2.512961908	0.518764361	-4.844129813	1.27E-06	0.005346068	KRT6B
ENSG00000259974	91.438626	1.949666038	0.402988783	4.838015651	1.31E-06	0.005346068	LINC00261
ENSG00000112303	4.56690601	1.637565211	0.339999048	4.816381747	1.46E-06	0.005346068	VNN2
ENSG00000105697	5.619986718	1.942779989	0.40360077	4.813618144	1.48E-06	0.005346068	HAMP
ENSG00000106541	580.1433913	1.005472787	0.210783293	4.770173057	1.84E-06	0.005532462	AGR2
ENSG00000167772	39.33585768	1.219886495	0.263434492	4.630701486	3.64E-06	0.009388731	ANGPTL4
ENSG00000221944	15.59583945	1.011867592	0.222927761	4.53899321	5.65E-06	0.0127418	TIGD1
ENSG00000101695	25.06056055	0.923484511	0.204821008	4.508739208	6.52E-06	0.013067439	RNF125
ENSG00000215182	121.853454	2.528354834	0.572799148	4.414033863	1.01E-05	0.016662982	MUC5AC
ENSG00000154451	32.98490905	1.226512726	0.277890145	4.413660398	1.02E-05	0.016662982	GBP5
ENSG00000115523	59.64665462	1.123702257	0.25647109	4.38139931	1.18E-05	0.017721341	GNLY
ENSG00000078596	13.39379433	1.09341838	0.251632725	4.345294829	1.39E-05	0.018430636	ITM2A
ENSG00000169715	100.5620577	0.994145691	0.229114289	4.339082007	1.43E-05	0.018430636	MT1E
ENSG00000111732	3.757034839	1.526103154	0.353472548	4.317458774	1.58E-05	0.018976081	AICDA
ENSG00000181773	2.816154602	1.610962163	0.379695773	4.242770866	2.21E-05	0.024104125	GPR3
ENSG00000153395	110.471051	0.620843895	0.146552916	4.236312121	2.27E-05	0.024104125	LPCAT1
ENSG00000236438	4.488792351	-2.563919411	0.607620762	-4.219604684	2.45E-05	0.024519334	FAM157A
ENSG00000146205	18.23352125	1.539242461	0.366485913	4.200004437	2.67E-05	0.02533395	ANO7
ENSG00000074695	436.580665	0.415016709	0.099096684	4.187997949	2.81E-05	0.025376188	LMAN1
ENSG00000177627	7.714951257	-1.072562069	0.260088905	-4.123828611	3.73E-05	0.031999719	C12orf54
ENSG00000163449	6.010419731	-1.329533701	0.326691823	-4.069687728	4.71E-05	0.03858964	TMEM169
ENSG00000090382	462.1610975	1.169996327	0.289085656	4.047230649	5.18E-05	0.040183653	LYZ
ENSG00000265681	313.6101086	0.450560496	0.111548519	4.039143691	5.36E-05	0.040183653	RPL17
ENSG00000010219	32.24489483	0.800712289	0.198673061	4.030301273	5.57E-05	0.040183653	DYRK4
ENSG00000258655	3.458901702	1.282656955	0.321290317	3.992205456	6.55E-05	0.045405185	ARHGAP5-AS1
ENSG00000187653	4.903378004	1.219526221	0.306938781	3.973190415	7.09E-05	0.046318153	TMSB4XP8
ENSG00000047457	7.978356781	1.846812472	0.465208366	3.969860839	7.19E-05	0.046318153	CP
ENSG00000229950	6.299615732	1.636937441	0.414730746	3.946988394	7.91E-05	0.049214408	TFAP2A-AS1

Table 4.1 Signature 1: generation in COIN. Using DESeq2 package for differential gene expression analysis, 29 genes have come up significantly when comparing PD patients versus non-PD in the COIN dataset. This set of genes was used to generate a signature termed signature 1 where all positive log2FoldChange genes have meaned together and all the negative log2FoldChange genes have meaned together and the final score of the signature is calculated by subtracting the two means.

	Cohort	OR	95% CI	p-value
Sig 1	COIN	3.66	2.23-5.99	0.000
	FOCUS-5FUFA	0.98	0.44-2.20	0.962
	FOCUS-Ox	1.13	0.10-12.67	0.919
	FOxTROT	2.83	0.86-9.30	0.086

Table 4.2: Testing signature 1 in COIN, FOCUS, and FOxTROT. Signature 1 has been generated in COIN dataset. Outcome analysis shows that it is significant in COIN. Outcome analysis shows the signature insignificant in 5FUFA arm of focus, in FOLFOX arm and neither in interaction analysis in the two arms of FOCUS. Signature 1 had the same trend in FOxTROT biopsies where high score of signature 1 is associated with poor response to oxaliplatin.

4.2.2 Signature 1: test in FOCUS and FOxTROT

Using the DEG list from COIN, I have generated scores both in FOCUS and FOxTROT in the same way mentioned previously. In FOCUS, I linked significant Ensembl IDs from COIN to entrez IDs using biomaRt R package with Genome Reference Consortium Human Build 37 (GRCh37) (reference assembly accession GCF_000001405.13) to get matches in FOCUS. Genes linked with multiple probesets have been collapsed using mean per entrez id. Upon calculating the signature in FOCUS, 2 entrez IDs were lost. Those are entez ID 7117 (ENSG00000187653) corresponding to *TMSB4XP8* and entrez ID 101928670 which corresponds to *FAM157A* is not found on Xcel array. The first analysis I did was to test signature one in the MSS cases of FOCUS. This signature has been tested using logistic regression with binarised response as an endpoint. The binarisation has been done taking PD samples as poor response while CR, PR and SD (Non-PD) as good response. In the 5FUFA arm (PD = 70, Non-PD = 211) the analysis has not yielded significant results (OR = 0.98, 95%CI (0.44- 2.20), P = 0.96) (Table 4.2). In the oxaliplatin MSS arm (PD = 7, Non-PD = 63), signature one has not yielded significance (OR = 1.13, 95%CI (0.10- 12.67), P = 0.92) (Table 4.2). Interaction analysis (PD = 81, Non-PD = 274) has been conducted between the MSS 5FUFA arm and the MSS oxaliplatin arm and signature one was not predictive of oxaliplatin response (OR = 1.04, 95%CI (0.08-13.07), P = 0.98) (Table 4.2).

Despite the fact that this thesis focuses on MSS cases only, it was logical to test this signature in all samples and compare to the MSS setting since signature one has been generated in all COIN samples and not MSS only cases. In the randomised FOCUS trial, such signature used as a continuous score in patients treated with single agent 5FUFA (poor responders n= 76, good responders n = 211) did not show any signal (OR=1.09 (0.51-2.35), P=0.82). The FOLFOX arm was statistically underpowered (7 poor

responders/67 good responders) but it did show a trend in the expected direction (OR=3.25 (0.38-27.72), P=0.28) where the interaction with the 5FUFA arm was inevitably not significant (OR=2.84 (0.14-57.07), P=0.49) (Table 4.3) .

In FOxTROT, I linked the 29 Ensembl IDs from signature one to entrez ID using biomaRt R package in the exact way that was described previously. FOxTROT has been presented in chapter 1 section 1.7. Genes linked with multiple probesets have been collapsed using mean per entrez ID. The same number of genes (n = 2) that have been lost upon signature generation in the FOCUS trial have been lost in FOxTROT because expression of both trials was done on the Xcel array. I tested signature one in FOxTROT MSS using logistic regression binarised to treatment as an endpoint. Marked regression along with moderate regression were combined as good response while mild regression with no regression were joined as poor response to oxaliplatin. The first analysis has been done on the biopsies of FOxTROT (poor responders n = 64, good responders n = 24), the result shows the same trend observed in COIN of higher signature score associated with poor response (OR = 2.83 (0.86- 9.30), P = 0.09) (Table 4.2). Taking all samples in FOxTROT rather than MSS cases only (poor responders n= 69, good responders n=24), the signature was significantly associated with poor response (OR=3.31 (1.03-10.60), P=0.04) (Table 4.3). This is particularly interesting and promising as FOxTROT is a neoadjuvant clinical trial and a biomarker of oxaliplatin lack of response might be particularly important clinically in neoadjuvant treatment.

Signature one has been generated from all COIN samples and not only MSS.

Dataset	Sig 1 in MSS- results	Sig 1 in all samples- results
FOCUS-5FUFA	OR=0.98 (0.44-2.20), P=0.962	OR=1.09 (0.51-2.35), P=0.823
FOCUS-Ox	OR=1.13 (0.10-12.67), P=0.920	OR=3.25 (0.38-27.72), P=0.280
FOCUS-Interaction	OR=1.04 (0.08-13.07), P=0.978	OR=2.84 (0.14-57.07), P=0.496
FOxTROT	OR=2.83 (0.86-9.30), P=0.086	OR=3.31 (1.03-10.60), P=0.044*

Table 4.3 Signature 1 tested in MSS setting versus all samples. Because signature one has been generated in COIN all samples, I have tested the signature in all samples of FOCUS and FOxTROT and compared the results to the MSS setting. Adding samples to all the analysis resulted in stronger OR where oxaliplatin was added, but in 5FUFA setting, the OR was completely flat assuring that this signature is specific for oxaliplatin. Further, in FOxTROT, a neoadjuvant clinical trial, testing signature one in all samples resulted in higher OR than MSS and a significant P-value which is very promising for such a biomarker.

4.2.3 Signature two: generation in Focus

In this analysis, the FOCUS oxaliplatin arm is used (PD = 7, Non-PD = 63). The primary progressors are selected as the PD patients. This binarisation is different from the one used in chapter two and three where I have binarised as good responders to treatment (CR+PR) versus poor responders (PD+SD). Here the PD cases are grouped alone because as explained in section 1.5 of this thesis, our aim in this chapter is to find a signature predictive to oxaliplatin resistance. Thus, the rationale is to select a resistant subgroup showing a biology that could be used to predict such resistance in other cohorts. Limma at a cut off <0.1 identified 6 probesets corresponding to five genes (*MFI2, ALS2CR12, ABCB9, LOC105371371, ZC4H2*) and one not linked to any gene (Figure 4.4A). Those 6 probesets were able to separate patients between responders and non-responders in the oxaliplatin arm (Figure 4.4B). This potential signature has been tested in the control arm and did not stratify patients to 5FUFA response, suggesting that it may be specific to oxaliplatin (Figure 4.4C). Signature two has been tested in FOCUS FOLFOX arm to make sure that it is statistically associated with poor response. The results are shown in table 4.4. High signature score has resulted in OR=1.91 (1.73-2.10, P= $2x10^{-16}$). Testing the signature in the control arm was flat (OR=1.17 (0.97-1.42), P=0.09), and the interaction analysis between both arms has been significant (OR=1.63 (1.21-2.17), P=0.00147).

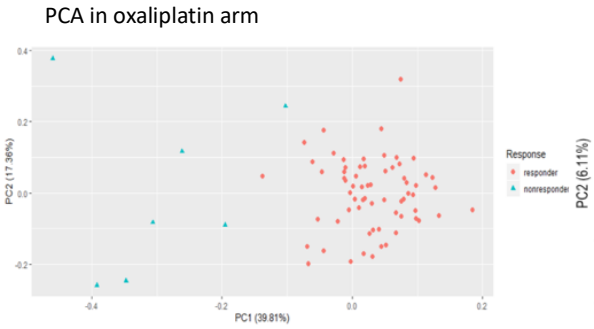
Detailed analysis of the probeset unlinked to any gene and which is 2707 bp long showed that it mapped to an intron of HIPK2 gene (Figure 4.4D). All of the individual probesets align to this consensus sequence and has no off-target alignment in the genome.

Pathway analysis, using over representation analysis with Kegg pathways was conducted to check if those genes belong to a common pathway, but it was not informative. This is probably to the low number of genes being tested. Furthermore, Gene Set Enrichment Analysis (GSEA) was performed by sorting the logFC of the genes and using gseKEGG() function for KEGG pathways as GSEA gene sets but again the analysis did not yield any informative outcome. This can be attributed again to low number of genes or low logFC in these significant genes

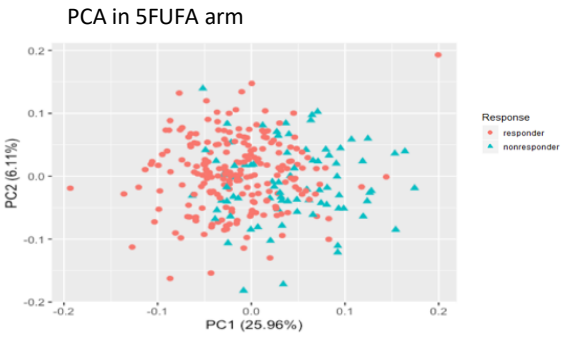
A

Probe.Set.ID	logFC	AveExpr	p.value	adjust.p.value	Gene.Symbol
ADXECEMUTR.1397_at	-1.4069669	3.82482726	3.96E-08	0.00264072	---
ADXECAD.1972_x_at	-0.941645	6.07352758	1.36E-06	0.04215125	MFI2
ADXECAD.18046_at	-0.8028881	3.91748809	1.90E-06	0.04215125	ALS2CR12
ADXEC.27397.C1_at	-1.5166774	6.34761482	4.28E-06	0.07051704	ABCB9
ADXECADA.13556_at	-0.7749504	3.68317225	5.29E-06	0.07051704	LOC105371371
ADXECAD.9184_at	-1.0676049	4.11748061	6.71E-06	0.07456251	ZC4H2

B



C



D

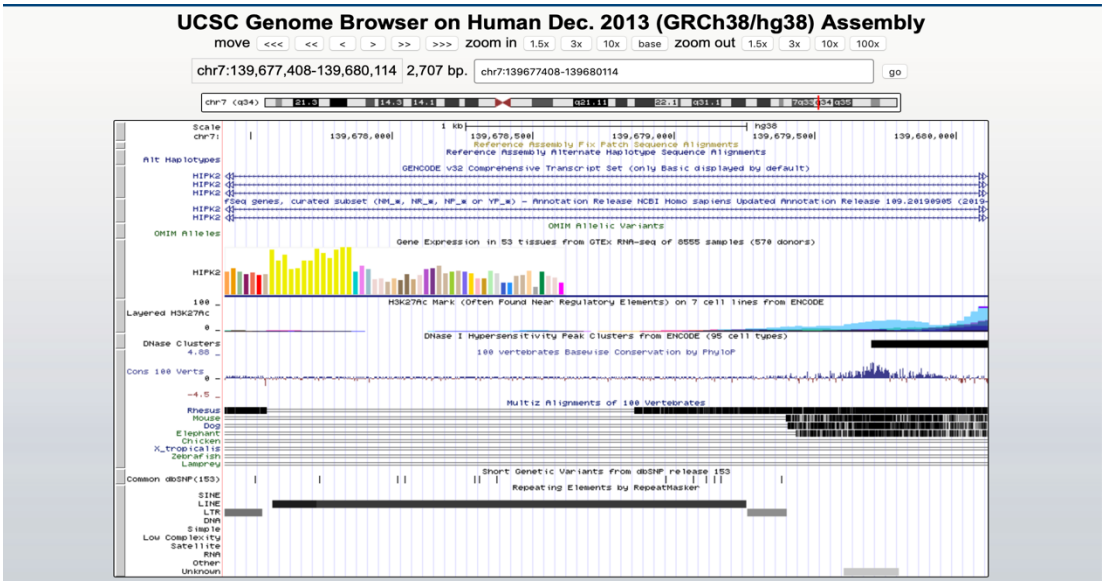


Figure 4.4 Signature 2: generation in FOCUS. (A) Limma analysis in PD versus non-PD patients generated six probesets corresponding to five genes. (B) Those six probesets were able to stratify patients in the oxaliplatin arm (C). The probesets identified did not stratify patients in the control 5FUFA arm suggesting that it could be specific to oxaliplatin.

4.2.4 Signature two: testing in COIN and FOxTROT

The list generated from limma in FOCUS all had logFC below zero thus when calculating the signature in any dataset, the subtraction step described earlier is omitted. The probesets identified have been linked to entrez IDs using biomaRT package in R.

For COIN, mean per entrez expression has been used for signature generation in the same way that signature one has been generated. Outcome analysis using logistic regression with binarised response to treatment in COIN MSS patients (PD = 29, Non-PD = 82) have not yielded significant results (OR = 1.00 (0.99-1.01), P = 0.75) (Table 4.4). Similar analysis has been conducted in FOxTROT with no significant result for validation of signature two. This signature has then been tested in Foxtrot biopsies MSS samples using the same binarised response as used to validate signature one (e.g. no regression grouped with mild regression as poor responders (n = 64); marked regression grouped with moderate regression (n = 24) as good responders). The analysis has not yielded a significant result although it shows a similar trend observed in FOCUS with high score associated with poor response (OR = 1.57 (0.27-9.07), P = 0.61) (Table 4.4).

	Cohort	OR	95% CI	p-value
Sig 2	FOCUS-Ox	1.91	1.73-2.10	2.00E-16 ***
	FOCUS-5FUFA	1.17	0.97-1.42	0.092
	FOCUS-Interaction	1.62	1.21-2.17	0.00147 **
	COIN	1.00	0.99-1.01	0.754
	FOxTROT	1.57	0.27-9.07	0.613

Table 4.4 Signature 2: testing in COIN and FOxTROT. Signature 2 did not validate in COIN MSS cases neither in FOxTROT MSS biopsies although the latter had a similar trend to what seen in FOCUS and associated with poor response yet it did not reach significance.

4.2.5 Signature 3: Generation in FOxTROT

Mindful of the low statistical power, differential gene expression analysis was run on FOxTROT as a discovery set with poor responders (n = 69) and good responders (n = 24) using limma, but the analysis yielded zero significant probesets.

4.3 Methylation signature

4.3.1 Generation in FOCUS

Using ChAMP, differential methylation analysis was conducted on the 5FUFA MSS arm and the oxaliplatin MSS arm separately. From the oxaliplatin arm (n = 62, PD = 5, Non-PD =57), 4948 DMPs corresponding to 3452 genes were identified and also 43 DMRs. From the 5FUFA arm (n = 247, PD=63, Non-PD= 184), 58 DMPs corresponding to 43 genes were identified along with 81 DMRs. The signal is unique to the oxaliplatin arm as the probes are not common with the signal identified in the control arm. This is a very interesting research pipeline to follow for future work. Further QC and machine learning methods may be used to better understand the robustness of these signals.

4.3.2 Generation in FOxTROT

Performing the same ChAMP analysis on FOxTROT MSS cases (n = 37, poor responders = 16, good responders= 21), did not yield any DMPs and only 18 DMRs. FOxTROT may be used for future methylation work. A summary of the methylation results conducted in both FOCUS and FOxTROT is shown in table 4.5.

Dataset	DMP	Genes	Common Probes/Genes between arms	DMR
FOCUS – Ox arm N= 62	4948	3452	0	43
FOCUS – 5FUFA arm N= 247	58	43	0	81
FOxTROT N= 37	0	0	0	18

Table 4.5 Differential methylation analysis in FOCUS and FOxTROT. In the FOCUS MSS oxaliplatin arm 4949 DMPs are identified corresponding to 3452 genes. In the FOCUS MSS 5FUFA arm 58 DMPs have been identified corresponding to 43 genes. The signal is unique for each arm as there are zero common genes between the two arms. FoXTROT analysis shows no DMPs due to the fact of lows numbers of samples in this dataset.

4.4 Exploratory analysis of DNA mutations and copy number alterations

Next, exploratory analysis in DNA mutations that have not been tested as prespecified biomarkers of CRC were tested in this section. Total number of genes in the targeted panel is 78. Any genes whose variability was less than 5% and not used in the previous candidate analysis (n = 62 in FOCUS) were removed due to low statistical power, leaving 16 genes in FOCUS. In the 5FUFA MSS arm (n = 270, CR+PR= 91, SD+PD = 179), logistic regression with binarised response to treatment was conducted and yielded no significant genes prognostic of 5FUFA treatment (Figure 4.5A). Similarly, in the oxaliplatin arm MSS (n = 70, CR+PR= 47, SD+PD = 23) there were no significant genes identified prognostic for response to oxaliplatin treatment (Figure 4.5B). Interaction analysis was performed to check if any genes were predictive of oxaliplatin

response (5FUFA arm n = 270, oxaliplatin arm n = 70). None of the mutated genes tested were predictive of oxaliplatin response (Figure 4.5C).

Similarly, copy number variations were explored. Due to the low resolution of the targeted NGS panel, these were run only at the chromosomal arm level. Two of the interrogated variables were found associated with response in 5FUFA arm (n = 270, CR+PR= 91, SD+PD = 179) after FDR correction: 13q (OR= 0.12, 95%CI (0.04-0.39, P=0.0004 FDR=0.015) and 18q (19.57, 95%CI (2.84-134.76), P=0.0025 FDR=0.048) (Figure 4.6A). No significant results were found in the oxaliplatin arm (n = 64, CR+PR = 43, SD+PD = 21) (Figure 4.6B). Finally, interaction analysis between the two arms had no significant results (Figure 4.6C).

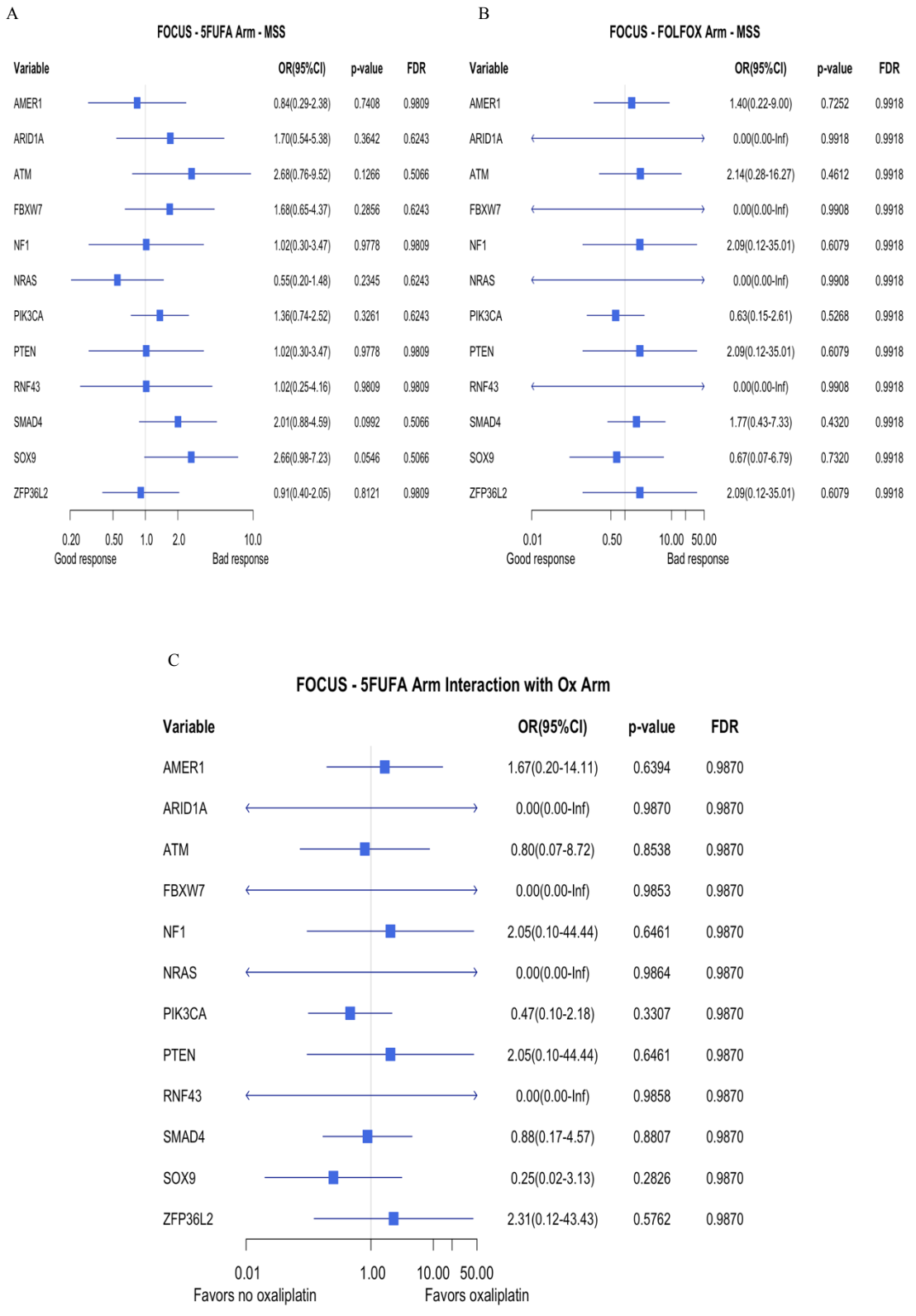


Figure 4.5 Exploratory analysis of DNA mutations in FOCUS. None of the tested mutations show significant result after correcting for false positive results by FDR in the control 5FUFA arm (A) and oxaliplatin arm (B). Prediction for oxaliplatin response tested by interaction analysis also shows negative results (C).

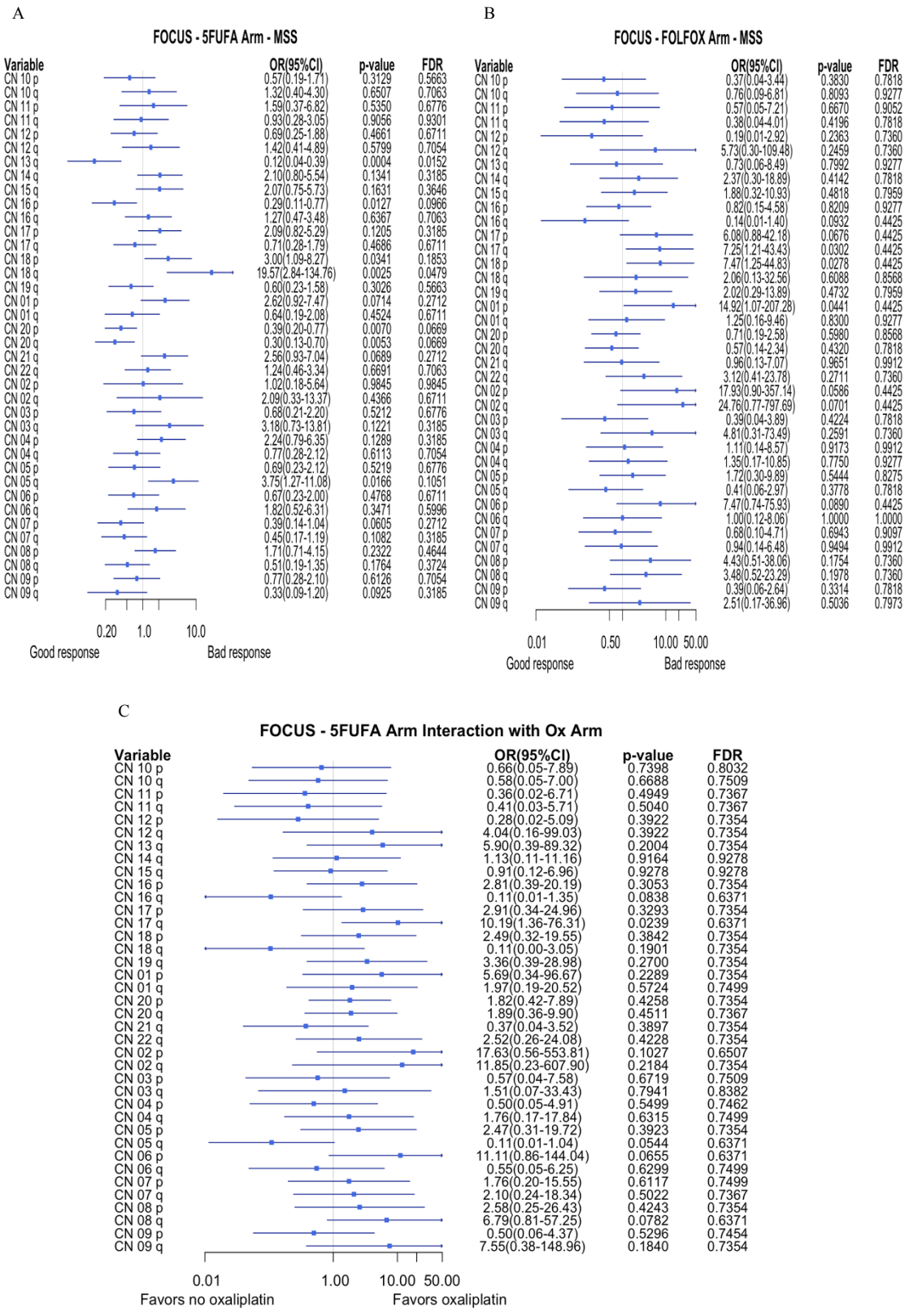


Figure 4.6 Exploratory analysis of copy number variation in FOCUS. In 5FUFA arm (A) 13q (OR= 0.12, 95%CI (0.04-0.39, P=0.0004 FDR=0.015) and 18q (19.57, 95%CI (2.84-134.76), P=0.0025 FDR=0.048) are both significant for response to 5FUFA. No significant result is seen in the oxaliplatin arm (B) neither in interaction analysis between the two arms (C).

4.5 Summary

In this chapter, I generated two signatures for oxaliplatin from differential expression analysis between poor responders and good responders to treatment. Signature one, which consists of 29 DEGs, is generated using DESeq2 in the COIN clinical trial where patients are treated with oxaliplatin based chemotherapy. Signature one high score is associated with poor response to oxaliplatin in COIN as expected. Testing this signature in FOCUS MSS cases gives a flat result in the 5FUFA, oxaliplatin arm and interaction analysis. However, testing signature one in all samples of FOCUS, showed no signal in the control (5FUFA) arm, but shows a trend in the right direction in the FOLFOX arm and interaction analysis. Failure to reach significance, despite a strong OR, is because the non progressors in the FOLFOX arm is hugely underpowered (n=7). This signature shows a trend to be associated with poor response to oxaliplatin when tested in FOxTROT MSS cases. However, when tested in all samples of FOxTROT, signature one validates and is significant.

Signature two, composed of 5 genes, is generated using limma in the FOCUS clinical trial oxaliplatin arm where patients are treated with FOLFOX. Signature two was able to stratify patients in the oxaliplatin arm but not in the 5FUFA suggesting that it may be specific to oxaliplatin. Testing signature two in COIN and FOxTROT MSS samples has not resulted in significant results. I then did differential methylation analysis using CHAMP in both the 5FUFA arm and the oxaliplatin arm of FOCUS. Both analyses have resulted in significant results that are worth following up in the future. Finally, in this chapter, I performed some exploratory analysis with DNA mutations and copy number variation in FOCUS. In 5FUFA arm 13q and 18q are both significant for response to 5FUFA. 18q has been reported in colorectal cancer but 13q is a new finding that can be explored further in future work.

Chapter 5 Discussion

Colorectal cancer is the second most common cancer in terms of mortality in the world. Adjuvant chemotherapy after surgery is the standard of care for advanced stage II, stage III and IV. The most common approved chemotherapeutic regimen is FOLFOX composed of folinic acid (leucovorin), 5-fluorouracil, and oxaliplatin (Poston, 2011). Oxaliplatin is a third-generation platinum-based compound whose most common and dominant mechanism of action is through intra-strand crosslinks that causes cell lesions and death (Faivre, 2003). It has a lot of side effects that narrow its therapeutic index. It affects the hematopoietic system causing neutropenia, the peripheral nerves causing acute or chronic neuropathy and finally the gastrointestinal tract causing nausea and diarrhoea (Koutras, 2004). Despite the improvement in outcome seen in advanced disease when treated with oxaliplatin, a high percentage of cases show lack of response and resistance while bearing all the side effects of this chemotherapeutic drug. Thus, my aim for this thesis is to discover or validate a biomarker of oxaliplatin response with potential use in the clinic to guide therapy decisions.

The multidisciplinary S:CORT consortium's primary hypothesis was that DDRD positive predicts improved survival following FOLFOX treatment, similar to what has been shown in breast cancer where DDRD positive predicts response to DNA damaging (anthracycline and cyclophosphamide) chemotherapy. Testing this hypothesis in the FOCUS clinical trial, with patients randomised to a control arm of 5FUFA treatment and an investigation arm with FOLFOX treatment, and FOxTROT a neoadjuvant trial, showed that DDRD negative patients presented a trend toward improved outcome for oxaliplatin. The results showed significance in response and a trend in PFS and OS. Moreover, it has been shown that DDRD positive is strongly associated with MSI and CMS1 (Malla, 2021). Since the original hypothesis of DDRD status being a predictor to

oxaliplatin outcome did not validate in colorectal cancer, I started my thesis searching predictors for response to oxaliplatin. Response has been selected as the primary end point because it gives the most optimal view on oxaliplatin resistance since it presents a more direct measure of the drug effect compared to survival endpoints where several factors may contribute to the results seen rather than the effect of a particular biomarker identified. For example, early events shown in PFS data may not always be directly linked to lack of effects of the treatment, and conversely late events may not always be linked to a good effect uniquely due to the treatment. In addition, in our FOCUS dataset, it is clear how the PD subset of patients are so different from other responders (Figure 1.6). They are clearly not being benefited at all by the treatment while the others show at least some benefit. Thus, this thesis focused on the progressors group of patients to find a biomarker for lack of response to oxaliplatin. Clinically, a biomarker for response is an interesting end point for neoadjuvant settings where treatment before surgery is crucial to have an effect on tumour shrinkage and maximise chances of a successful surgery. Accordingly, response is a valuable endpoint in the neoadjuvant setting.

Furthermore, this thesis has explored finding and validating a biomarker in MSS cases only in light of the changing paradigm in which MSI patients are treated increasingly with immunotherapy. MSI group which is immune rich and defined by loss of MMR machinery, shows good response to checkpoint inhibitors and immune therapy, most reasonably due to the fact that they have high mutational load and high immune infiltration. A monoclonal anti-PD1 antibody (pembrolizumab) has been approved for MSI-H in all kind of cancers, thus giving oxaliplatin to this group of patients is not the first choice that clinicians are following recently. For this reason, finding a biomarker in the MSS patients, where oxaliplatin based chemotherapy is still the treatment of choice in metastatic setting, is of high clinical significance.

I first investigated 13 multiomic candidate biomarkers known for their prominent role in CRC biology. I have found that *KRAS* mutation is associated with poor response and survival upon 5FUFA treatment in MSS metastatic colorectal cancer patients (Figure 2.4A and 2.5C). This is consistent with the literature that clearly shows KRAS mutation within MSS patients results in poor prognosis both in the adjuvant (Domingo, 2018; Taieb, 2016) and metastatic setting (Richman, 2009). However, I found KRAS mutation favours oxaliplatin as is predictive of oxaliplatin better response compared to 5FUFA alone (Figure 2.4C). Intriguingly, this association did not result in different survival patterns as already published in this dataset (Richman, 2009). However, the trends observed for response and PFS are the same. Further analyses with additional KRAS mutation data from the Focus trial might provide more insights about the strength on the link on response and PFS. In our results, *KRAS* mutants showed better response with the addition of oxaliplatin than *KRAS* wild types. Oncogenic *KRAS* induces DNA replication stress by promoting aberrations in replication fork progression, which leads to DNA damage and genomic instability. The cells respond to this signal by activating DDR pathways. Mutant *KRAS* keeps the oncogenic signal on causing higher replication stress. This results in replication fork stalling forming DNA breaks and damage. This catastrophic level of damage will consequently lead to less effective DNA damage responses, particularly NER due to high correlation between replication stress response and NER (Bélanger, 2016). When treated with oxaliplatin, cells with *KRAS* mutation will have replication stress inhibiting NER potentially leading to higher sensitivity to oxaliplatin as there would be no repair of intrastrand crosslink, the major crosslinks caused by oxaliplatin. Furthermore, besides the lesions caused by oxaliplatin, replication stress cause accumulation of DNA damage in mitotic cells causing mitotic catastrophe and cell death (Gastl, 2020).

Interestingly, I also found that DDRD positive favours no oxaliplatin and is predictive of oxaliplatin worse response (Figure 2.4C). This result is in line with the results from Malla et al. where it has been shown that DDRD negative benefits from oxaliplatin based chemotherapy. Similarly, we show that DDRD positive in MSS metastatic setting favours no oxaliplatin and predicts worse response. However, Malla et al. has shown that DDRD positive is strongly associated with MSI, while we show that DDRD positive is predictive of poor response to oxaliplatin in the absence of MSI. Thus, strong association of DDRD positivity with MSI is not confounding the association of poor response to oxaliplatin and thus those cases will not respond to oxaliplatin and might not respond to immune therapy either.

Then I have investigated 32 candidate biomarkers identified from a systematic literature review that I conducted based on oxaliplatin resistance and resistant mechanisms in the literature. One of our exclusion criteria for the literature review was to exclude literature reported in journals with below five impact factor. This criterion was put into place to avoid selecting a high number of candidate biomarkers leading to multiple testing problems and increase type I errors associated with that. There are obvious associated risks that come with this exclusion criterion be it missing out on additional candidates that could be bona fide biomarkers as the impact factor of journals do not always reflect the strength of the data reported, neither the importance of research conducted. The ideal method was to conduct a systematic literature review using PRISMA which stands for Preferred Reposting Items for Systematic Reviews and Meta Analysis. PRISMA facilitates transparent reporting of literature reviews. Its methodology consists of a 27 item checklist and a 4-phase flow diagram (Page, 2021). This approach would have

increased the objectivity and transparency of my literature review and would have given me the chance to assess manuscripts from any impact factor journals and select biomarkers based on biology along with the strength of the methods used and results reported avoiding the associated risk of missing and not testing a crucial candidate variable in oxaliplatin resistance.

Our data show that in the 5FUFA arm three biomarkers have been associated with treatment response. *BMP4* high expression is associated with poor response to 5FU. *BMP4* is a gene that encodes for bone morphogenic protein 4 and is a member of the TGF-β family. It has been reported that BMP4 is involved in cell proliferation and key player in EMT which drives chemoresistance. It has been recently shown that BMP4 is correlated with poor prognosis in hepatocellular carcinoma by inducing EMT signaling and causing oxaliplatin resistance through activation of MEK/ERK/ELK1 signaling (Ma, 2017). We did not see *BMP4* gene to have an influence on oxaliplatin response but when overexpressed, it was associated with poor response to 5FU alone. The second biomarker that we have found to be associated with poor response to 5FU is *CXCL8*. It encodes interleukin-8 and is one of the most upregulated chemokines in colorectal cancer. It has been shown that CXCL8 is upregulated in HCT116 5FU resistance cell line (Dabkeviciene, 2015). Our data validate this finding in a clinical setting where we show high expression of the gene drives poor response to 5FU. The third biomarker that correlated with response to 5FU is *AKT1* gene which translates AKT kinase which plays a role in many signaling pathways, one of which is AKT pathway. The gene high expression has resulted in poor response. It has been shown that mechanistically AKT signaling is upregulated leading to SLUG overexpression and facilitating the development of resistance to oxaliplatin in HCT116 cell lines, consequently leading to

the induction of EMT (Wei, 2020). Our data reveal that *AKT1* gene over expression leads to poor response to 5FU and not oxaliplatin as was observed in cell lines.

Next, I did interaction analysis between the 5FUFA arm and the oxaliplatin arm that would identify predictive biomarkers of response to oxaliplatin. Three biomarkers (*CDC6*, *NFE2L2*, and DNA damage repair score) have validated in the FOCUS trial to be predictive of response to oxaliplatin in MSS metastatic setting (Figure 3.4C). *CDC6* high expression favours no oxaliplatin and predicts lack of response. Mechanistically, it has been shown that *CDC6* is a key regulator of DNA replication where it is part of a complex that adds mini chromosome maintenance proteins to replication origins to initiate DNA synthesis (Whitmire, 2002). Moreover, *CDC6* acts upstream of a complex composed of p-21(CIP1) or p27 (KIP1), CDK2 and cyclin A or E to promote cell proliferation and survival (Okayama, 2012). Recently, *CDC6* has been shown to be a target of human antigen R (HuR) and acts downstream of it to promote colorectal cancer tumourigenesis and oxaliplatin resistance. Mechanistically, HuR binds to 3'UTR of CDC6 and upregulates its expression in colorectal cancer cell lines. *CDC6* overexpression in those cell lines resulted in increased colorectal cancer cell malignancy and oxaliplatin resistance, while *CDC6* knockdown resulted in decreased malignancy and reversed oxaliplatin resistance. In colorectal cancer tissues, *CDC6* levels have been shown to be increased (Cai, 2019). Here, we tested *CDC6* expression in a randomised metastatic colorectal clinical trial setting and we show that its high expression is predictive of oxaliplatin worse response than 5FUFA in MSS cases. Even upon binarisng the expression score based on a cut-off that has been generated using an independent dataset, we show that *CDC6* binarised high score is significant and predicts oxaliplatin worse response (Figure 3.7C). By this analysis, we confirmed the role *CDC6* might be playing in a clinical setting which suggests potential for targeting *CDC6* to reverse oxaliplatin

resistance in colorectal cancer. HuR role has been studied in carcinogenesis and there are available HuR inhibitors like small molecule MS-444 (Meisner, 2007) but those molecules are difficult to develop because they modulate protein-protein interactions which are hard to target. Targeting *CDC6* might be a better and more feasible option to overcome oxaliplatin resistance after we have confirmed its role in a clinical context.

The second biomarker for oxaliplatin we identified in this work is *NFE2L2*. The protein that this gene translates is Nrf2 and is overexpressed in colorectal cancer with both nuclear and cytoplasmic localisations(Lin, 2016). It has also been shown that Nrf2 overexpression leads to 5FU resistance in colorectal cancer where Nrf2 and the antioxidant response element (ARE) pathway increases antioxidant and gene expressions leading to decreased ROS- induced DNA damage and apoptosis and thus chemoresistance (Akhdar, 2009). Moreover, Nrf2 plays a particular role in oxaliplatin resistance through the activation of Keap1/Nrf2 pathway (Wang, 2014). Recently, it has been shown that cytoplasmic Nrf2 may play a more important role than nuclear Nrf2 in resistance to both 5FU and might play the same role in resistance to oxaliplatin in colon cancer cells where mechanistically cNrf2 induced PSMD4 expression to cause chemoresistance via the NF-κB/AKT/β-catenin/ZEB1 signalling cascade (Cheng, 2018). Thus, we selected *NFE2L2* gene to test its role in oxaliplatin resistance in a clinical metastatic setting. Contrary to what is expected, our data show that high expression of this gene actually favours oxaliplatin, not the opposite, and predicts better response to FOLFOX than 5FUFA alone in MSS cases (Figure 3.7C). Our group has previously tested Nrf2 role in colorectal cancer by generating an *NRF2* metagene composed of 36 Nrf2 regulated genes and has shown its prognostic value for 5FUFA treatment. High expression of this metagene resulted in worse survival in the 5FUFA arm of the FOCUS trial (O'Cathail, 2019). The metagene did not include *NRF2* gene where we show here its predictive value in

predicting better response to oxaliplatin. Thus, the addition of oxaliplatin to 5FUFA might lead to better response in MSS metastatic colorectal cancer than the treatment with 5FUFA alone with high *NRF2* expression cases.

The third biomarker identified is the DNA damage repair signature. This DNA repair score has been shown to be prognostic for OS and PFS of patients with colorectal cancer treated with FOLFOX. The score is composed of 12 genes. The signature includes bad prognostic genes (*HMGN1, RPA2, GTF2H2, SMUG1, ERCC1*) and the good prognosis genes (*SRCAP, XRCC2, DMC1, GTF2H3, BARD1, MSH3, FAAP24*) (Combes, 2019). Those genes belong to nucleotide excision repair (NER) pathway (*ERCC1*, *GTF2H2, HMGN1, RPA2,* and *GTF2H3)* and homologous recombination (HR) (*SRCAP, XRCC2, DMC1, BARD1, RPA2)* highlighting the importance of DNA damage response in oxaliplatin mechanism of action. The major DNA damage, caused by oxaliplatin 1,2 GG intra-strand crosslinks, is repaired by NER (Slyskova, 2018). ATR which is the major protein in HR pathway has been shown to have a role in oxaliplatin resistance highlighting the importance of HR in oxaliplatin resistance (Combes, 2019). Combes et al tested this DNA repair score in FOLFOX setting in two datasets (Tsuji and Del Rio) and linked it to survival to be prognostic in this setting. We, not only, have validated this result in the FOCUS metastatic clinical trial where high score of the DNA repair score is associated with poor response to oxaliplatin (Figure 3.7B), but we also demonstrate that DNA repair score is a predictive biomarker for lack of response to oxaliplatin upon interaction analysis between 5FUFA arm and FOLFOX arm of FOCUS MSS cases (Figure 3.7C). Finally, the way that we generate the score is different than the method used in the original paper which is subjective to the cohorts used in the analysis, while our method can be applied to any dataset. Our method takes the mean of expression of the positive genes, along with the mean of expression of the negative genes and each sample DNA repair

signature score is then generated by subtracting the two means. I validated this method with Tsuji and Del Rio datasets to give the exact same results with survival as the ones shown in Combes et al. Thus, our data were able to show that the DNA repair score is not only a prognostic biomarker for oxaliplatin treatment but, also, we identified it as a predictive biomarker in a randomised controlled clinical trial.

We tested *CDC6*, *NFE2L2* and the DNA signature score as continuous biomarkers and then used an independent dataset to binarise them in FOCUS and rerun outcome analysis. One limitation of using a binary biomarker is that it is less sensitive to change than the continuous one. For example, in our candidate genes and DNA repair score, any change in expression would translate as a direct effect on continuous biomarker measure, however this change in expression needs to be significant for it to have a measurable effect on the binarised biomarker. Thus, continuous biomarkers are more sensitive to change and minimal changes in expression could be important on the biomarker and its association with outcome. It has been reported that the information lost when transforming continuous data into a binary response measure translates into a loss of power to detect differences between treatments in mixed treatment comparison models (Schmitz, 2012). Furthermore, dichotomising a biomarker into patient groups assumes the relationship between the biomarker and outcome is flat for patients within the same risk group not adequately reflecting the linear relationship between the biomarker and the outcome(Polley & Dignam, 2021). There is a preference from statisticians not to categorise a continuous biomarker at all. A supported approach is to model a linear regression model between a continuous biomarker and the outcome of interest without introducing cut points. The advantage of this method is retaining valuable sensitive information in the data and the ability of reproducibility of the method and result comparison in different datasets (Polley & Dignam, 2021).

However, binary or categorical markers are more accepted in the oncology field from medical practitioners as having a clear optimal cut point(s) stratifying patients into high risk or low risk groups may have clinical advantages. This kind of categorization makes it easier to interpret the effects of a biomarker on the outcome via an odds ratio or risk ratio. It also helps clinicians by providing objective criteria in the selection of treatment options (Woo & Kim, 2020). To aid clinical decision making, medical practitioners are accustomed to discretizing a biomarker measured on a quantitative scale (percentiles of the data, for example, like selecting mean or median of the biomarker) into different risk categories based on some partition of the scale, commonly called cut points. This practice is natural, as it is desirable to define patient groups sharing a similar expected prognosis, thus categorical or binary biomarkers are more acceptable in the medical practice (Polley & Dignam, 2021). However, those cut points that are derived arbitrarily are less accurate and lead to information loss if compared with the continuous counterpart. To avoid those issues, one solution is to model the relationship between the biomarker and the outcome data via regression and then derive a decision matrix to select the cut point based on highest sensitivity and specificity (Y. Zhang & Molinaro, 2022). This was the method we used in dichotomising the biomarkers giving us stronger confidence in the observed results. Finally, in the medical literature, there is a prevalent problem of optimising a biomarker based on outcome and then deriving the cut point from the same dataset (Polley & Dignam, 2021). In our analysis, we avoided such issue by using a cut point derived from an independent publicly available dataset based on the same outcome endpoint. Nonetheless, continuous biomarkers can sometimes be more optimal choices, supported by statisticians and are slowly becoming more accepted by medical practitioners. Instead of a binary biomarker, a more supportive method is the derivation of a prognostic risk

model. Oncologists can assess efficacy versus toxicity versus response to evaluate a biomarker indicating lack of response to treatment.

Finally, a new RNA signature has been generated and validated in high quality clinical trial CRC data identifying patients with early progression on oxaliplatin plus 5FU. Signature one, composed of 29 genes, was generated in COIN clinical trial where FOLFOX is the treatment of choice. This signature has been generated in all samples and not MSS only. Upon testing it in the FOCUS trial, the control arm 5FUFA is flat whether all samples are used in the analysis or MSS only cases, while oxaliplatin arm shows the same trend as the discovery set where high signature score is associated with poor response to oxaliplatin. Interaction analysis between the two arms show the same trend that high score of this signature is predictive of worse response to oxaliplatin (Table 4.2). Validation of signature one in the FOCUS trial shows high odds ratio, but did not reach significance due to the fact of very low number of PD patients in the oxaliplatin arm (n=7) compared to the non-PD samples (n=63). We are hugely under powered and for this reason we could see the trend but the signature needs to be tested in a bigger sample cohort for validation and significance. In the FOxTROT trial which is a neoadjuvant clinical trial with FOLFOX treatment, signature one high score is associated with poor response with a result close to reaching significance in MSS samples. When taking all samples into the analysis, the signature validates with significance (Table 4.4). Despite heterogeneity in the clinical settings used (COIN and FOCUS versus FOxTORT), this signature promises to provide useful information for oxaliplatin stratification. This is particularly promising and interesting as such a signature for oxaliplatin lack of response is important in neoadjuvant settings where response to treatment is an important endpoint pre-surgery. We generated a second signature of resistance, composed of 5 genes, from the FOCUS trial oxaliplatin arm with PD versus non-PD samples. The signature did not

validate in COIN or FOxTROT which is expected because this analysis is exploratory one with limma cutoff set to 0.1 and thus the signature is not expected to be replicated in other cohorts as it was weakly significant in the discovery set. Moreover, there is a huge difference between FOCUS and FOxTROT as datasets. FOCUS accrued patients with metastatic disease, with both synchronous and metachronous cases, and the other is early stage. FOCUS is an advanced disease setting, while FOxTROT is neoadjuvant. The sample type differs, FOCUS is mostly primary resections while FOxTROT is biopsies. Response has been assessed with very different methods: CT scan over time in FOCUS, compared to pathological response in the resected specimen in FOxTROT. Even the response categories differ and their frequencies are not comparable between the two trials. Response is measured in metastases in FOCUS while in the primary tissue in Foxtrot. Thus, signature two not validating in Foxtrot is furthermore not surprising. Finally, one of the genes has been identified in the systematic literature review and has come up in the differential gene expression analysis in the COIN trial and is one of the genes that comprised signature one. *MUC5AC,* a gene translating a secretary mucin, when the protein is over-expressed in patients and cell lines results in higher malignancy and resistance to 5FU and oxaliplatin. In normal physiological conditions, *MUC5AC* is not expressed in colon mucosa, whereas in colon cancer and its precursor lesions, *MUC5AC* expression is observed (Pothuraju, 2020) . Because this gene has come up twice in two independent methods and analyses for oxaliplatin resistance, it is worth investigating it further and assessing its role in colorectal cancer and particularly oxaliplatin resistance.

In chapter two and three of this thesis, binarisation of the response categories of the FOCUS trial has been to group SD with PD patients. The reason for this binarisation was to gain statistical power and be able to conduct the analysis reported. The number of patients in the control is 270 in total divided into CR =1, PR=90, SD=109, PD =70.

In the FOLFOX arm which is our investigation arm of interest, the total number of patients is 70 divided into CR=1, PR=46, SD=16, PD=7. The group of interest in this analysis are the PD patients: 7 in the FOLFOX arm is hugely underpowered and for that reason we decided to binarise the PD patients with the SD ones to increase the number of poor responders to 23 in comparison to 7. This gave us statistical power to conduct the analysis. Moreover, we can be less stringent in the methods used on the basis that we were conducting a hypothesis driven analysis rather than testing in a hypothesis free manner as the case in chapter 4. In the latter, we were more stringent with patients group selection, despite losing statistical power, we still conducted the validation testing of signature 1 (derived from COIN) in FOCUS FOLFOX arm of 7 PD patients. This validation showed the same trend observed in COIN but did not reach significance due to the low number of patients in the PD group and loss of statistical power.

Other future steps for this study are to test all our predictive biomarkers for response identified from our candidate analysis (*KRAS* mutation, DDRD positive, *CDC6*, *NFE2L2*, and DNA repair score) in a bigger cohort to confirm our results. Also, test the candidates in FOxTROT due to the importance of those biomarkers of response in neoadjuvant setting. Similarly, test signature one in a bigger cohort with higher numbers of PD in the FOLFOX setting, as a major limitation of this study is low numbers of PD samples in the oxaliplatin arm, thus there is a necessity to check if this signature will gain significance with poor response to oxaliplatin similarly to what we observed here in FOxTROT. Another limitation in our study is using primary resections while having data from metastatic disease. This is particularly not ideal for RNA expression as the different tumour microenvironment may result in different transcriptomic patterns, but has little effect on mutation data which is mostly consistent across match paired primaries and metastases. However, our results are of high importance since they are conducted in a

randomised controlled clinical trial which is ideal to identify a predictive biomarker of response due to the possibility of interaction analysis conducted between the control arm which is void of the treatment of interest (in this case oxaliplatin) and the investigation arm.

Bibliography

Adam, R., Avisar, E., Ariche, A., Giachetti, S., Azoulay, D., Castaing, D., Kunstlinger, F., Levi, F., & Bismuth, F. (2001). Five-year survival following hepatic resection after neoadjuvant therapy for nonresectable colorectal [liver] metastases. *Annals of Surgical Oncology, 8,* 347-353. https://doi.org/10.1007/s10434-001-0347-3

Adams, R. A., Meade, A. M., Seymour, M. T., Wilson, R. H., Madi, A., Fisher, D., Kenny, S. L., Kay, E., Hodgkinson, E., Pope, M., Rogers, P., Wasan, H., Falk, S., Gollins, S., Hickish, T., Bessell, E. M., Propper, D., Kennedy, M. J., Kaplan, R., & Maughan, T. S. (2011). Intermittent versus continuous oxaliplatin and fluoropyrimidine combination chemotherapy for first-line treatment of advanced colorectal cancer: results of the randomised phase 3 MRC COIN trial. *The Lancet Oncology*, *12*(7), 642–653. https://doi.org/10.1016/S1470-2045(11)70102-4

Ahmad, S. (2010). Platinum-DNA interactions and subsequent cellular processes controlling sensitivity to anticancer platinum complexes. *Chemistry & Biodiversity*, *7*(3), 543–566. https://doi.org/10.1002/CBDV.200800340

Akhdar, H., Loyer, P., Rauch, C., Corlu, A., Guillouzo, A., & Morel, F. (2009). Involvement of Nrf2 activation in resistance to 5-fluorouracil in human colon cancer HT-29 cells. *European Journal of Cancer (Oxford, England : 1990)*, *45*(12), 2219–2227. https://doi.org/10.1016/J.EJCA.2009.05.017

Alcindor, T., & Beauger, N. (2011). Oxaliplatin: a review in the era of molecularly targeted therapy. *Current Oncology*, *18*(1), 18. https://doi.org/10.3747/CO.V18I1.708

Alderdice, M., Richman, S. D., Gollins, S., Stewart, J. P., Hurt, C., Adams, R., McCorry, A. M. B., Roddy, A. C., Vimalachandran, D., Isella, C., Medico, E., Maughan, T., McArt, D. G., Lawler, M., & Dunne, P. D. (2018). Prospective patient stratification into robust cancer-cell intrinsic subtypes from colorectal cancer biopsies. *Journal of Pathology*, *245*(1), 19–28. https://doi.org/10.1002/path.5051

AlDubayan, S. H., Giannakis, M., Moore, N. D., Han, G. C., Reardon, B., Hamada, T., Mu, X. J., Nishihara, R., Qian, Z., Liu, L., Yurgelun, M. B., Syngal, S., Garraway, L. A., Ogino, S., Fuchs, C. S., & van Allen, E. M. (2018). Inherited DNA-Repair Defects in Colorectal Cancer. *American Journal of Human Genetics*, *102*(3), 401–414. https://doi.org/10.1016/J.AJHG.2018.01.018

Ali, R., Al-Kawaz, A., Toss, M. S., Green, A. R., Miligy, I. M., Mesquita, K. A., Seedhouse, C., Mirza, S., Band, V., Rakha, E. A., & Madhusudan, S. (2018). Targeting PARP1 in XRCC1-deficient sporadic invasive breast cancer or preinvasive ductal carcinoma in situ induces synthetic lethality and chemoprevention. *Cancer Research*, *78*(24), 6818–6827. https://doi.org/10.1158/0008-5472.CAN-18-0633

Allen, W. L., Dunne, P. D., McDade, S., Scanlon, E., Loughrey, M., Coleman, H. G., McCann, C., McLaughlin, K., Nemeth, Z., Syed, N. A., Jithesh, P. V., Arthur, K., Wilson, R., Coyle, V. M., McArt, D., Murray, G. I., Samuel, L., Nuciforo, P., Jimenez, J., … Longley, D. B. (2018). Transcriptional subtyping and CD8 immunohistochemistry identify poor prognosis stage II/III colorectal cancer patients who benefit from adjuvant chemotherapy. *Journal of Clinical Oncology Precision Oncology*, *2018*(2), 1–15. https://doi.org/10.1200/PO.17.00241

Arango, D., Wilson, A. J., Shi, Q., Corner, G. A., Arañes, M. J., Nicholas, C., Lesser, M., Mariadason, J. M., & Augenlicht, L. H. (2004). Molecular mechanisms of action and prediction of response to oxaliplatin in colorectal cancer cells. *British Journal of Cancer*, *91*(11), 1931–1946. https://doi.org/10.1038/SJ.BJC.6602215

Armaghany, T., Wilson, J. D., Chu, Q., & Mills, G. (2012). Genetic Alterations in Colorectal Cancer. *Gastrointestinal Cancer Research: GCR*, *5*(1), 19. https://doi.org/10.3390/cancers13092025

Auclin, E., Zaanan, A., Vernerey, D., Douard, R., Gallois, C., Laurent-Puig, P., Bonnetain, F., & Taieb, J. (2017). Subgroups and prognostication in stage III colon cancer: future perspectives for adjuvant therapy. *Annals of Oncology*, *28*(5), 958–968. https://doi.org/10.1093/ANNONC/MDX030

Auvinen, A., Isola, J., Visakorpi, T., Koivula, T., Virtanen, S., & Hakama, M. (1994). Overexpression of p53 and long-term survival in colon carcinoma. *British Journal of Cancer*, *70*(2), 293–296. https://doi.org/10.1038/bjc.1994.295

Bakkenist, C. J., Lee, J. J., & Schmitz, J. C. (2018). ATM is required for the repair of oxaliplatin-induced DNA damage in colorectal cancer. *Clinical Colorectal Cancer*, *17*(4), 255. https://doi.org/10.1016/J.CLCC.2018.09.001

Bécouarn, Y., Ychou, M., Ducreux, M., Borel, C., Bertheault-Cvitkovic, F., Seitz, J. F., Nasca, S., Nguyen, T. D., Paillot, B., Raoul, J. L., Duffour, J., Fandi, A., Dupont-André, G., & Rougier, P. (1998). Phase II trial of oxaliplatin as first-line chemotherapy in metastatic colorectal cancer patients. *Journal of Clinical Oncology*, *16*(8), 2739–2744. https://doi.org/10.1200/JCO.1998.16.8.2739

Bélanger, F., Angers, J. P., Fortier, É., Hammond-Martel, I., Costantino, S., Drobetsky, E., & Wurtele, H. (2016). Mutations in replicative stress response pathways are associated with S phase-specific defects in nucleotide excision repair. *The Journal of Biological Chemistry*, *291*(2), 522. https://doi.org/10.1074/JBC.M115.685883

Benson, A. B., Schrag, D., Somerfield, M. R., Cohen, A. M., Figueredo, A. T., Flynn, P. J., Krzyzanowska, M. K., Maroun, J., McAllister, P., Van Cutsem, E., Brouwers, M., Charette, M., & Haller, D. G. (2004). American society of clinical oncology recommendations on adjuvant chemotherapy for stage ii colon cancer. *Journal of Clinical Oncology*, *22*(16), 3408–3419. https://doi.org/10.1200/JCO.2004.05.063

Berg, K. C. G., Sveen, A., Høland, M., Alagaratnam, S., Berg, M., Danielsen, S. A., Nesbakken, A., Søreide, K., & Lothe, R. A. (2019). Gene expression profiles of CMS2-epithelial/canonical colorectal cancers are largely driven by DNA copy

number gains. *Oncogene*, *38*(33), 6109–6122. https://doi.org/10.1038/S41388-019-0868-5

Bhandari, V., Li, C. H., Bristow, R. G., & Boutros, P. C. (2020). Divergent mutational processes distinguish hypoxic and normoxic tumours. *Nature Communications*, *11*(1), 1–10. https://doi.org/10.1038/s41467-019-14052-x

Bilgin, B., Sendur, M. A. N., Bülent Akıncı, M., Şener Dede, D., & Yalçın, B. (2017). Targeting the PD-1 pathway: a new hope for gastrointestinal cancers. *Current Medical Research and Opinion, 33*(4), 749–759. https://doi.org/10.1080/03007995.2017.1279132

Biswas, R., Bugde, P., He, J., Merien, F., Lu, J., Liu, D. X., Myint, K., Liu, J., McKeage, M., & Li, Y. (2019). Transport-mediated oxaliplatin resistance associated with endogenous overexpression of MRP2 in Caco-2 and PANC-1 cells. *Cancers*, *11*(9). https://doi.org/10.3390/cancers11091330

Blanco, F. F., Jimbo, M., Wulfkuhle, J., Gallagher, I., Deng, J., Enyenihi, L., Meisner-Kober, N., Londin, E., Rigoutsos, I., Sawicki, J. A., Risbud, M. V., Witkiewicz, A. K., McCue, P. A., Jiang, W., Rui, H., Yeo, C. J., Petricoin, E., Winter, J. M., & Brody, J. R. (2016). The mRNA-binding protein HuR promotes hypoxia-induced chemoresistance through posttranscriptional regulation of the proto-oncogene PIM1 in pancreatic cancer cells. *Oncogene*, *35*(19), 2529–2541. https://doi.org/10.1038/onc.2015.325

Bokemeyer, C., Cutsem, E. van, Rougier, P., Ciardiello, F., Heeger, S., Schlichting, M., Celik, I., & Köhne, C. H. (2012). Addition of cetuximab to chemotherapy as first-line treatment for KRAS wild-type metastatic colorectal cancer: pooled analysis of the CRYSTAL and OPUS randomised clinical trials. *European Journal of Cancer*, *48*(10), 1466–1475. https://doi.org/10.1016/J.EJCA.2012.02.057

Braun, M. S., Adab, F., Bradley, C., McAdam, K., Thomas, G., Wadd, N. J., Rea, D., Philips, R., Twelves, C., Bozzino, J., MacMillan, C., Saunders, M. P., Counsell, R., Anderson, H., McDonald, A., Stewart, J., Robinson, A., Davies, S., Richards, F. J., & Seymour, M. T. (2003). Modified de Gramont with oxaliplatin in the first-line treatment of advanced colorectal cancer. *British Journal of Cancer*, *89*(7), 1155–1158. https://doi.org/10.1038/sj.bjc.6601237

Bruno, P. M., Liu, Y., Park, G. Y., Murai, J., Koch, C. E., Eisen, T. J., Pritchard, J. R., Pommier, Y., Lippard, S. J., & Hemann, M. T. (2017). A subset of platinum-containing chemotherapeutic agents kills cells by inducing ribosome biogenesis stress. *Nature Medicine*, *23*(4), 461–471. https://doi.org/10.1038/nm.4291

Budzowska, M., & Kanaar, R. (2009). Mechanisms of dealing with DNA damage-induced replication problems. *Cell Biochemistry and Biophysics*, *53*(1), 17–31. https://doi.org/10.1007/S12013-008-9039-Y

Buffa, F. M., Harris, A. L., West, C. M., & Miller, C. J. (2010). Large meta-analysis of multiple cancers reveals a common, compact and highly prognostic hypoxia

metagene. *British Journal of Cancer, 102*(2), 428–435. https://doi.org/10.1038/sj.bjc.6605450

Burger, H., Loos, W. J., Eechoute, K., Verweij, J., Mathijssen, R. H. J., & Wiemer, E. A. C. (2011). Drug transporters of platinum-based anticancer agents and their clinical significance. *Drug Resistance Updates: Reviews and Commentaries in Antimicrobial and Anticancer Chemotherapy*, *14*(1), 22–34. https://doi.org/10.1016/J.DRUP.2010.12.002

Burrell, R. A., McClelland, S. E., Endesfelder, D., Groth, P., Weller, M. C., Shaikh, N., Domingo, E., Kanu, N., Dewhurst, S. M., Gronroos, E., Chew, S. K., Rowan, A. J., Schenk, A., Sheffer, M., Howell, M., Kschischo, M., Behrens, A., Helleday, T., Bartek, J., … Swanton, C. (2013). Replication stress links structural and numerical cancer chromosomal instability. *Nature, 494*(7438), 492–496. https://doi.org/10.1038/nature11935

Cai, J., Wang, H., Jiao, X., Huang, R., Qin, Q., Zhang, J., Chen, H., Feng, D., Tian, X., & Wang, H. (2019). The RNA-binding protein HuR confers oxaliplatin resistance of colorectal cancer by upregulating CdC6. *Molecular Cancer Therapeutics*, *18*(7), 1243–1254. https://doi.org/10.1158/1535-7163.MCT-18-0945

Calon, A., Espinet, E., Palomo-Ponce, S., Tauriello, D. V. F., Iglesias, M., Céspedes, M. V., Sevillano, M., Nadal, C., Jung, P., Zhang, X. H. F., Byrom, D., Riera, A., Rossell, D., Mangues, R., Massagué, J., Sancho, E., & Batlle, E. (2012). Dependency of colorectal cancer on a TGF-β-driven program in stromal cells for metastasis initiation. *Cancer Cell*, *22*(5), 571–584. https://doi.org/10.1016/J.CCR.2012.08.013

Castillo-Lizardo, M., Henneke, G., & Viguera, E. (2014). Replication slippage of the thermophilic DNA polymerases B and D from the Euryarchaeota Pyrococcus abyssi. *Frontiers in Microbiology*, *5*, 403. https://doi.org/10.3389/FMICB.2014.00403

Causse, S. Z., Marcion, G., Chanteloup, G., Uyanik, B., Boudesco, C., Grigorash, B. B., Douhard, R., Dias, A. M. M., Dumetier, B., Dondaine, L., Gozzi, G. J., Moussay, E., Paggetti, J., Mirjolet, C., de Thonel, A., Dubrez, L., Demidov, O. N., Gobbo, J., & Garrido, C. (2019). HSP110 translocates to the nucleus upon genotoxic chemotherapy and promotes DNA repair in colorectal cancer cells. *Oncogene*, *38*(15), 2767–2777. https://doi.org/10.1038/s41388-018-0616-2

Ceccaldi, R., Sarangi, P., & D'Andrea, A. D. (2016). The Fanconi anaemia pathway: new players and new functions. *Nature Reviews Molecular Cell Biology*, *17*(6), 337–349. https://doi.org/10.1038/NRM.2016.48

Chang, K., Willis, J. A., Reumers, J., Taggart, M. W., San Lucas, F. A., Thirumurthi, S., Kanth, P., Delker, D. A., Hagedorn, C. H., Lynch, P. M., Ellis, L. M., Hawk, E. T., Scheet, P. A., Kopetz, S., Arts, J., Guinney, J., Dienstmann, R., & Vilar, E. (2018). Colorectal premalignancy is associated with consensus molecular subtypes 1 and 2. *Annals of Oncology*, *29*(10), 2061–2067. https://doi.org/10.1093/ANNONC/MDY337

Chatterjee, N., & Walker, G. C. (2017). Mechanisms of DNA damage, repair and mutagenesis. *Environmental and Molecular Mutagenesis*, *58*(5), 235. https://doi.org/10.1002/EM.22087

Cheeseman, S. L., Joel, S. P., Chester, J. D., Wilson, G., Dent, J. T., Richards, F. J., & Seymour, M. T. (2002). A 'modified de Gramont' regimen of fluorouracil, alone and with oxaliplatin, for advanced colorectal cancer. *British Journal of Cancer*, *87*(4), 393–399. https://doi.org/10.1038/sj.bjc.6600467

Chen, Z. H., Qi, J. J., Wu, Q. N., Lu, J. H., Liu, Z. X., Wang, Y., Hu, P. S., Li, T., Lin, J. F., Wu, X. Y., Miao, L., Zeng, Z. L., Xie, D., Ju, H. Q., Xu, R. H., & Wang, F. (2019). Eukaryotic initiation factor 4A2 promotes experimental metastasis and oxaliplatin resistance in colorectal cancer. *Journal of Experimental and Clinical Cancer Research*, *38*(1). https://doi.org/10.1186/s13046-019-1178-z

Cheng, Y. M., Lin, P. L., Wu, D. W., Wang, L., Huang, C. C., & Lee, H. (2018). PSMD4 is a novel therapeutic target in chemoresistant colorectal cancer activated by cytoplasmic localization of Nrf2. *Oncotarget*, *9*(41), 26342–26352. https://doi.org/10.18632/oncotarget.25254

Ciccia, A., & Elledge, S. J. (2010). The DNA damage response: making it safe to play with knives. *Molecular Cell*, *40*(2), 179–204. https://doi.org/10.1016/J.MOLCEL.2010.09.019

Clancy, C., Burke, J. P., Kalady, M. F., & Coffey, J. C. (2013). BRAF mutation is associated with distinct clinicopathological characteristics in colorectal cancer: a systematic review and meta-analysis. *Colorectal Disease: The Official Journal of the Association of Coloproctology of Great Britain and Ireland*, *15*(12). https://doi.org/10.1111/CODI.12427

Cohen, R., Taieb, J., Fiskum, J., Yothers, G., Goldberg, R., Yoshino, T., Alberts, S., Allegra, C., de Gramont, A., Seitz, J. F., O'Connell, M., Haller, D., Wolmark, N., Erlichman, C., Zaniboni, A., Lonardi, S., Kerr, R., Grothey, A., Sinicrope, F. A., … Shi, Q. (2021). Microsatellite instability in patients with stage III colon cancer receiving fluoropyrimidine with or without oxaliplatin: an accent pooled analysis of 12 adjuvant trials. *Journal of Clinical Oncology*, *39*(6), 642–651. https://doi.org/10.1200/JCO.20.01600

Combes, E., Andrade, A. F., Tosi, D., Michaud, H. A., Coquel, F., Garambois, V., Desigaud, D., Jarlier, M., Coquelle, A., Pasero, P., Bonnefoy, N., Moreaux, J., Martineau, P., Rio, M. del, Beijersbergen, R. L., Vezzio-Vie, N., & Gongora, C. (2019). Inhibition of ataxia-telangiectasia mutated and Rad3-related (ATR) overcomes oxaliplatin resistance and promotes antitumor immunity in colorectal cancer. *Cancer Research*, *79*(11), 2933–2946. https://doi.org/10.1158/0008-5472.CAN-18-2807

Conti, J., & Thomas, G. (2011). The role of tumour stroma in colorectal cancer invasion and metastasis. *Cancers*, *3*(2), 2160. https://doi.org/10.3390/CANCERS3022160

Cotte, A. K., Aires, V., Fredon, M., Limagne, E., Derangère, V., Thibaudin, M., Humblin, E., Scagliarini, A., De Barros, J. P. P., Hillon, P., Ghiringhelli, F., & Delmas, D. (2018). Lysophosphatidylcholine acyltransferase 2-mediated lipid droplet production supports colorectal cancer chemoresistance. *Nature Communications*, *9*(1). https://doi.org/10.1038/s41467-017-02732-5

Craig, S. G., Humphries, M. P., Alderdice, M., Bingham, V., Richman, S. D., Loughrey, M. B., Coleman, H. G., Viratham-Pulsawatdi, A., McCombe, K., Murray, G. I., Blake, A., Domingo, E., Robineau, J., Brown, L., Fisher, D., Seymour, M. T., Quirke, P., Bankhead, P., McQuaid, S., … Salto-Tellez, M. (2020). Immune status is prognostic for poor survival in colorectal cancer patients and is associated with tumour hypoxia. *British Journal of Cancer*, *123*(8), 1280. https://doi.org/10.1038/S41416-020-0985-5

Cui, G., Yuan, A., Vonen, B., & Florholmen, J. (2009). Progressive cellular response in the lamina propria of the colorectal adenoma-carcinoma sequence. *Histopathology*, *54*(5), 550–560. https://doi.org/10.1111/J.1365-2559.2009.03273.X

Cunningham D, Zalcberg J, Maroun J, R, J., S, C., TS, M., M, V., J, S., M, G. B., & T, F. (2002). Efficacy, tolerability and management of raltitrexed (Tomudex) monotherapy in patients with advanced colorectal cancer. a review of phase II/III trials. *European Journal of Cancer*, *38*(4), 478–486. https://doi.org/10.1016/S0959-8049(01)00413-0

Dabkeviciene, D., Jonusiene, V., Zitkute, V., Zalyte, E., Grigaitis, P., Kirveliene, V., & Sasnauskiene, A. (2015). The role of interleukin-8 (CXCL8) and CXCR2 in acquired chemoresistance of human colorectal carcinoma cells HCT116. *Medical Oncology*, *32*(12). https://doi.org/10.1007/s12032-015-0703-y

Daniel G. Haller, M. (2001). Update on chemotherapy for advanced colorectal cancer. *Oncology, 15*(3), 11-15. https://doi.org/11301834

de Gramont, A., Figer, A., Seymour, M., Homerin, M., Hmissi, A., Cassidy, J., Boni, C., Cortes-Funes, H., Cervantes, A., Freyer, G., Papamichael, D., le Bail, N., Louvet, C., Hendler, D., de Braud, F., Wilson, C., Morvan, F., & Bonetti, A. (2000). Leucovorin and fluorouracil with or without oxaliplatin as first-line treatment in advanced colorectal cancer. *Journal of Clinical Oncology*, *18*(16), 2938–2947. https://doi.org/10.1200/JCO.2000.18.16.2938

De Gramont, A., Vignoud, J., Tournigand, C., Louvet, C., André, T., Varette, C., Raymond, E., Moreau, S., Le Bail, N., & Krulik, M. (1997). Oxaliplatin with high-dose leucovorin and 5-fluorouracil 48-hour continuous infusion in pretreated metastatic colorectal cancer. *European Journal of Cancer Part A*, *33*(2), 214–219. https://doi.org/10.1016/S0959-8049(96)00370-X

de Porras, V. R., Bystrup, S., Las Heras, S. C. De, Musulén, E., Palomero, L., Alonso, M. H., Nieto, R., Arango, D., Moreno, V., Queralt, C., Manzano, J. L., Layos, L., Bugés, C., & Martinez-Balibrea, E. (2019). Tumor expression of cyclin-dependent kinase 5 (Cdk5) is a prognostic biomarker and predicts outcome of oxaliplatin-

treated metastatic colorectal cancer patients. *Cancers*, *11*(10). https://doi.org/10.3390/cancers11101540

de Roock, W., Claes, B., Bernasconi, D., de Schutter, J., Biesmans, B., Fountzilas, G., Kalogeras, K. T., Kotoula, V., Papamichael, D., Laurent-Puig, P., Penault-Llorca, F., Rougier, P., Vincenzi, B., Santini, D., Tonini, G., Cappuzzo, F., Frattini, M., Molinari, F., Saletti, P., … Tejpar, S. (2010). Effects of KRAS, BRAF, NRAS, and PIK3CA mutations on the efficacy of cetuximab plus chemotherapy in chemotherapy-refractory metastatic colorectal cancer: a retrospective consortium analysis. *The Lancet. Oncology*, *11*(8), 753–762. https://doi.org/10.1016/S1470-2045(10)70130-3

Deuis, J. R., Zimmermann, K., Romanovsky, A. A., Possani, L. D., Cabot, P. J., Lewis, R. J., & Vetter, I. (2013). An animal model of oxaliplatin-induced cold allodynia reveals a crucial role for Nav1.6 in peripheral pain pathways. *Pain*, *154*(9), 1749. https://doi.org/10.1016/J.PAIN.2013.05.032

Di Nicolantonio, F., Martini, M., Molinari, F., Sartore-Bianchi, A., Arena, S., Saletti, P., de Dosso, S., Mazzucchelli, L., Frattini, M., Siena, S., & Bardelli, A. (2008). Wild-type BRAF is required for response to panitumumab or cetuximab in metastatic colorectal cancer. *Journal of Clinical Oncology*, *26*(35), 5705–5712. https://doi.org/10.1200/JCO.2008.18.0786

Dilruba, S., & Kalayda, G. v. (2016). Platinum-based drugs: past, present and future. *Cancer Chemotherapy and Pharmacology 2016 77:6*, *77*(6), 1103–1124. https://doi.org/10.1007/S00280-016-2976-Z

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., & Gingeras, T. R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*, *29*(1), 15. https://doi.org/10.1093/BIOINFORMATICS/BTS635

Domingo, E., Camps, C., Kaisaki, P. J., Parsons, M. J., Mouradov, D., Pentony, M. M., Makino, S., Palmieri, M., Ward, R. L., Hawkins, N. J., Gibbs, P., Askautrud, H., Oukrif, D., Wang, H., Wood, J., Tomlinson, E., Bark, Y., Kaur, K., Johnstone, E. C., … Tomlinson, I. (2018). Mutation burden and other molecular markers of prognosis in colorectal cancer treated with curative intent: results from the QUASAR 2 clinical trial and an Australian community-based series. *The Lancet. Gastroenterology & Hepatology*, *3*(9), 635. https://doi.org/10.1016/S2468-1253(18)30117-1

Domingo, E., Ramamoorthy, R., Oukrif, D., Rosmarin, D., Presz, M., Wang, H., Pulker, H., Lockstone, H., Hveem, T., Cranston, T., Danielsen, H., Novelli, M., Davidson, B., Xu, Z. Z., Molloy, P., Johnstone, E., Holmes, C., Midgley, R., Kerr, D., … Tomlinson, I. (2013). Use of multivariate analysis to suggest a new molecular classification of colorectal cancer. *The Journal of Pathology*, *229*(3), 441–448. https://doi.org/10.1002/PATH.4139

Dong, Y., Wang, Z., Xie, G. feng, Li, C., Zuo, W. wei, Meng, G., Xu, C. ping, & Li, J. jun. (2017). Pregnane X receptor is associated with unfavorable survival and

induces chemotherapeutic resistance by transcriptional activating multidrug resistance-related protein 3 in colorectal cancer. *Molecular Cancer*, *16*(1). https://doi.org/10.1186/s12943-017-0641-8

Douillard, J.-Y., Oliner, K. S., Siena, S., Tabernero, J., Burkes, R., Barugel, M., Humblet, Y., Bodoky, G., Cunningham, D., Jassem, J., Rivera, F., Kocákova, I., Ruff, P., Błasińska-Morawiec, M., Šmakal, M., Canon, J. L., Rother, M., Williams, R., Rong, A., … Patterson, S. D. (2013). Panitumumab–FOLFOX4 Treatment and RAS Mutations in Colorectal Cancer. *New England Journal of Medicine*, *369*(11), 1023–1034. https://doi.org/10.1056/NEJMOA1305275

Downward, J. (2003). Targeting RAS signalling pathways in cancer therapy. *Nature Reviews Cancer 2003 3:1*, *3*(1), 11–22. https://doi.org/10.1038/NRC969

Dunne, P. D., Alderdice, M., O'Reilly, P. G., Roddy, A. C., McCorry, A. M. B., Richman, S., Maughan, T., McDade, S. S., Johnston, P. G., Longley, D. B., Kay, E., McArt, D. G., & Lawler, M. (2017). Cancer-cell intrinsic gene expression signatures overcome intratumoural heterogeneity bias in colorectal cancer patient classification. *Nature Communications*, *8*, 15657. https://doi.org/10.1038/NCOMMS15657

Dutton, L. R., Hoare, O. P., McCorry, A. M. B., Redmond, K. L., Adam, N. E., Canamara, S., Bingham, V., Mullan, P. B., Lawler, M., Dunne, P. D., & Brazil, D. P. (2019). Fibroblast-derived Gremlin1 localises to epithelial cells at the base of the intestinal crypt. *Oncotarget*, *10*(45), 4630–4639. https://doi.org/10.18632/ONCOTARGET.27050

Edge, S. B., & Compton, C. C. (2010). The American Joint Committee on Cancer: the 7th edition of the AJCC cancer staging manual and the future of TNM. *Annals of Surgical Oncology*, *17*(6), 1471–1474. https://doi.org/10.1245/S10434-010-0985-4

Eisenhauer, E. A., Therasse, P., Bogaerts, J., Schwartz, L. H., Sargent, D., Ford, R., Dancey, J., Arbuck, S., Gwyther, S., Mooney, M., Rubinstein, L., Shankar, L., Dodd, L., Kaplan, R., Lacombe, D., & Verweij, J. (2009). New response evaluation criteria in solid tumours: Revised RECIST guideline (version 1.1). *European Journal of Cancer*, *45*(2), 228–247. https://doi.org/10.1016/j.ejca.2008.10.026

Elsaleh, H., Powell, B., Mccaul, K., Grieu, F., Grant, R., Joseph, D., & Iacopetta, B. (2001). P53 alteration and microsatellite instability have predictive value for survival benefit from chemotherapy in stage III colorectal carcinoma 1. *Clinical Cancer Research, 7*(5), 1343-1349. https://doi.org/ 1557-3265

Esteller, M. (2007). Cancer epigenomics: DNA methylomes and histone-modification maps. *Nature Reviews Genetics*, *8*(4), 286–298. https://doi.org/10.1038/NRG2005

DeBraud, F., Munzone, E., Nolè, F., De Pas, T., Biffi, R., Brienza, S., Aapro, MS. (1998). Synergistic activity of oxaliplatin and 5-fluorouracil in patients with metastatic colorectal cancer with progressive disease while on or after 5-

fluorouracil. *American Journal of Clinical Oncology*, *21*(3), 279–283. https://doi.org/10.1097/00000421-199806000-00015

Faivre, S., Chan, D., Salinas, R., Woynarowska, B., & Woynarowski, J. M. (2003). DNA strand breaks and apoptosis induced by oxaliplatin in cancer cells. *Biochemical Pharmacology*, *66*(2), 225–237. https://doi.org/10.1016/S0006-2952(03)00260-0

Fang, L., Yang, Z., Zhou, J., Tung, J. Y., Hsiao, C. D., Wang, L., Deng, Y., Wang, P., Wang, J., & Lee, M. H. (2015). Circadian clock gene CRY2 degradation is involved in chemoresistance of colorectal cancer. *Molecular Cancer Therapeutics*, *14*(6), 1476–1487. https://doi.org/10.1158/1535-7163.MCT-15-0030

Fang, Z., Gong, C., Yu, S., Zhou, W., Hassan, W., Li, H., Wang, X., Hu, Y., Gu, K., Chen, X., Hong, B., Bao, Y., Chen, X., Zhang, X., & Liu, H. (2018). NFYB-induced high expression of E2F1 contributes to oxaliplatin resistance in colorectal cancer via the enhancement of CHK1 signaling. *Cancer Letters*, *415*, 58–72. https://doi.org/10.1016/j.canlet.2017.11.040

Fearnhead, N. S., Britton, M. P., & Bodmer, W. F. (2001). The ABC of APC. *Human Molecular Genetics*, *10*(7), 721–733. https://doi.org/10.1093/hmg/10.7.721

Fearon, E. F., & Vogelstein, B. (1990). A genetic model for colorectal tumorigenesis. *Cell, 61*(5), 759-67. https://doi.org/10.1016/0092-8674(90)90186-i

Fischel, J. L., Formento, P., Ciccolini, J., Rostagno, P., Etienne, M. C., Catalin, J., & Milano, G. (2002). Impact of the oxaliplatin-5 fluorouracil-folinic acid combination on respective intracellular determinants of drug activity. *British Journal of Cancer*, *86*(7), 1162–1168. https://doi.org/10.1038/SJ.BJC.6600185

Floriani, I., Torri, V., Rulli, E., Garavaglia, D., Compagnoni, A., Salvolini, L., & Giovagnoni, A. (2010). Performance of imaging modalities in diagnosis of liver metastases from colorectal cancer: A systematic review and meta-analysis. *Journal of Magnetic Resonance Imaging*, *31*(1), 19–31. https://doi.org/10.1002/JMRI.22010

Fontana, E., Eason, K., Cervantes, A., Salazar, R., & Sadanandam, A. (2019). Context matters—consensus molecular subtypes of colorectal cancer as biomarkers for clinical trials. *Annals of Oncology*, *30*(4), 520–527. https://doi.org/10.1093/ANNONC/MDZ052

Gafter-Gvili, A., Zingerman, B., Rozen-Zvi, B., Ori, Y., Green, H., Lubin, I., Malachi, T., Gafter, U., & Herman-Edelstein, M. (2013). Oxidative stress-induced DNA damage and repair in human peripheral blood mononuclear cells: protective role of hemoglobin. *PloS One*, *8*(7). https://doi.org/10.1371/JOURNAL.PONE.0068341

Galon, J., Costes, A., Sanchez-Cabo, F., Kirilovsky, A., Mlecnik, B., Lagorce-Pagès, C., Tosolini, M., Camus, M., Berger, A., Wind, P., Zinzindohoué, F., Bruneval, P., Cugnenc, P. H., Trajanoski, Z., Fridman, W. H., & Pagès, F. (2006). Type, density, and location of immune cells within human colorectal tumors predict clinical

outcome. *Science*, *313*(5795), 1960–1964. https://doi.org/10.1126/SCIENCE.1129139

Galon, J., Mlecnik, B., Marliot, F., Ou, F.-S., Bifulco, C. B., Lugli, A., Zlobec, I., Rau, T. T., Hartmann, A., Masucci, G. V., Zavadova, E., Ohashi, P., Roehrl, M. H. A., Kawakami, Y., Torigoe, T., Ascierto, P. A., Marincola, F., Sargent, D. J., Fox, B. A., & Pages, F. (2016). Validation of the Immunoscore (IM) as a prognostic marker in stage I/II/III colon cancer: Results of a worldwide consortium-based analysis of 1,336 patients. *Journal of Clinical Oncology*, *34*(15), 3500–3500. https://doi.org/10.1200/JCO.2016.34.15_SUPPL.3500

Gastl, B., Klotz-Noack, K., Klinger, B., Ispasanie, S., Salib, K. H. F., Zuber, J., Mamlouk, S., Bublitz, N., Blüthgen, N., Horst, D., Morkel, M., Schäfer, R., & Sers, C. (2020). Reduced replication origin licensing selectively kills KRAS-mutant colorectal cancer cells via mitotic catastrophe. *Cell Death & Disease 2020 11:7*, *11*(7), 1–15. https://doi.org/10.1038/s41419-020-2704-9

Geiersbach, K. B., & Samowitz, W. S. (2011). Microsatellite instability and colorectal cancer. *Archives of Pathology & Laboratory Medicine*, *135*(10), 1269–1277. https://doi.org/10.5858/ARPA.2011-0035-RA

Gervaz, P., Bucher, P., & Morel, P. (2004). Two colons-two cancers: Paradigm shift and clinical implications. *Journal of Surgical Oncology*, *88*(4), 261–266. https://doi.org/10.1002/jso.20156

Gkekas, I., Novotny, J., Pecen, L., Strigård, K., Palmqvist, R., & Gunnarsson, U. (2017). Microsatellite instability as a prognostic factor in stage II colon cancer patients, a meta-analysis of published literature. *Anticancer Research*, *37*(12), 6563–6574. https://doi.org/10.21873/ANTICANRES.12113

Glaire, M. A., Domingo, E., Sveen, A., Bruun, J., Nesbakken, A., Nicholson, G., Novelli, M., Lawson, K., Oukrif, D., Kildal, W., Danielsen, H. E., Kerr, R., Kerr, D., Tomlinson, I., Lothe, R. A., & Church, D. N. (2019). Tumour-infiltrating CD8+ lymphocytes and colorectal cancer recurrence by tumour and nodal stage. *British Journal of Cancer*, *121*(6), 474–482. https://doi.org/10.1038/s41416-019-0540-4

Glickman, M. E., Rao, S. R., & Schultz, M. R. (2014). False discovery rate control is a recommended alternative to Bonferroni-type adjustments in health studies. *Journal of Clinical Epidemiology*, *67*(8), 850–857. https://doi.org/10.1016/J.JCLINEPI.2014.03.012

Goka, E. T., Chaturvedi, P., Lopez, D. T. M., De La Garza, A., & Lippman, M. E. (2019). RAC1b overexpression confers resistance to chemotherapy treatment in colorectal cancer. *Molecular Cancer Therapeutics*, *18*(5), 957–968. https://doi.org/10.1158/1535-7163.MCT-18-0955

Goldberg, R. M., Sargent, D. J., Morton, R. F., Fuchs, C. S., Ramanathan, R. K., Williamson, S. K., Findlay, B. P., Pitot, H. C., & Alberts, S. R. (2004). A randomized controlled trial of fluorouracil plus leucovorin, irinotecan, and

oxaliplatin combinations in patients with previously untreated metastatic colorectal cancer. *Journal of Clinical Oncology*, *22*(1), 23–30. https://doi.org/10.1200/JCO.2004.09.046

Gourdier, I., del Rio, M., Crabbé, L., Candeil, L., Copois, V., Ychou, M., Auffray, C., Martineau, P., Mechti, N., Pommier, Y., & Pau, B. (2002). Drug specific resistance to oxaliplatin is associated with apoptosis defect in a cellular model of colon carcinoma. *FEBS Letters*, *529*(2–3), 232–236. https://doi.org/10.1016/S0014-5793(02)03347-1

Graham, M. A., Lockwood, G. F., Greenslade, D., Brienza, S., Bayssas, M., & Gamelin, E. (2000). Clinical pharmacokinetics of oxaliplatin: a critical review. *Clinical Cancer Research*, *6*(4), 1205-1218. https://pubmed.ncbi.nlm.nih.gov/10778943/

Grothey A, & Schmoll HJ. (2001). New chemotherapy approaches in colorectal cancer. *Current Opinion in Oncology*, *13*(4), 275–286. https://doi.org/10.1097/00001622-200107000-00011

Guan, Y., Kraus, S. G., Quaney, M. J., Daniels, M. A., Mitchem, J. B., & Teixeiro, E. (2020). FOLFOX chemotherapy ameliorates CD8 T lymphocyte exhaustion and enhances checkpoint blockade efficacy in colorectal cancer. *Frontiers in Oncology*, *0*, 586. https://doi.org/10.3389/FONC.2020.00586

Guastadisegni, C., Colafranceschi, M., Ottini, L., & Dogliotti, E. (2010). Microsatellite instability as a marker of prognosis and response to therapy: a meta-analysis of colorectal cancer survival data. *European Journal of Cancer*, *46*(15), 2788–2798. https://doi.org/10.1016/J.EJCA.2010.05.009

Guibert, S., & Weber, M. (2013). Functions of DNA methylation and hydroxymethylation in mammalian development. *Current Topics in Developmental Biology*, *104*, 47–83. https://doi.org/10.1016/B978-0-12-416027-9.00002-4

Guinney, J., Dienstmann, R., Wang, X., de Reyniès, A., Schlicker, A., Soneson, C., Marisa, L., Roepman, P., Nyamundanda, G., Angelino, P., Bot, B. M., Morris, J. S., Simon, I. M., Gerster, S., Fessler, E., de Sousa .E Melo, F., Missiaglia, E., Ramay, H., Barras, D., … Tejpar, S. (2015). The consensus molecular subtypes of colorectal cancer. *Nature Medicine*, *21*(11), 1350–1356. https://doi.org/10.1038/nm.3967

Hayward, R. L., Macpherson, J. S., Cummings, J., Monia, B. P., Smyth, J. F., & Jodrell, D. I. (2004). Enhanced oxaliplatin-induced apoptosis following antisense Bcl-xl down-regulation is p53 and Bax dependent: Genetic evidence for specificity of the antisense effect. *Molecular Cancer Therapeutics*, *3*(2), 169–178. https://pubmed.ncbi.nlm.nih.gov/14985457/

He, J. Ming, Wang, F. Chao, Qi, H. Bing, Li, Y., & Liang, H. Jie. (2009). Down-regulation of alphav integrin by retroviral delivery of small interfering RNA reduces multicellular resistance of HT29. *Cancer Letters*, *284*(2), 182–188. https://doi.org/10.1016/J.CANLET.2009.04.023

Hoesel, B., & Schmid, J. A. (2013). The complexity of NF-κB signaling in inflammation and cancer. *Molecular Cancer*, *12*(1), 1–15. https://doi.org/10.1186/1476-4598-12-86

Hongo, K., Tsuno, N. H., Kawai, K., Sasaki, K., Kaneko, M., Hiyoshi, M., Murono, K., Tada, N., Nirei, T., Sunami, E., Takahashi, K., Nagawa, H., Kitayama, J., & Watanabe, T. (2013). Hypoxia enhances colon cancer migration and invasion through promotion of epithelial-mesenchymal transition. *Journal of Surgical Research*, *182*(1), 75–84. https://doi.org/10.1016/j.jss.2012.08.034

Howell, S. B., Safaei, R., Larson, C. A., & Sailor, M. J. (2010). Copper Transporters and the Cellular Pharmacology of the Platinum-Containing Cancer Drugs. *Molecular Pharmacology*, *77*(6), 887–894. https://doi.org/10.1124/MOL.109.063172

Hsu, H. H., Kuo, W. W., Shih, H. N., Cheng, S. F., Yang, C. K., Chen, M. C., Tu, C. C., Viswanadha, V. P., Liao, P. H., & Huang, C. Y. (2019). FOXC1 Regulation of miR-31-5p confers oxaliplatin resistance by targeting LATS2 in colorectal cancer. *Cancers*, *11*(10). https://doi.org/10.3390/cancers11101576

Hua, Y., Zhu, Y., Zhang, J., Zhu, Z., Ning, Z., Chen, H., Liu, L., Chen, Z., & Meng, Z. (2018). MiR-122 targets X-linked inhibitor of apoptosis protein to sensitize oxaliplatin-resistant colorectal cancer cells to oxaliplatin-mediated cytotoxicity. *Cellular Physiology and Biochemistry*, *51*(5), 2148–2159. https://doi.org/10.1159/000495832

Hughes, L. A. E., Khalid-de Bakker, C. A. J., Smits, K. M., van den Brandt, P. A., Jonkers, D., Ahuja, N., Herman, J. G., Weijenberg, M. P., & van Engeland, M. (2012). The CpG island methylator phenotype in colorectal cancer: Progress and problems. *Biochimica et Biophysica Acta - Reviews on Cancer*, *1825*(1), 77–85. https://doi.org/10.1016/j.bbcan.2011.10.005

Huijbers, A., Tollenaar, R. A. E. M., Pelt, G. W. V., Zeestraten, E. C. M., Dutton, S., McConkey, C. C., Domingo, E., Smit, V. T. H. B. M., Midgley, R., Warren, B. F., Johnstone, E. C., Kerr, D. J., & Mesker, W. E. (2013). The proportion of tumor-stroma as a strong prognosticator for stage II and III colon cancer patients: validation in the VICTOR trial. *Annals of Oncology*, *24*(1), 179–185. https://doi.org/10.1093/ANNONC/MDS246

Hutchins, G., Southward, K., Handley, K., Magill, L., Beaumont, C., Stahlschmidt, J., Richman, S., Chambers, P., Seymour, M., Kerr, D., Gray, R., & Quirke, P. (2011). Value of mismatch repair, KRAS, and BRAF mutations in predicting recurrence and benefits from chemotherapy in colorectal cancer. *Journal of Clinical Oncology*, *29*(10), 1261–1270. https://doi.org/10.1200/JCO.2010.30.1366

Imamura, Y., Morikawa, T., Liao, X., Lochhead, P., Kuchiba, A., Yamauchi, M., Qian, Z. R., Nishihara, R., Meyerhardt, J. A., Haigis, K. M., Fuchs, C. S., & Ogino, S. (2012). Specific mutations in KRAS codons 12 and 13, and patient prognosis in

1075 BRAF wild-type colorectal cancers. *Clinical Cancer Research*, *18*(17), 4753–4763. https://doi.org/10.1158/1078-0432.CCR-11-3210

Isella, C., Brundu, F., Bellomo, S. E., Galimi, F., Zanella, E., Porporato, R., Petti, C., Fiori, A., Orzan, F., Senetta, R., Boccaccio, C., Ficarra, E., Marchionni, L., Trusolino, L., Medico, E., & Bertotti, A. (2017). Selective analysis of cancer-cell intrinsic transcriptional traits defines novel clinically relevant subtypes of colorectal cancer. *Nature Communications*, *8*(1), 1–16. https://doi.org/10.1038/ncomms15107

Isella, C., Terrasi, A., Bellomo, S. E., Petti, C., Galatola, G., Muratore, A., Mellano, A., Senetta, R., Cassenti, A., Sonetto, C., Inghirami, G., Trusolino, L., Fekete, Z., de Ridder, M., Cassoni, P., Storme, G., Bertotti, A., & Medico, E. (2015). Stromal contribution to the colorectal cancer transcriptome. *Nature Genetics*, *47*(4), 312–319. https://doi.org/10.1038/NG.3224

Itatani, Y., Kawada, K., & Sakai, Y. (2019). Transforming growth factor-β signaling pathway in colorectal cancer and its tumor microenvironment. *International Journal of Molecular Sciences*, *20*(23). https://doi.org/10.3390/ijms20235822

Jackson, S. P., & Bartek, J. (2009). The DNA-damage response in human biology and disease. *Nature*, *461*(7267), 1071. https://doi.org/10.1038/NATURE08467

Jackstadt, R., van Hooff, S. R., Leach, J. D., Cortes-Lavaud, X., Lohuis, J. O., Ridgway, R. A., Wouters, V. M., Roper, J., Kendall, T. J., Roxburgh, C. S., Horgan, P. G., Nixon, C., Nourse, C., Gunzer, M., Clark, W., Hedley, A., Yilmaz, O. H., Rashid, M., Bailey, P., … Sansom, O. J. (2019). Epithelial NOTCH signaling rewires the tumor microenvironment of colorectal cancer to drive poor-prognosis subtypes and metastasis. *Cancer Cell*, *36*(3), 319-336. https://doi.org/10.1016/J.CCELL.2019.08.003

Jančík, S., Hajdúch, M., Drábek, J., & Radzioch, D. (2010). Clinical relevance of KRAS in human cancers. *Journal of Biomedicine and Biotechnology*, 2010, 13. https://doi.org/10.1155/2010/150960

Jiang, T., Ye, L., Han, Z., Liu, Y., Yang, Y., Peng, Z., & Fan, J. (2017). MiR-19b-3p promotes colon cancer proliferation and oxaliplatin-based chemoresistance by targeting SMAD4: Validation by bioinformatics and experimental analyses. *Journal of Experimental and Clinical Cancer Research*, *36*(1). https://doi.org/10.1186/s13046-017-0602-5

Johnson, K. R., Ringland, C., Stokes, B. J., Anthony, D. M., Freemantle, N., Irs, A., Hill, S. R., & Ward, R. L. (2006). Response rate or time to progression as predictors of survival in trials of metastatic colorectal cancer or non-small-cell lung cancer: a meta-analysis. *The Lancet. Oncology*, *7*(9), 741–746. https://doi.org/10.1016/S1470-2045(06)70800-2

Jong, N. N., Nakanishi, T., Liu, J. J., Tamai, I., & McKeage, M. J. (2011). Oxaliplatin transport mediated by organic cation/carnitine transporters OCTN1 and OCTN2 in overexpressing human embryonic kidney 293 cells and rat dorsal root ganglion

neurons. *The Journal of Pharmacology and Experimental Therapeutics*, *338*(2), 537–547. https://doi.org/10.1124/JPET.111.181297

Jongen, J. M. J., van der Waals, L. M., Trumpi, K., Laoukili, J., Peters, N. A., van Schelven, S. J. S., Govaert, K. M., Rinkes, I. H. M. B., & Kranenburg, O. (2017). Downregulation of DNA repair proteins and increased DNA damage in hypoxic colon cancer cells is a therapeutically exploitable vulnerability. *Oncotarget*, *8*(49), 86296–86311. https://doi.org/10.18632/ONCOTARGET.21145

Ju, H. Q., Lu, Y. X., Wu, Q. N., Liu, J., Zeng, Z. L., Mo, H. Y., Chen, Y., Tian, T., Wang, Y., Kang, T. B., Xie, D., Zeng, M. S., Huang, P., & Xu, R. H. (2017). Disrupting G6PD-mediated Redox homeostasis enhances chemosensitivity in colorectal cancer. *Oncogene*, *36*(45), 6282–6292. https://doi.org/10.1038/onc.2017.227

Jung, Y., & Lippard, S. J. (2003). Multiple states of stalled T7 RNA polymerase at DNA lesions generated by platinum anticancer agents. *The Journal of Biological Chemistry*, *278*(52), 52084–52092. https://doi.org/10.1074/JBC.M310120200

Kaminski, M. F., Regula, J., Kraszewska, E., Polkowski, M., Wojciechowska, U., Didkowska, J., Zwierko, M., Rupinski, M., Nowacki, M. P., & Butruk, E. (2010). Quality indicators for colonoscopy and the risk of interval cancer. *New England Journal of Medicine*, *362*(19), 1795–1803. https://doi.org/10.1056/NEJMOA0907667/SUPPL_FILE/NEJM_KAMINSKI_1795SA1.PDF

Kawakami, H., Zaanan, A., & Sinicrope, F. A. (2015). Microsatellite instability testing and its role in the management of colorectal cancer. *Current Treatment Options in Oncology*, *16*(7), 30. https://doi.org/10.1007/s11864-015-0348-2

Kawanishi, S., Hiraku, Y., Pinlaor, S., & Ma, N. (2006). Oxidative and nitrative DNA damage in animals and patients with inflammatory diseases in relation to inflammation-related carcinogenesis. *Biological Chemistry*, *387*(4), 365–372. https://doi.org/10.1515/BC.2006.049

Kekelidze, M., D'Errico, L., Pansini, M., Tyndall, A., & Hohmann, J. (2013). Colorectal cancer: Current imaging methods and future perspectives for the diagnosis, staging and therapeutic response evaluation. *World Journal of Gastroenterology*, *19*(46), 8502. https://doi.org/10.3748/WJG.V19.I46.8502

Kerr, D. J., Domingo, E., & Kerr, R. (2016). Is sidedness prognostically important across all stages of colorectal cancer? *The Lancet Oncology*, *17*(11), 1480–1482. https://doi.org/10.1016/S1470-2045(16)30431-4

Klaus, A., & Birchmeier, W. (2008). Wnt signalling and its impact on development and cancer. *Nature Reviews Cancer*, *8*(5), 387–398. https://doi.org/10.1038/NRC2389

Kondo, T., Kanai, M., Kou, T., Sakuma, T., Mochizuki, H., Kamada, M., Nakatsui, M., Uza, N., Kodama, Y., Masui, T., Takaori, K., Matsumoto, S., Miyake, H., Okuno, Y., & Muto, M. (2018). Association between homologous recombination repair

gene mutations and response to oxaliplatin in pancreatic cancer. *Oncotarget*, *9*(28), 19817. https://doi.org/10.18632/ONCOTARGET.24865

Kopetz, S., Grothey, A., Yaeger, R., van Cutsem, E., Desai, J., Yoshino, T., Wasan, H., Ciardiello, F., Loupakis, F., Hong, Y. S., Steeghs, N., Guren, T. K., Arkenau, H.-T., Garcia-Alfonso, P., Pfeiffer, P., Orlov, S., Lonardi, S., Elez, E., Kim, T.-W., … Tabernero, J. (2019). Encorafenib, binimetinib, and cetuximab in BRAF V600E-mutated colorectal cancer. *The New England Journal of Medicine*, *381*(17), 1632–1643. https://doi.org/10.1056/NEJMOA1908075

Koutras, A. K., Makatsoris, T., Paliogianni, F., Kopsida, G., Onyenadum, A., Gogos, C. A., Mouzaki, A., & Kalofonos, H. P. (2004). Oxaliplatin induced acute onset thrombocytopenia, hemorrhage and hemolysis. *Oncology*, *67*(2), 179–182. https://doi.org/10.1159/000081006

Kovtun, I. v., Liu, Y., Bjoras, M., Klungland, A., Wilson, S. H., & McMurray, C. T. (2007). OGG1 initiates age-dependent CAG trinucleotide expansion in somatic cells. *Nature*, *447*(7143), 447. https://doi.org/10.1038/NATURE05778

Krausova, M., & Korinek, V. (2014). Wnt signaling in adult intestinal stem cells and cancer. *Cellular Signalling*, *26*(3), 570–579. https://doi.org/10.1016/J.CELLSIG.2013.11.032

Krokan, H. E., & Bjørås, M. (2013). Base excision repair. *Cold Spring Harbor Perspectives in Biology*, *5*(4), 1–22. https://doi.org/10.1101/CSHPERSPECT.A012583

Kuipers, E. J., Grady, W. M., Lieberman, D., Seufferlein, T., Sung, J. J., Boelens, P. G., Velde, C. J. H. van de, & Watanabe, T. (2015). Colorectal cancer. *Nature Reviews Disease Primers*, *1*, 15065. https://doi.org/10.1038/NRDP.2015.65

Kunkel, T. A. (2009). Evolving views of DNA replication (in)fidelity. *Cold Spring Harbor Symposia on Quantitative Biology*, *74*, 91–101. https://doi.org/10.1101/SQB.2009.74.027

Lan, Y. T., Chang, S. C., Fen-Yau Li, A., Lin, T. C., Chen, W. S., Jiang, J. K., Yang, S. H., Wang, H. S., & Lin, J. K. (2007). P53 protein accumulation as a prognostic marker in sporadic colorectal cancer. *International Journal of Colorectal Disease*, *22*(5), 499–506. https://doi.org/10.1007/s00384-006-0194-6

Lanza, G., Maestri, I., Dubini, A., Gafa', R., Santini, A., Ferretti, S., & Cavazzini, L. (1996). p53 Expression in colorectal cancer: Relation to tumor type, DNA ploidy pattern, and short-term survival. *American Journal of Clinical Pathology*, *105*(5), 604–612. https://doi.org/10.1093/ajcp/105.5.604

Lao, V. V., & Grady, W. M. (2011). Epigenetics and colorectal cancer. *Nature Reviews Gastroenterology and Hepatology*, *8*(12), 686–700. https://doi.org/10.1038/nrgastro.2011.173

Ledys, F., Klopfenstein, Q., Truntzer, C., Arnould, L., Vincent, J., Bengrine, L., Remark, R., Boidot, R., Ladoire, S., Ghiringhelli, F., & Derangere, V. (2018). RAS status and neoadjuvant chemotherapy impact CD8+ cells and tumor HLA class I expression in liver metastatic colorectal cancer. *Journal for ImmunoTherapy of Cancer*, *6*(1), 123. https://doi.org/10.1186/S40425-018-0438-3

Leguisamo, N. M., Gloria, H. C., Kalil, A. N., Martins, T. v., Azambuja, D. B., Meira, L. B., & Saffi, J. (2017). Base excision repair imbalance in colorectal cancer has prognostic value and modulates response to chemotherapy. *Oncotarget*, *8*(33), 54199. https://doi.org/10.18632/ONCOTARGET.14909

Lévi, F., Zidani, R., & Misset, J. L. (1997). Randomised multicentre trial of chronotherapy with oxaliplatin, fluorouracil, and folinic acid in metastatic colorectal cancer. *Lancet*, *350*(9079), 681–686. https://doi.org/10.1016/S0140-6736(97)03358-8

Lewis, S. (2007). Regression analysis. *Practical Neurology*, *7*(4), 259–264. https://doi.org/10.1136/JNNP.2007.120055

Li, P., Zhang, X., Wang, H., Wang, L., Liu, T., Du, L., Yang, Y., & Wang, C. (2017). MALAT1 is associated with poor response to oxaliplatin-based chemotherapy in colorectal cancer patients and promotes chemoresistance through EZH2. *Molecular Cancer Therapeutics*, *16*(4), 739–751. https://doi.org/10.1158/1535-7163.MCT-16-0591

Lièvre, A., Bachet, J. B., Boige, V., Cayre, A., le Corre, D., Buc, E., Ychou, M., Bouché, O., Landi, B., Louvet, C., André, T., Bibeau, F., Diebold, M. D., Rougier, P., Ducreux, M., Tomasic, G., Emile, J. F., Penault-Llorca, F., & Laurent-Puig, P. (2008a). KRAS mutations as an independent prognostic factor in patients with advanced colorectal cancer treated with cetuximab. *Journal of Clinical Oncology*, *26*(3), 374–379. https://doi.org/10.1200/JCO.2007.12.5906

Lièvre, A., Bachet, J. B., Boige, V., Cayre, A., le Corre, D., Buc, E., Ychou, M., Bouché, O., Landi, B., Louvet, C., André, T., Bibeau, F., Diebold, M. D., Rougier, P., Ducreux, M., Tomasic, G., Emile, J. F., Penault-Llorca, F., & Laurent-Puig, P. (2008b). KRAS mutations as an independent prognostic factor in patients with advanced colorectal cancer treated with cetuximab. *Journal of Clinical Oncology*, *26*(3), 374–379. https://doi.org/10.1200/JCO.2007.12.5906

Lin, P. L., Chang, J. T., Wu, D. W., Huang, C. C., & Lee, H. (2016). Cytoplasmic localization of Nrf2 promotes colorectal cancer with more aggressive tumors via upregulation of PSMD4. *Free Radical Biology & Medicine*, *95*, 121–132. https://doi.org/10.1016/J.FREERADBIOMED.2016.03.014

Lin, S., Lei, K., Du, W., Yang, L., Shi, H., Gao, Y., Yin, P., Liang, X., & Liu, J. (2016). Enhancement of oxaliplatin sensitivity in human colorectal cancer by hypericin mediated photodynamic therapy via ROS-related mechanism. *The International Journal of Biochemistry & Cell Biology*, *71*, 24–34. https://doi.org/10.1016/J.BIOCEL.2015.12.003

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, *15*(12), 1–21. https://doi.org/10.1186/S13059-014-0550-8/FIGURES/9

Lynch, H. T., & de la Chapelle, A. (2003). Hereditary colorectal cancer. *The New England Journal of Medicine*, *348*(10), 919–932. https://doi.org/10.1056/NEJMRA012242

Ma, J., Zeng, S., Zhang, Y., Deng, G., Qu, Y., Guo, C., Yin, L., Han, Y., Cai, C., Li, Y., Wang, G., Bonkovsky, H. L., & Shen, H. (2017). BMP4 promotes oxaliplatin resistance by an induction of epithelial-mesenchymal transition via MEK1/ERK/ELK1 signaling in hepatocellular carcinoma. *Cancer Letters*, *411*, 117–129. https://doi.org/10.1016/j.canlet.2017.09.041

Maindrault-Gœbel, F., De Gramont, A., Louvet, C., André, T., Carola, E., Gilles, V., Lotz, J. P., Tournigand, C., Mabro, M., Molitor, J. L., Artru, P., Izrael, V., & Krulik, M. (2000). Evaluation of oxaliplatin dose intensity in bimonthly leucovorin and 48-hour 5-fluorouracil continuous infusion regimens (FOLFOX) in pretreated metastatic colorectal cancer. *Annals of Oncology*, *11*(11), 1477–1483. https://doi.org/10.1023/a:1026520812351

Maindrault-Gœbel, F., De Gramont, A., Louvet, C., André, T., Carola, E., Mabro, M., Artru, P., Gilles, V., Lotz, J. P., Izrael, V., & Krulik, M. (2001). High-dose intensity oxaliplatin added to the simplified bimonthly leucovorin and 5-fluorouracil regimen as second-line therapy for metastatic colorectal cancer (FOLFOX 7). *European Journal of Cancer*, *37*(8), 1000–1005. https://doi.org/10.1016/S0959-8049(01)00068-5

Maindrault-Goebel, F., Louvet, C., André, T., Carola, E., Lotz, J. P., Molitor, J. L., Garcia, M. L., Gilles-Amar, V., Izrael, V., Krulik, M., & De Gramont, A. (1999). Oxaliplatin added to the simplified bimonthly leucovorin and 5-fluorouracil regimen as second-line therapy for metastatic colorectal cancer (FOLFOX6). *European Journal of Cancer*, *35*(9), 1338–1342. https://doi.org/10.1016/S0959-8049(99)00149-5

Malla, S. B., Fisher, D. J., Domingo, E., Blake, A., Hassanieh, S., Redmond, K. L., Richman, S. D., Youdell, M., Walker, S. M., Logan, G. E., Chatzipli, A., Amirkhah, R., Humphries, M. P., Craig, S. G., McDermott, U., Seymour, M. T., Morton, D. G., Quirke, P., West, N. P., … Dunne, P. D. (2021). In-depth clinical and biological exploration of DNA damage immune response as a biomarker for oxaliplatin use in colorectal cancer. *Clinical Cancer Research*, *27*(1), 288-300. https://doi.org/10.1158/1078-0432.CCR-20-3237

Man, J., Yu, X., Huang, H., Zhou, W., Xiang, C., Huang, H., Miele, L., Liu, Z., Bebek, G., Bao, S., & Yu, J. S. (2018). Hypoxic induction of vasorin regulates Notch1 turnover to maintain glioma stem-like cells. *Cell Stem Cell*, *22*(1), 104-118.e6. https://doi.org/10.1016/j.stem.2017.10.005

Mandard, A.-M. , Dalibard, F. , Mandard, J.-C. , Marnay, J. , Henry-Amar, M. , Petiot, J.-F., & Gignoux, M. (1994). Pathologic assessment of tumor regression after

preoperative chemoradiotherapy of esophageal carcinoma. *Cancers, 73*(11), 2680-2686. https://doi.org/10.1002/1097-0142(19940601)73:11<2680::aid-cncr2820731105>3.0.co;2-c

Marks, E. I., Matera, R., Olszewski, A. J., Yakirevich, E., El-Deiry, W. S., Safran, H., & Carneiro, B. A. (2021). Mutations in DNA repair genes and clinical outcomes of patients with metastatic colorectal cancer receiving oxaliplatin or irinotecan-containing regimens. *American Journal of Clinical Oncology: Cancer Clinical Trials*, *44*(2), 68–73. https://doi.org/10.1097/COC.0000000000000785

Martinez-Balibrea, E., Martínez-Cardus, A., Gines, A., Ruiz De Porras, V., Moutinho, C., Layos, L., Manzano, J. L., Buges, C., Bystrup, S., Esteller, M., & Abad, A. (2015). Tumor-related molecular mechanisms of oxaliplatin resistance. *Molecular Cancer Therapeutics*, *14*(8), 1767–1776. https://doi.org/10.1158/1535-7163.MCT-14-0636

Martini, G., Dienstmann, R., Ros, J., Baraibar, I., Cuadra-Urteaga, J. L., Salva, F., Ciardiello, D., Mulet, N., Argiles, G., Tabernero, J., & Elez, E. (2020). Molecular subtypes and the evolution of treatment management in metastatic colorectal cancer. *Therapeutic Advances in Medical Oncology*, *12*. https://doi.org/10.1177/1758835920936089

Marzouk, O., & Schofield, J. (2011). Review of histopathological and molecular prognostic features in colorectal cancer. *Cancers*, *3*, 2767–2810. https://doi.org/10.3390/cancers3022767

McGavin JK, & Goa KL. (2001). Capecitabine: a review of its use in the treatment of advanced or metastatic colorectal cancer. *Drugs*, *61*(15), 2309–2326. https://doi.org/10.2165/00003495-200161150-00015

Meisner, N. C., Hintersteiner, M., Mueller, K., Bauer, R., Seifert, J. M., Naegeli, H. U., Ottl, J., Oberer, L., Guenat, C., Moss, S., Harrer, N., Woisetschlaeger, M., Buehler, C., Uhl, V., & Auer, M. (2007). Identification and mechanistic characterization of low-molecular-weight inhibitors for HuR. *Nature Chemical Biology*, *3*(8), 508–515. https://doi.org/10.1038/NCHEMBIO.2007.14

Mi, Y., Mu, L., Huang, K., Hu, Y., Yan, C., Zhao, H., Ma, C., Li, X., Tao, D., & Qin, J. (2020). Hypoxic colorectal cancer cells promote metastasis of normoxic cancer cells depending on IL-8/p65 signaling pathway. *Cell Death and Disease*, *11*(7), 1–11. https://doi.org/10.1038/s41419-020-02797-z

Mitry, E., Guiu, B., Cosconea, S., Jooste, V., Faivre, J., & Bouvier, A. M. (2010). Epidemiology, management and prognosis of colorectal cancer with lung metastases: a 30-year population-based study. *Gut*, *59*(10), 1383–1388. https://doi.org/10.1136/GUT.2010.211557

Modest, D. P., Ricard, I., Heinemann, V., Hegewisch-Becker, S., Schmiegel, W., Porschen, R., Stintzing, S., Graeven, U., Arnold, D., von Weikersthal, L. F., Giessen-Jung, C., Stahler, A., Schmoll, H. J., Jung, A., Kirchner, T., Tannapfel, A., & Reinacher-Schick, A. (2016). Outcome according to KRAS, NRAS and BRAF-

mutation as well as KRAS mutation variants: pooled analysis of five randomized trials in metastatic colorectal cancer by the AIO colorectal cancer study group. *Annals of Oncology*, *27*(9), 1746–1753. https://doi.org/10.1093/ANNONC/MDW261

Molassiotis, A., Cheng, H. L., Lopez, V., Au, J. S. K., Chan, A., Bandla, A., Leung, K. T., Li, Y. C., Wong, K. H., Suen, L. K. P., Chan, C. W., Yorke, J., Farrell, C., & Sundar, R. (2019). Are we mis-estimating chemotherapy-induced peripheral neuropathy? Analysis of assessment methodologies from a prospective, multinational, longitudinal cohort study of patients receiving neurotoxic chemotherapy. *BMC Cancer*, *19*(1). https://doi.org/10.1186/S12885-019-5302-4

Monnier, Y., Farmer, P., Bieler, G., Imaizumi, N., Sengstag, T., Alghisi, G. C., Stehle, J. C., Ciarloni, L., Andrejevic-Blant, S., Moeckli, R., Mirimanoff, R. O., Goodman, S. L., Delorenzi, M., & Rüegg, C. (2008). CYR61 and alphaVbeta5 integrin cooperate to promote invasion and metastasis of tumors growing in preirradiated stroma. *Cancer Research*, *68*(18), 7323–7331. https://doi.org/10.1158/0008-5472.CAN-08-0841

Morikawa, T., Kuchiba, A., Liao, X., Imamura, Y., Yamauchi, M., Qian, Z. R., Nishihara, R., Sato, K., Meyerhardt, J. A., Fuchs, C. S., & Ogino, S. (2012). Tumor TP53 expression status, body mass index and prognosis in colorectal cancer. *International Journal of Cancer*, *131*(5), 1169–1178. https://doi.org/10.1002/ijc.26495

Msc, H., Mrcp, M., Frcpath, J., Bsc, J., Cheadle, J. P., Medical,) ;, Dphil, M., Fisher, D., Kenny, S. L., Ba, K., Mitchell Bhsc, J. K., Kaplan, R., Kaplan, R., Maughan, T. S., Adams, R. A., Smith, C. G., Meade, A. M., Seymour, M. T., Wilson, R. H., … Cheadle, J. P. (2011). Addition of cetuximab to oxaliplatin-based first-line combination chemotherapy for treatment of advanced colorectal cancer: results of the randomised phase 3 MRC COIN trial. *Lancet*, *377*, 2103–2117. https://doi.org/10.1016/S0140

Chang, M. T., Asthana, S., Gao, S. P., Lee, B. H., Chapman, J. S., Kandoth, C., Gao, J., Socci, N. D., Solit, D. B., Olshen, A. B., Schultz, N., & Taylor, B. S. (2016). Identifying recurrent mutations in cancer reveals widespread lineage diversity and mutational specificity. *Nature Biotechnology*, *34*(2), 155–163. https://doi.org/10.1038/NBT.3391

Mulligan, J. M., Hill, L. A., Deharo, S., Irwin, G., Boyle, D., Keating, K. E., Raji, O. Y., McDyer, F. A., O'Brien, E., Bylesjo, M., Quinn, J. E., Lindor, N. M., Mullan, P. B., James, C. R., Walker, S. M., Kerr, P., James, J., Davison, T. S., Proutski, V., … Kennedy, R. D. (2014). Identification and validation of an anthracycline/cyclophosphamide-based chemotherapy response assay in breast cancer. *Journal of the National Cancer Institute*, *106*(1). https://doi.org/10.1093/JNCI/DJT335

Mulligan, J. M., Hill, L. A., Deharo, S., Irwin, G., Boyle, D., Keating, K. E., Raji, O. Y., McDyer, F. A., O'Brien, E., Bylesjo, M., Quinn, J. E., Lindor, N. M., Mullan, P. B., James, C. R., Walker, S. M., Kerr, P., James, J., Davison, T. S., Proutski, V.,

… Kennedy, R. D. (2014). Identification and validation of an anthracycline/cyclophosphamide–based chemotherapy response assay in breast cancer. *Journal of the National Cancer Institute*, *106*(1). https://doi.org/10.1093/JNCI/DJT335

Muzny, D. M., Bainbridge, M. N., Chang, K., Dinh, H. H., Drummond, J. A., Fowler, G., Kovar, C. L., Lewis, L. R., Morgan, M. B., Newsham, I. F., Reid, J. G., Santibanez, J., Shinbrot, E., Trevino, L. R., Wu, Y. Q., Wang, M., Gunaratne, P., Donehower, L. A., Creighton, C. J., … Thomson., E. (2012). Comprehensive molecular characterization of human colon and rectal cancer. *Nature*, *487*(7407), 330–337. https://doi.org/10.1038/nature11252

Nicum S, Midgley R, & Kerr DJ. (2000). Chemotherapy for colorectal cancer. *Journal of the Royal Society of Medicine*, *93*(8), 416–419. https://doi.org/10.1177/014107680009300807

Noske, A., Lipka, S., Budczies, J., Müller, K., Loddenkemper, C., Buhr, H. J., & Kruschewski, M. (2009). Combination of p53 expression and p21 loss has an independent prognostic impact on sporadic colorectal cancer. *Oncology Reports*, *22*(1), 3–9. https://doi.org/10.3892/or_00000398

Noubissi, F. K., Elcheva, I., Bhatia, N., Shakoori, A., Ougolkov, A., Liu, J., Minamoto, T., Ross, J., Fuchs, S. Y., & Spiegelman, V. S. (2006). CRD-BP mediates stabilization of βTrCP1 and c-myc mRNA in response to β-catenin signalling. *Nature*, *441*(7095), 898–901. https://doi.org/10.1038/nature04839

O'cathail, S. M., Wu, C.-H., Lewis, A., Holmes, C., Hawkins, M. A., & Maughan, T. (2019). A metagene of NRF2 expression is a prognostic biomarker in all stage colorectal cancer. *BioRxiv*. https://doi.org/10.1101/690974

Ogino, S., Nosho, K., Irahara, N., Shima, K., Baba, Y., Kirkner, G. J., Meyerhardt, J. A., & Fuchs, C. S. (2009). Prognostic significance and molecular associations of 18q loss of heterozygosity: a cohort study of microsatellite stable colorectal cancers. *Journal of Clinical Oncology*, *27*(27), 4591–4598. https://doi.org/10.1200/JCO.2009.22.8858

Oh, M., McBride, A., Yun, S., Bhattacharjee, S., Slack, M., Martin, J. R., Jeter, J., & Abraham, I. (2018). BRCA1 and BRCA2 gene mutations and colorectal cancer risk: systematic review and meta-analysis. *Journal of the National Cancer Institute*, *110*(11), 1178-1189. https://doi.org/10.1093/JNCI/DJY148

Okayama, H. (2012). CDC6: a trifunctional AAA+ ATPase that plays a central role in controlling the G(1)-S transition and cell survival. *Journal of Biochemistry*, *152*(4), 297–303. https://doi.org/10.1093/JB/MVS083

Okita, A., Takahashi, S., Ouchi, K., Inoue, M., Watanabe, M., Endo, M., Honda, H., Yamada, Y., & Ishioka, C. (2018). Consensus molecular subtypes classification of colorectal cancer as a predictive factor for chemotherapeutic efficacy against metastatic colorectal cancer. *Oncotarget*, *9*(27), 18698. https://doi.org/10.18632/ONCOTARGET.24617

Oldenhuis, C. N. A. M., Oosting, S. F., Gietema, J. A., & de Vries, E. G. E. (2008). Prognostic versus predictive value of biomarkers in oncology. *European Journal of Cancer*, *44*(7), 946–953. https://doi.org/10.1016/J.EJCA.2008.03.006

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *British Medical Journal*, *372*. https://doi.org/10.1136/BMJ.N71

Pagès, F., Berger, A., Camus, M., Sanchez-Cabo, F., Costes, A., Molidor, R., Mlecnik, B., Kirilovsky, A., Nilsson, M., Damotte, D., Meatchi, T., Bruneval, P., Cugnenc, P.-H., Trajanoski, Z., Fridman, W.-H., & Galon, J. (2009). Effector memory T cells, early metastasis, and survival in colorectal cancer. *New England Journal of Medicine*, *353*(25), 2654–2666. https://doi.org/10.1056/NEJMOA051424

Pancione, M., Forte, N., Fucci, A., Sabatino, L., Febbraro, A., Di Blasi, A., Daniele, B., Parente, D., & Colantuoni, V. (2010). Prognostic role of β-catenin and p53 expression in the metastatic progression of sporadic colorectal cancer. *Human Pathology*, *41*(6), 867–876. https://doi.org/10.1016/j.humpath.2009.09.019

Perona, R., & Sánchez-Pérez, I. (2004). Control of oncogenesis and cancer therapy resistance. *British Journal of Cancer*, *90*(3), 573–577. https://doi.org/10.1038/SJ.BJC.6601552

Petrelli, F., Ghidini, M., Cabiddu, M., Pezzica, E., Corti, D., Turati, L., Costanzo, A., Varricchio, A., Ghidini, A., Barni, S., & Tomasello, G. (2019). Microsatellite Instability and Survival in Stage II Colorectal Cancer: A Systematic Review and Meta-analysis. *Anticancer Research*, *39*(12), 6431–6441. https://doi.org/10.21873/ANTICANRES.13857

Pickhardt, P. J., Hassan, C., Halligan, S., & Marmo, R. (2011). Colorectal cancer: CT colonography and colonoscopy for detection-systematic review and meta-analysis. *Radiology*, *259*(2), 393–405. https://doi.org/10.1148/RADIOL.11101887/-/DC1

Piedbois, P., & Buyse, M. (2008). Endpoints and surrogate endpoints in colorectal cancer: a review of recent developments. *Current Opinion in Oncology*, *20*(4), 466-471. https://doi.org/10.1097/CCO.0b013e32830218fe

Pino, M. S., & Chung, D. C. (2010). The chromosomal instability pathway in colon cancer. *Gastroenterology*, *138*(6), 2059. https://doi.org/10.1053/J.GASTRO.2009.12.065

Pino, M. S., Kikuchi, H., Zeng, M., Herraiz, M. T., Sperduti, I., Berger, D., Park, D. Y., Iafrate, A. J., Zukerberg, L. R., & Chung, D. C. (2010). Epithelial to mesenchymal transition is impaired in colon cancer cells with microsatellite instability. *Gastroenterology*, *138*(4), 1406–1417. https://doi.org/10.1053/J.GASTRO.2009.12.010

Pogue-Geile, K. L., Andre, T., Song, N., Lipchik, C., Wang, Y., Kim, R. S., Feng, H., Gavin, P., Laethem, J.-L. van, Srinivasan, A., Hickish, T., Jacobs, S. A., Tabernero, J., Lucas, P. C., Gramont, A. de, Wolmark, N., Flejou, J.-F., & Paik, S. (2019). Association of colon cancer (CC) molecular signatures with prognosis and oxaliplatin prediction-benefit in the MOSAIC Trial (Multicenter International Study of Oxaliplatin/5FU-LV in the Adjuvant Treatment of Colon Cancer). *Journal of Clinical Oncology*, *37*(15), 3503–3503. https://doi.org/10.1200/JCO.2019.37.15_SUPPL.3503

Polley, M. Y. C., & Dignam, J. J. (2021). Statistical considerations in the evaluation of continuous biomarkers. *Journal of Nuclear Medicine*, *62*(5), 605–611. https://doi.org/10.2967/JNUMED.120.251520

Poston, G. J., Tait, D., O'Connell, S., Bennett, A., & Berendse, S. (2011). Diagnosis and management of colorectal cancer: summary of NICE guidance. *British Medical Journal*, *343*(7831), 1010–1012. https://doi.org/10.1136/BMJ.D6751

Pothuraju, R., Rachagani, S., Krishn, S. R., Chaudhary, S., Nimmakayala, R. K., Siddiqui, J. A., Ganguly, K., Lakshmanan, I., Cox, J. L., Mallya, K., Kaur, S., & Batra, S. K. (2020). Molecular implications of MUC5AC-CD44 axis in colorectal cancer progression and chemoresistance. *Molecular Cancer*, *19*(1). https://doi.org/10.1186/s12943-020-01156-y

Puccini, A., Marshall, J. L., & Salem, M. E. (2018). Molecular variances between right- and left-sided colon cancers. *Current Colorectal Cancer Reports*, *14*(5), 152–158. https://doi.org/10.1007/s11888-018-0411-z

Qi, L., Chen, J., Yang, Y., & Hu, W. (2020). Hypoxia correlates with poor survival and m2 macrophage infiltration in colorectal cancer. *Frontiers in Oncology*, *0*, 2491. https://doi.org/10.3389/FONC.2020.566430

Ranjha, L., Howard, S. M., & Cejka, P. (2018). Main steps in DNA double-strand break repair: an introduction to homologous recombination and related processes. *Chromosoma*, *127*(2), 187–214. https://doi.org/10.1007/S00412-017-0658-1

Rao, D., Mallick, A. B., Augustine, T., Daroqui, C., Jiffry, J., Merla, A., Chaudhary, I., Seetharam, R., Sood, A., Gajavelli, S., Aparo, S., Rajdev, L., Kaubisch, A., Chuy, J., Negassa, A., Mariadason, J. M., Maitra, R., & Goel, S. (2019). Excision repair cross-complementing group-1 (ERCC1) induction kinetics and polymorphism are markers of inferior outcome in patients with colorectal cancer treated with oxaliplatin. *Oncotarget*, *10*(53), 5510–5522. https://doi.org/10.18632/oncotarget.27140

Richman, S. D., Seymour, M. T., Chambers, P., Elliott, F., Daly, C. L., Meade, A. M., Taylor, G., Barrett, J. H., & Quirke, P. (2009). KRAS and BRAF mutations in advanced colorectal cancer are associated with poor prognosis but do not preclude benefit from oxaliplatin or irinotecan: results from the MRC FOCUS trial. *Journal of Clinical Oncology*, *27*(35), 5931–5937. https://doi.org/10.1200/JCO.2009.22.4295

Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., & Smyth, G. K. (2015). Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Research*, *43*(7), 47. https://doi.org/10.1093/NAR/GKV007

Roberts, D. L., Williams, K. J., Cowen, R. L., Barathova, M., Eustace, A. J., Brittain-Dissont, S., Tilby, M. J., Pearson, D. G., Ottley, C. J., Stratford, I. J., & Dive, C. (2009). Contribution of HIF-1 and drug penetrance to oxaliplatin resistance in hypoxic colorectal cancer cells. *British Journal of Cancer*, *101*(8), 1290. https://doi.org/10.1038/SJ.BJC.6605311

Roock, W. De, Claes, B., Bernasconi, D., Schutter, J. De, Biesmans, B., Fountzilas, G., Kalogeras, K. T., Kotoula, V., Papamichael, D., Laurent-Puig, P., Penault-Llorca, F., Rougier, P., Vincenzi, B., Santini, D., Tonini, G., Cappuzzo, F., Frattini, M., Molinari, F., Saletti, P., … Tejpar, S. (2010). Effects of KRAS, BRAF, NRAS, and PIK3CA mutations on the efficacy of cetuximab plus chemotherapy in chemotherapy-refractory metastatic colorectal cancer: a retrospective consortium analysis. *The Lancet Oncology*, *11*(8), 753–762. https://doi.org/10.1016/S1470-2045(10)70130-3

Roth, A. D., Delorenzi, M., Tejpar, S., Yan, P., Klingbiel, D., Fiocca, R., D'Ario, G., Cisar, L., Labianca, R., Cunningham, D., Nordlinger, B., Bosman, F., & van Cutsem, E. (2012). Integrated analysis of molecular and clinical prognostic factors in stage II/III colon cancer. *Journal of the National Cancer Institute*, *104*(21), 1635–1646. https://doi.org/10.1093/JNCI/DJS427

Roth, A. D., Tejpar, S., Delorenzi, M., Yan, P., Fiocca, R., Klingbiel, D., Dietrich, D., Biesmans, B., Bodoky, G., Barone, C., Aranda, E., Nordlinger, B., Cisar, L., Labianca, R., Cunningham, D., van Cutsem, E., & Bosman, F. (2010). Prognostic role of KRAS and BRAF in stage II and III resected colon cancer: results of the translational study on the PETACC-3, EORTC 40993, SAKK 60-00 trial. *Journal of Clinical Oncology*, *28*(3), 466–474. https://doi.org/10.1200/JCO.2009.23.3452

Rougier P. (2001). Palliative and adjuvant chemotherapy in colorectal cancer. *European Journal of Cancer*, *37*(7), 189-202. https://doi.org/10.1016/S0959-8049(01)80020-4

Safina, A. F., Varga, A. E., Bianchi, A., Zheng, Q., Kunnev, D., Liang, P., & Bakin, A. v. (2009). Ras alters epithelial-mesenchymal transition in response to TGFbeta by reducing actin fibers and cell-matrix adhesion. *Cell Cycle*, *8*(2), 284–298. https://doi.org/10.4161/CC.8.2.7590

Saif, M. W., & Reardon, J. (2005). Management of oxaliplatin-induced peripheral neuropathy. *Therapeutics and Clinical Risk Management*, *1*(4), 249-258. https://pubmed.ncbi.nlm.nih.gov/18360567/

Sałat, K. (2020). Chemotherapy-induced peripheral neuropathy—part 2: focus on the prevention of oxaliplatin-induced neurotoxicity. *Pharmacological Reports*, *72*(3), 508. https://doi.org/10.1007/S43440-020-00106-1

Samuels, Y., Wang, Z., Bardelli, A., Silliman, N., Ptak, J., Szabo, S., Yan, H., Gazdar, A., Powell, S. M., Riggins, G. J., Willson, J. K. V., Markowitz, S., Kinzler, K. W., Vogelstein, B., & Velculescu, V. E. (2004). High frequency of mutations of the PIK3CA gene in human cancers. *Science*, *304*(5670), 554. https://doi.org/10.1126/SCIENCE.1096502

Sarshekeh, A. M., Advani, S., Overman, M. J., Manyam, G., Kee, B. K., Fogelman, D. R., Dasari, A., Raghav, K., Vilar, E., Manuel, S., Shureiqi, I., Wolff, R. A., Patel, K. P., Luthra, R., Shaw, K., Eng, C., Maru, D. M., Routbort, M. J., Meric-Bernstam, F., & Kopetz, S. (2017). Association of SMAD4 mutation with patient demographics, tumor characteristics, and clinical outcomes in colorectal cancer. *PloS One*, *12*(3), e0173345. https://doi.org/10.1371/JOURNAL.PONE.0173345

Sawayama, H., Miyamoto, Y., Ogawa, K., Yoshida, N., & Baba, H. (2020). Investigation of colorectal cancer in accordance with consensus molecular subtype classification. *Annals of Gastroenterological Surgery*, *4*(5), 528. https://doi.org/10.1002/AGS3.12362

Schirripa, M., Biason, P., Lonardi, S., Pella, N., Simona Pino, M., Urbano, F., Antoniotti, C., Cremolini, C., Corallo, S., Pietrantonio, F., Gelsomino, F., Cascinu, S., Orlandi, A., Munari, G., Malapelle, U., Saggio, S., Fontanini, G., Rugge, M., Mescoli, C., … Fassan, M. (2019). Class 1, 2, and 3 BRAF-mutated metastatic colorectal cancer: a detailed clinical, pathologic, and molecular characterization. *Clinical Cancer Research*, *25*(13), 3954–3961. https://doi.org/10.1158/1078-0432.CCR-19-0311

Schmitz, S., Adams, R., & Walsh, C. (2012). The use of continuous data versus binary data in MTC models: A case study in rheumatoid arthritis. *BMC Medical Research Methodology*, *12*(1), 1–17. https://doi.org/10.1186/1471-2288-12-167/TABLES/6

Schmoll, H. J. (2002). The role of oxaliplatin in the treatment of advanced metastatic colorectal cancer: Prospects and future directions. *Seminars in Oncology*, *29*(5), 34–39. https://doi.org/10.1016/S0093-7754(02)90018-7

Schweizer, L., & Varmus, H. (2003). Wnt/Wingless signaling through β-catenin requires the function of both LRP/Arrow and frizzled classes of receptors. *BMC Cell Biology*, *4*(1), 1–11. https://doi.org/10.1186/1471-2121-4-4/FIGURES/5

Seligmann, J. F., Fisher, D., Elliott, F., Richman, S., Butler, R., Cheadle, J., Adams, R. A., Quirke, P., Maughan, T., Seymour, M. T., & Middleton, G. W. (2015). Exploring the poor outcomes of BRAF mutant (BRAF mut) advanced colorectal cancer (aCRC): Analysis from 2,530 patients (pts) in randomized clinical trials (RCTs). *Journal of Clinical Oncology*, *33*(15), 3509–3509. https://doi.org/10.1200/JCO.2015.33.15_SUPPL.3509

Seligmann, J. F., Fisher, D. J., Brown, L. C., Adams, R. A., Graham, J., Quirke, P., Richman, S. D., Butler, R., Domingo, E., Blake, A., Yates, E., Braun, M., Collinson, F., Jones, R., Brown, E., de Winton, E., Humphrey, T. C., Parmar, M., Kaplan, R., … Maughan, T. S. (2021). Inhibition of WEE1 is effective in TP53-

and RAS-mutant metastatic colorectal cancer: a randomized trial (FOCUS4-C) comparing adavosertib (AZD1775) with active monitoring. *Journal of Clinical Oncology*, *39*(33), 3705–3715. https://doi.org/10.1200/JCO.21.01435

Seoane, J., & Gomis, R. R. (2017). TGF-β Family Signaling in Tumor Suppression and Cancer Progression. *Cold Spring Harbor Perspectives in Biology*, *9*(12), a022277. https://doi.org/10.1101/CSHPERSPECT.A022277

Seymour, M. T., Maughan, T. S., Ledermann, J. A., Topham, C., James, R., Gwyther, S. J., Smith, D. B., Shepherd, S., Maraveyas, A., Ferry, D. R., Meade, A. M., Thompson, L., Griffiths, G. O., Parmar, M. K., & Stephens, R. J. (2007). Different strategies of sequential and combination chemotherapy for patients with poor prognosis advanced colorectal cancer (MRC FOCUS): a randomised controlled trial. *Lancet*, *370*(9582), 143–152. https://doi.org/10.1016/S0140-6736(07)61087-3

Seymour, M. T., & Morton, D. (2019). FOxTROT: an international randomised controlled trial in 1052 patients (pts) evaluating neoadjuvant chemotherapy (NAC) for colon cancer. *Journal of Clinical Oncology*, *37*(15), 3504–3504. https://doi.org/10.1200/JCO.2019.37.15_SUPPL.3504

Slyskova, J., Sabatella, M., Ribeiro-Silva, C., Stok, C., Theil, A. F., Vermeulen, W., & Lans, H. (2018a). Base and nucleotide excision repair facilitate resolution of platinum drugs-induced transcription blockage. *Nucleic Acids Research*, *46*(18), 9537. https://doi.org/10.1093/NAR/GKY764

Slyskova, J., Sabatella, M., Ribeiro-Silva, C., Stok, C., Theil, A. F., Vermeulen, W., & Lans, H. (2018b). Base and nucleotide excision repair facilitate resolution of platinum drugs-induced transcription blockage. *Nucleic Acids Research*, *46*(18), 9537–9549. https://doi.org/10.1093/NAR/GKY764

Smeby, J., Sveen, A., Bergsland, C. H., Eilertsen, I. A., Danielsen, S. A., Eide, P. W., Hektoen, M., Guren, M. G., Nesbakken, A., Bruun, J., & Lothe, R. A. (2019). Exploratory analyses of consensus molecular subtype-dependent associations of TP53 mutations with immunomodulation and prognosis in colorectal cancer. *ESMO Open*, *4*(3). https://doi.org/10.1136/ESMOOPEN-2019-000523

Smeby, J., Sveen, A., Merok, M. A., Danielsen, S. A., Eilertsen, I. A., Guren, M. G., Dienstmann, R., Nesbakken, A., & Lothe, R. A. (2018). CMS-dependent prognostic impact of KRAS and BRAFV600E mutations in primary colorectal cancer. *Annals of Oncology*, *29*(5), 1227. https://doi.org/10.1093/ANNONC/MDY085

Song, N., Pogue-Geile, K. L., Gavin, P. G., Yothers, G., Kim, S. R., Johnson, N. L., Lipchik, C., Allegra, C. J., Petrelli, N. J., O'Connell, M. J., Wolmark, N., & Paik, S. (2016a). Clinical outcome from oxaliplatin treatment in stage II/III colon cancer according to intrinsic subtypes: secondary analysis of NSABP C-07/NRG oncology randomized clinical trial. *Journal of the American Medical Association Oncology*, *2*(9), 1162–1169. https://doi.org/10.1001/JAMAONCOL.2016.2314

Soong, R., Grieu, F., Robbins, P., Dix, B., Chen, D., Parsons, R., House, A., & Iacopetta, B. (1997). p53 Alterations are associated with improved prognosis in distal colonic carcinomas. *Clinical Cancer Research*, *3*(8), 1405–1411. https://pubmed.ncbi.nlm.nih.gov/9815825/

Stojanovska, V., Prakash, M., McQuade, R., Fraser, S., Apostolopoulos, V., Sakkal, S., & Nurgali, K. (2019). Oxaliplatin treatment alters systemic immune responses. *BioMed Research International,* e4650695. https://doi.org/10.1155/2019/4650695

Stracker, T. H., & Petrini, J. H. J. (2011). The MRE11 complex: starting from the ends. *Nature Reviews Molecular Cell Biology*, *12*(2), 90–103. https://doi.org/10.1038/NRM3047

Sun, Y., Jiang, X., Chen, S., Fernandes, N., & Price, B. D. (2005). A role for the Tip60 histone acetyltransferase in the acetylation and activation of ATM. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(37), 13182–13187. https://doi.org/10.1073/PNAS.0504211102

Sundar, R., Miranda, S., Rodrigues, D. N., Chénard-Poirier, M., Dolling, D., Clarke, M., Figueiredo, I., Bertan, C., Yuan, W., Ferreira, A., Chistova, R., Boysen, G., Perez, D. R., Tunariu, N., Mateo, J., Wotherspoon, A., Chau, I., Cunningham, D., Valeri, N., … de Bono, J. (2018). Ataxia telangiectasia mutated protein loss and benefit from oxaliplatin-based chemotherapy in colorectal cancer. *Clinical Colorectal Cancer*, *17*(4), 280–284. https://doi.org/10.1016/J.CLCC.2018.05.011

Sweeney, S. M., Cerami, E., Baras, A., Pugh, T. J., Schultz, N., Stricker, T., Lindsay, J., del Vecchio Fitz, C., Kumari, P., Micheel, C., Shaw, K., Gao, J., Moore, N., Stricker, T., Kandoth, C., Reardon, B., Lepisto, E., Gardos, S., Dang, K., … Shaw, K. (2017). AACR project GENIE: powering precision medicine through an international consortium. *Cancer Discovery*, *7*(8), 818–831. https://doi.org/10.1158/2159-8290.CD-17-0151

Taieb, J., Zaanan, A., Malicot, K. le, Julié, C., Blons, H., Mineur, L., Bennouna, J., Tabernero, J., Mini, E., Folprecht, G., Laethem, J. L. van, Lepage, C., Emile, J.-F., & Laurent-Puig, P. (2016). Prognostic effect of BRAF and kras mutations in patients with stage III colon cancer treated with leucovorin, fluorouracil, and oxaliplatin with or without cetuximab: a post hoc analysis of the PETACC-8 trial. *Journal of the American Medical Association Oncology*, *2*(5), 643–653. https://doi.org/10.1001/JAMAONCOL.2015.5225

Talevich, E., Shain, A. H., Botton, T., & Bastian, B. C. (2016). CNVkit: genome-wide copy number detection and visualization from targeted dna sequencing. *PLOS Computational Biology*, *12*(4), e1004873. https://doi.org/10.1371/JOURNAL.PCBI.1004873

Tang, P. A., Bentzen, S. M., Chen, E. X., & Siu, L. L. (2007). Surrogate end points for median overall survival in metastatic colorectal cancer: literature-based analysis from 39 randomized controlled trials of first-line chemotherapy. *Journal of Clinical Oncology*, *25*(29), 4562–4568. https://doi.org/10.1200/JCO.2006.08.1935

Ten Hoorn, S., Trinh, A., de Jong, J., Koens, L., & Vermeulen, L. (2018). Classification of Colorectal Cancer in Molecular Subtypes by Immunohistochemistry. *Methods in Molecular Biology (Clifton, N.J.)*, *1765*, 179–191. https://doi.org/10.1007/978-1-4939-7765-9_11

Tesniere, A., Schlemmer, F., Boige, V., Kepp, O., Martins, I., Ghiringhelli, F., Aymeric, L., Michaud, M., Apetoh, L., Barault, L., Mendiboure, J., Pignon, J. P., Jooste, V., van Endert, P., Ducreux, M., Zitvogel, L., Piard, F., & Kroemer, G. (2010). Immunogenic death of colon cancer cells treated with oxaliplatin. *Oncogene*, *29*(4), 482–491. https://doi.org/10.1038/ONC.2009.356

The Chip Analysis Methylation Pipeline. (n.d.). Retrieved December 3, 2021, from https://www.bioconductor.org/packages/devel/bioc/vignettes/ChAMP/inst/doc/ChAMP.html

Tian, Y., Morris, T. J., Webster, A. P., Yang, Z., Beck, S., Feber, A., & Teschendorff, A. E. (2017). ChAMP: updated methylation analysis pipeline for Illumina BeadChips. *Bioinformatics*, *33*(24), 3982–3984. https://doi.org/10.1093/BIOINFORMATICS/BTX513

Tournigand, C., André, T., Achille, E., Lledo, G., Flesh, M., Mery-Mignard, D., Quinaux, E., Couteau, C., Buyse, M., Ganem, G., Landi, B., Colin, P., Louvet, C., & de Gramont, A. (2004). FOLFIRI followed by FOLFOX6 or the reverse sequence in advanced colorectal cancer: a randomized GERCOR study. *Journal of Clinical Oncology*, *22*(2), 229–237. https://doi.org/10.1200/JCO.2004.05.113

Trinh, A., Trumpi, K., de Sousa E Melo, F., Wang, X., de Jong, J. H., Fessler, E., Kuppen, P. J. K., Reimers, M. S., Swets, M., Koopman, M., Nagtegaal, I. D., Jansen, M., Hooijer, G. K. J., Offerhaus, G. J. A., Kranenburg, O., Punt, C. J., Medema, J. P., Markowetz, F., & Vermeulen, L. (2017). Practical and robust identification of molecular subtypes in colorectal cancer by immunohistochemistry. *Clinical Cancer Research*, *23*(2), 387–398. https://doi.org/10.1158/1078-0432.CCR-16-0680

Tsuji, S., Midorikawa, Y., Takahashi, T., Yagi, K., Takayama, T., Yoshida, K., Sugiyama, Y., & Aburatani, H. (2012). Potential responders to FOLFOX therapy for colorectal cancer by Random Forests analysis. *British Journal of Cancer*, *106*, 126–132. https://doi.org/10.1038/bjc.2011.505

Tsujino, T., Seshimo, I., Yamamoto, H., Chew, Y. N., Ezumi, K., Takemasa, I., Ikeda, M., Sekimoto, M., Matsuura, N., & Monden, M. (2007). Stromal myofibroblasts predict disease recurrence for colorectal cancer. *Clinical Cancer Research*, *13*(7), 2082–2090. https://doi.org/10.1158/1078-0432.CCR-06-2191

Ueno, H., Jones, A. M., Wilkinson, K. H., Jass, J. R., & Talbot, I. C. (2004). Histological categorisation of fibrotic cancer stroma in advanced rectal cancer. *Gut*, *53*(4), 581–586. https://doi.org/10.1136/GUT.2003.028365

Van Pelt, G., van Krieken, H., E M Tollenaar, R. A., Brandt Sørensen, F., van Pelt, G. W., Hansen, T. F., Bastiaannet, E., Kjaer-Frifeldt, S., Han van Krieken, J. J., AEM

Tollenaar, R., Sørensen, F. B., & Mesker, W. E. (2016). Stroma-high lymph node involvement predicts poor survival more accurately for patients with stage III colon cancer. *Journal of Medical and Surgical Pathology, 1*(2), 116-124. http://dx.doi.org/10.4172/jmsp.1000116

Vasen, H. F. A., Tomlinson, I., & Castells, A. (2015). Clinical management of hereditary colorectal cancer syndromes. *Nature Reviews Gastroenterology and Hepatology*, *12*(2), 88–97. https://doi.org/10.1038/nrgastro.2014.229

Vermeulen, L., de Sousa E Melo, F., van der Heijden, M., Cameron, K., de Jong, J. H., Borovski, T., Tuynman, J. B., Todaro, M., Merz, C., Rodermond, H., Sprick, M. R., Kemper, K., Richel, D. J., Stassi, G., & Medema, J. P. (2010). Wnt activity defines colon cancer stem cells and is regulated by the microenvironment. *Nature Cell Biology*, *12*(5), 468–476. https://doi.org/10.1038/NCB2048

Vogelstein, B., Fearon, E., Hamilton, S., Kern, S., Preisinger, A., Leppert, M., Smits, A., Bos, J. (1988). Genetic alterations during colorectal-tumor development. *New England Journal of Medicine*, *319*(9), 525–532. https://doi.org/10.1056/NEJM198809013190901

Walther, A., Houlston, R., & Tomlinson, I. (2008). Association between chromosomal instability and prognosis in colorectal cancer: a meta-analysis. *Gut*, *57*(7), 941–950. https://doi.org/10.1136/GUT.2007.135004

Wang, L., Patel, U., Ghosh, L., & Banerjee, S. (1992). DNA polymerase beta mutations in human colorectal cancer. *Cancer Research*, *52*(17), 4824–4827. https://pubmed.ncbi.nlm.nih.gov/1511447/

Wang, X. J., Li, Y., Luo, L., Wang, H., Chi, Z., Xin, A., Li, X., Wu, J., & Tang, X. (2014). Oxaliplatin activates the Keap1/Nrf2 antioxidant system conferring protection against the cytotoxicity of anticancer drugs. *Free Radical Biology & Medicine*, *70*, 68–77. https://doi.org/10.1016/J.FREERADBIOMED.2014.02.010

Wei, W., Ma, X. D., Jiang, G. M., Shi, B., Zhong, W., Sun, C. L., Zhao, L., Hou, Y. J., & Wang, H. (2020). The AKT/GSK3β-mediated Slug expression contributes to oxaliplatin resistance in colorectal cancer via upregulation of ERCC1. *Oncology Research*, *28*(4), 423–438. https://doi.org/10.3727/096504020X15877284857868

Weiser, M. R. (2018). AJCC 8th edition: colorectal cancer. *Annals of Surgical Oncology 2018 25:6*, *25*(6), 1454–1455. https://doi.org/10.1245/S10434-018-6462-1

West, N. P., Dattani, M., McShane, P., Hutchins, G., Grabsch, J., Mueller, W., Treanor, D., Quirke, P., & Grabsch, H. (2010). The proportion of tumour cells is an independent predictor for survival in colorectal cancer patients. *British Journal of Cancer 2010 102:10*, *102*(10), 1519–1523. https://doi.org/10.1038/sj.bjc.6605674

Whitmire, E., Khan, B., & Coué, M. (2002). CDC6 synthesis regulates replication competence in Xenopus oocytes. *Nature*, *419*(6908), 722–725. https://doi.org/10.1038/NATURE01032

Wilson, P. M., Danenberg, P. v., Johnston, P. G., Lenz, H. J., & Ladner, R. D. (2014). Standing the test of time: targeting thymidylate biosynthesis in cancer therapy. *Nature Reviews Clinical Oncology 2014 11:5*, *11*(5), 282–298. https://doi.org/10.1038/nrclinonc.2014.51

Wilson, R. H., Lehky, T., Thomas, R. R., Quinn, M. G., Floeter, M. K., & Grem, J. L. (2002). Acute oxaliplatin-induced peripheral nerve hyperexcitability. *Journal of Clinical Oncology*, *20*(7), 1767–1774. https://doi.org/10.1200/JCO.2002.07.056

Winterbourn, C. C. (2008). Reconciling the chemistry and biology of reactive oxygen species. *Nature Chemical Biology*, *4*(5), 278–286. https://doi.org/10.1038/NCHEMBIO.85

Woo, S. Y., & Kim, S. (2020). Determination of cutoff values for biomarkers in clinical studies. *Precision and Future Medicine*, *4*(1), 2–8. https://doi.org/10.23838/PFM.2019.00135

Woynarowski, J. M., Faivre, S., Herzig, M. C. S., Arnett, B., Chapman, W. G., Trevino, A. v., Raymond, E., Chaney, S. G., Vaisman, A., Varchenko, M., & Juniewicz, P. E. (2000). Oxaliplatin-induced damage of cellular DNA. *Molecular Pharmacology*, *58*(5), 920–927. https://doi.org/10.1124/MOL.58.5.920

Xu, K., Zhan, Y., Yuan, Z., Qiu, Y., Wang, H., Fan, G., Wang, J., Li, W., Cao, Y., Shen, X., Zhang, J., Liang, X., & Yin, P. (2019a). Hypoxia induces drug resistance in colorectal cancer through the HIF-1α/miR-338-5p/IL-6 feedback loop. *Molecular Therapy*, *27*(10), 1810–1824. https://doi.org/10.1016/j.ymthe.2019.05.017

Yamasaki, M., Kang, H. R., Homer, R. J., Chapoval, S. P., Soo, J. C., Byung, J. L., Elias, J. A., & Chun, G. L. (2008). P21 regulates TGF-beta1-induced pulmonary responses via a TNF-alpha-signaling pathway. *American Journal of Respiratory Cell and Molecular Biology*, *38*(3), 346–353. https://doi.org/10.1165/RCMB.2007-0276OC

Yang, G. Y., Xu, K. sen, Pan, Z. Q., Zhang, Z. Y., Mi, Y. T., Wang, J. S., Chen, R., & Niu, J. (2008). Integrin alpha v beta 6 mediates the potential for colon cancer cells to colonize in and metastasize to the liver. *Cancer Science*, *99*(5), 879–887. https://doi.org/10.1111/J.1349-7006.2008.00762.X

Yang, J., Parsons, J., Nicolay, N. H., Caporali, S., Harrington, C. F., Singh, R., Finch, D., Datri, S., Farmer, P. B., Johnston, P. G., McKenna, W. G., Dianov, G., & Sharma, R. A. (2010). Cells deficient in the base excision repair protein, DNA polymerase beta, are hypersensitive to oxaliplatin chemotherapy. *Oncogene*, *29*(3), 463–468. https://doi.org/10.1038/ONC.2009.327

Yin, X., Tang, B., Li, J. H., Wang, Y., Zhang, L., Xie, X. Y., Zhang, B. H., Qiu, S. J., Wu, W. Z., & Ren, Z. G. (2017). ID1 promotes hepatocellular carcinoma proliferation and confers chemoresistance to oxaliplatin by activating pentose

phosphate pathway. *Journal of Experimental and Clinical Cancer Research*, *36*(1). https://doi.org/10.1186/s13046-017-0637-7

Yin, Y., & Shen, W. H. (2008). PTEN: a new guardian of the genome. *Oncogene*, *27*(41), 5443–5453. https://doi.org/10.1038/ONC.2008.241

Yoshihara, K., Shahmoradgoli, M., Martínez, E., Vegesna, R., Kim, H., Torres-Garcia, W., Treviño, V., Shen, H., Laird, P. W., Levine, D. A., Carter, S. L., Getz, G., Stemke-Hale, K., Mills, G. B., & Verhaak, R. G. W. (2013). Inferring tumour purity and stromal and immune cell admixture from expression data. *Nature Communications 2013 4:1*, *4*(1), 1–11. https://doi.org/10.1038/ncomms3612

Yu, X., Fu, S., Lai, M., Baer, R., & Chen, J. (2006). BRCA1 ubiquitinates its phosphorylation-dependent binding partner CtIP. *Genes & Development*, *20*(13), 1721–1726. https://doi.org/10.1101/GAD.1431006

Zarei, M., Lal, S., Parker, S. J., Nevler, A., Vaziri-Gohar, A., Dukleska, K., Mambelli-Lisboa, N. C., Moffat, C., Blanco, F. F., Chand, S. N., Jimbo, M., Cozzitorto, J. A., Jiang, W., Yeo, C. J., Londin, E. R., Seifert, E. L., Metallo, C. M., Brody, J. R., & Winter, J. M. (2017). Posttranscriptional upregulation of IDH1 by HuR Establishes a powerful survival phenotype in pancreatic cancer cells. *Cancer Research*, *77*(16), 4460–4471. https://doi.org/10.1158/0008-5472.CAN-17-0015

Zhang, S., Lovejoy, K. S., Shima, J. E., Lagpacan, L. L., Shu, Y., Lapuk, A., Chen, Y., Komori, T., Gray, J. W., Chen, X., Lippard, S. J., & Giacomini, K. M. (2006). Organic cation transporters are determinants of oxaliplatin cytotoxicity. *Cancer Research*, *66*(17), 8847–8857. https://doi.org/10.1158/0008-5472.CAN-06-0769

Zhang, Y., & Molinaro, A. M. (2022). Categorizing continuous biomarkers: More cons than pros. *Neuro-Oncology Practice*, *9*(2), 81–82. https://doi.org/10.1093/NOP/NPAC012

Zhou, W., Laird, P. W., & Shen, H. (2017). Comprehensive characterization, annotation and innovative use of Infinium DNA methylation BeadChip probes. *Nucleic Acids Research*, *45*(4), e22. https://doi.org/10.1093/NAR/GKW967